Theoretical Computer Science 609 (2016) 384–394

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Extending the metric dimension to graphs with missing edges

Sabina Zejnilović [a,b,*], Dieter Mitsche [c], João Gomes [a], Bruno Sinopoli [b]

[a] Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal
[b] Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 5213, United States
[c] Université de Nice Sophia-Antipolis, Laboratoire J.-A. Dieudonné, Parc Valrose, 06108 Nice cedex 02, France

A R T I C L E   I N F O

Article history:
Received 16 March 2015
Received in revised form 4 August 2015
Accepted 18 October 2015
Available online 20 October 2015
Communicated by G. Ausiello

Keywords:
Graph theory
Metric dimension
Tree graph
Cycle graph
Grid
Graph boundary

A B S T R A C T

The metric dimension of a connected graph G is the minimum number of vertices in a subset S of the vertex set of G such that all other vertices are uniquely determined by their distances to the vertices in S. We define an extended metric dimension for graphs with some edges missing, which corresponds to the minimum number of vertices in a subset S such that all other vertices have unique distances to S in all minimally connected graphs that result from completing the original graph. This extension allows for incomplete knowledge of the underlying graph in applications such as localizing the source of infection. We give precise values for the extended metric dimension when the original graph's disconnected components are trees, cycles, grids, complete graphs, and we provide general upper bounds on this number in terms of the boundary of the graph.

1. Introduction

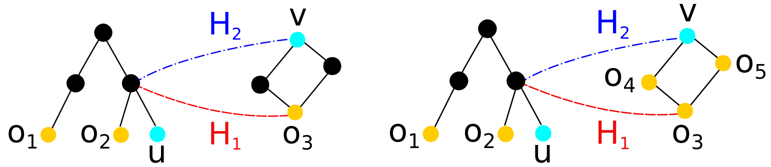
Let G be a finite, simple, connected graph with $|V(G)| = n$ vertices. For a subset $R \subseteq V(G)$ with $|R| = r$, and a vertex $v \in V(G)$, define $\boldsymbol{d}(v, R)$ to be the r-dimensional vector whose i-th coordinate $d(v, R)_i$ is the length of the shortest path between v and the i-th vertex of R. We call a set $R \subseteq V(G)$ a *resolving set* if for any pair of vertices $v, w \in V(G)$, $\boldsymbol{d}(v, R) \neq \boldsymbol{d}(w, R)$. Clearly, the entire vertex set $V(G)$ is always a resolving set, and so is $R = V(G) \setminus \{v\}$ for every vertex v. The *metric dimension* $\beta(G)$ is then the smallest cardinality of a resolving set. We have the trivial inequalities $1 \le \beta(G) \le n - 1$, with the lower bound attained for a path, and the upper bound for the complete graph. The metric dimension was introduced by Slater [10] in the mid-1970s, and by Harary and Melter [7]. As a start, Slater [10] determined the metric dimension of trees. Two decades later, Khuller, Raghavachari and Rosenfeld [9] gave a linear-time algorithm for computing the metric dimension of a tree, and characterized the graphs with metric dimensions 1 and 2. The metric dimension for many graph classes is known, including random graphs [1], and its calculation has also been extensively studied from a computational complexity point of view (see [5,6,9]).

In this paper[1] we extend the concept of metric dimension to graphs with some edges missing: suppose we are given a finite, simple graph $F = (V, E)$ with $|V| = n$ consisting of $k \ge 2$ connected components, denoted by C_i, for $i = 1, \dots k$. Denote the class $\mathcal{H}(F)$ to be the class of all possible connected graphs that can be constructed from F by adding $k - 1$

* Corresponding author at: Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisbon, Portugal.
E-mail addresses: sabinaz@cmu.edu (S. Zejnilović), dmitsche@unice.fr (D. Mitsche), jpg@isr.ist.utl.pt (J. Gomes), brunos@ece.cmu.edu (B. Sinopoli).

[1] A conference version of this paper was presented at [12].

http://dx.doi.org/10.1016/j.tcs.2015.10.022
0304-3975/



(a) Vertices o_1, o_2 and o_3 are included in the observed set O.

(b) Vertices o_1, o_2, o_3, o_4 and o_5 are included in the observed set O.

Fig. 1. An example of a partially observed network with two components. A missing edge is the one connecting the two components. In (*a*) distances of a vertex u from the set O in the graph H_1 are the same as the distances of a vertex v to the set O in the graph H_2: $d_{H_1}(u, o_1) = 4 = d_{H_2}(v, o_1)$, $d_{H_1}(u, o_2) = 2 = d_{H_2}(v, o_2)$ and $d_{H_1}(u, o_3) = 2 = d_{H_2}(v, o_3)$. Without knowing if the true graph is H_1 or H_2 the source cannot be correctly identified, as it can be either vertex u or v. In (*b*), two more vertices are included in set O. It can be checked that O is now a minimum cardinality extended resolving set. Now, the distances of the vertices u and v to the set O are different, as $d_{H_1}(u, o_4) = 3 \neq 1 = d_{H_2}(v, o_4)$ and $d_{H_1}(u, o_5) = 3 \neq 1 = d_{H_2}(v, o_5)$. Hence, the vertices u and v can be distinguished and the source can be unambiguously localized, even if it is not known exactly how the two components are connected.

edges. For a graph $H_1 \in \mathcal{H}(F)$, a vertex $u \in V$ and a set $O \subseteq V$, denote by $\boldsymbol{d}_{H_1}(u, O)$ the distance vector of u to the set O in the graph H_1, that is, $(d_{H_1}(u, O))_i$ is the length of the shortest path between u and the i-th vertex of O in the graph H_1. A set of vertices $O \subseteq V(F)$ such that for any two different vertices u and v, and any two graphs $H_1, H_2 \in \mathcal{H}(F)$, $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ is called an *extended resolving set* of F. The cardinality of a smallest extended resolving set of a graph F, denoted by $\gamma(F)$, is the *extended metric dimension* of F. Note that $\max_{H_i \in \mathcal{H}(F)} \beta(H_i) \leq \gamma(F) \leq n - 1$.

Motivation. The introduction of resolving sets by Slater [10] was motivated by the application of placement of a minimum number of sonar detectors in a network, while Khuller, Raghavachari and Rosenfeld [9] were interested in finding the minimum number of landmarks needed for robot navigation on a graph. Recently, the problem of finding the minimum number of agents whose infection times need to be observed in order to identify the first infected agent for a simplified diffusion model was cast as finding the metric dimension of the graph [11]. Similarly, to identify a rumor source in a network based on the times when the nodes first heard the rumor, observed nodes should form a resolving set.

However, in many practical applications, the network topology is not completely known, and only locally can the network be completely observed. For example, one wants to uniquely identify a source in a network possibly far away, but information about the presence/non-presence of edges is missing. More precisely, we want to find a subset of the vertices, from which we can identify a source uniquely, even when we only know that the graph has *some* edge connecting two (possibly far) components, and without knowing which edge it is. Hence, just by observing the distances between the nodes, and without knowing exactly how local components are connected, we wish to always unambiguously identify the source. An illustrative example is shown in Fig. 1.

We model incomplete network knowledge by assuming that the graph of interest is disconnected, with k components and $k - 1$ unobserved edges connecting the components, and we consequently introduce the concept of extended metric dimension. We are aware that our model is restrictive and is only a first step towards incomplete knowledge of the graph topology. A more general model, allowing the addition of more than $k - 1$ edges, and not necessarily only a spanning tree between the original components, is object of further research.

A similar, but different, approach was recently undertaken by [4]: their way of modeling incomplete information is the following: they call a set S doubly-resolving, if for any two vertices u, v there exist $x, y \in S$ such that $d(u, x) - d(u, y) \neq d(v, x) - d(v, y)$, and their goal is to find a doubly-resolving set of minimal cardinality. The motivation for the work [4] also stems from the application of source localization, but with the difference that the original activation time of the source is not known, while the graph structure is fully known.

Notation. For a connected graph G, $i, j \in V(G)$, denote an $i - j$-path to be a sequence of all different vertices $v_0 = i$, $v_1, \ldots, v_\ell = j$, such that for $i = 0, \ldots, \ell - 1$, $\{v_i, v_{i+1}\} \in E(G)$. Let $L(C_i)$ denote the set of all leaves of component C_i. Let $K(C_i)$ be the set of vertices of component C_i that have degree greater than two, and that are connected by paths of degree-two vertices to one or more leaves in C_i (when considering C_i as a separate graph and ignoring edges to other components). For a given vertex $c \in K(C_i)$, call the leaves connected to c via such degree-two-paths to be the *associated leaves* of c. Note that for a tree that is not a path each leaf is associated to exactly one vertex $c \in K(C_i)$. For a fixed component C_i of F, denote by S_i a minimum cardinality resolving set of C_i (so that $\beta(C_i) = |S_i|$). The $M \times N$-grid with $M, N \geq 2$, is the graph whose vertices correspond to the points in the plane with integer coordinates, x-coordinates being in the range $0, \ldots, M - 1$, y-coordinates in the range $0, \ldots, N - 1$, and two vertices are connected by an edge whenever the corresponding points are at Euclidean distance 1. The four vertices of degree two are called corner vertices.

For a connected graph G, a vertex v is a *boundary vertex* of u if $d_G(w, u) \leq d_G(v, u)$, for all w that are neighbors of v [3]. A vertex v is a boundary vertex of G if it is a boundary vertex of some vertex of G. The set of all boundary vertices of a vertex u is denoted as $\partial(u)$. The boundary of a vertex set $S \subseteq V$ is the set of vertices in G that are boundary vertices for some vertex $u \in S$. The *boundary* of graph G, $\partial(G)$, is the set of all boundary vertices of G. It is well known that the boundary is a resolving set, see [8]. For example, the boundary of a tree is the set of its leaves, whereas the boundary of a grid is the set of its 4 corner vertices, and the boundary of a cycle is the whole vertex set [8].

Statements of results. We state the main results of this paper which are then proved in the following sections.

Theorem 1.1. *Let F be a graph of k components, where each component is a tree. Then $\gamma(F) = \min_j \sum_{i=1, i\neq j}^{k} |L(C_i)| + |S_j|$, unless all components are isolated vertices, in which case $\gamma(F) = k-1$. In the first case, we may assume without loss of generality, that the minimum is attained for $j = k$. Then the set consisting of all leaves from components $1, \dots, k-1$ together with a minimum cardinality resolving set of the k-th component is a minimum cardinality extended resolving set of the graph F.*

Theorem 1.2. *Let F be a graph of k components, where each component is a complete graph. Let $\mathcal{I}_1$ denote the set of indices of components that are isolated vertices, $\mathcal{I}_2$ the set of indices of components that have only 2 vertices, and $\mathcal{I}_3$ the set of indices of components that contain at least 3 vertices. If $\mathcal{I}_1$ and $\mathcal{I}_2$ are empty, then $\gamma(F) = n - k$ and the set consisting of all but one vertex of each component is a minimum cardinality extended resolving set of the graph F. Otherwise, $\gamma(F) = \sum_{i\in\mathcal{I}_3}(|C_i| - 1) + 2|\mathcal{I}_2| + |\mathcal{I}_1| - 1$ (note that I_3 might be empty, in which case the contribution of the preceding sum over I_3 is zero). The set consisting of all but one vertex from each component of at least size 3 and all but one vertex from the components of sizes 1 or 2 is a minimum cardinality extended resolving set of the graph F.*

Theorem 1.3. *Let F be a graph of k components, where each component is a grid. Then $\gamma(F) = 3k - 1$. Let $O_i = \{r_1^i, r_2^i, r_3^i\}$ denote a set of three corner vertices from component C_i. Then $O = \cup_{i=1}^{k-1} O_i \cup S_k$ is a minimum cardinality extended resolving set of F.*

Theorem 1.4. *Let F be a graph of k components, where each component is a cycle of size greater than 3. Let k_e denote the number of components with an even number of vertices. Then $\gamma(F) = 2k + k_e - 1$, if $k_e > 0$, and $\gamma(F) = 2k$, otherwise. For a component C_i with an even number of vertices n_i, define $O_i = \{r_1^i, r_2^i, r_3^i\}$, where r_1^i, r_2^i are two neighboring vertices in C_i and r_3^i is a vertex at distance at least $\frac{n_i-2}{2}$ from both of them, also in C_i. For a component C_i with an odd number of vertices n_i, define $O_i = \{r_1^i, r_2^i\}$, where r_1^i and r_2^i are two vertices of C_i that are at distance $\frac{n_i-1}{2}$ from each other. If $k_e = 0$, $O = \cup_{i=1}^{k} O_i$ is a minimum cardinality extended resolving set of F. If $k_e > 0$, assume without loss of generality that C_k is a component with an even number of vertices. Then $O = \cup_{i=1}^{k-1} O_i \cup S_k$ is a minimum cardinality extended resolving set of F.*

For general graph classes we have the following results, the second one tightening the first one, as the boundary of a graph can be very large.

Theorem 1.5. *For any arbitrary graph F with k connected components, the set $O = \cup_{i=1}^{k-1} \partial(C_i) \cup S_k$ is an extended resolving set for F.*

Theorem 1.6. *Let F be an arbitrary graph with k connected components, let S_i be a resolving set of C_i, and let $O_i = S_i \cup \partial(S_i)$. Then $O = \cup_{i=1}^{k-1} O_i \cup S_k$ is an extended resolving set for F.*

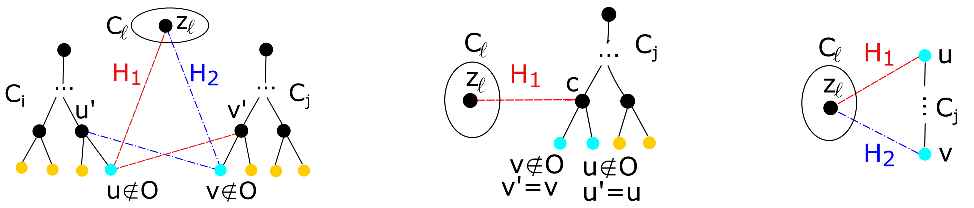
2. Proofs of main results for special graph classes

Proof of Theorem 1.1. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. We need to show that the set $O = \cup_{i=1}^{k-1} L(C_i) \cup S_k$ is a set of smallest cardinality for which $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ holds when all components are trees, unless all components are isolated vertices, in which case $O = \cup_{i=1}^{k-1} C_i$.

We first prove the claim of sufficiency. If both u and v are any two vertices in the same component, then u and v are distinguishable as the set of all the leaves of a tree is a resolving set. Hence we may assume $u \in V(C_i)$ and $v \in V(C_j)$ for $i \neq j$. We may also assume without loss of generality that $i < k$. Let p be the vertex in C_i and q the vertex in C_j, such that any path from a vertex in C_i to any vertex in C_j in H_2 contains a subpath $p - q$. Note that $d_{H_2}(p, q) \geq 1$. If u is a leaf, as it is contained in $L(C_i)$, it is distinguishable from v, since $0 = d_{H_1}(u, u) < d_{H_2}(u, v)$. If u is not a leaf, and $u = p$, then for any leaf $r \in L(C_i)$, $d_{H_2}(r, v) = d_{H_2}(r, p) + d_{H_2}(p, q) + d_{H_2}(q, v) \geq d_{H_1}(r, p) + d_{H_2}(p, q) > d_{H_1}(r, p)$. Thus, the two distance vectors are not equal either. Otherwise, if u is not a leaf, and $u \neq p$, let r be a leaf in $L(C_i)$ such that u is on the path from r to p (such a leaf clearly exists). Then $d_{H_2}(r, v) = d_{H_2}(r, u) + d_{H_2}(u, p) + d_{H_2}(p, q) + d_{H_2}(q, v) > d_{H_1}(r, u) + d_{H_1}(u, p) > d_{H_1}(r, u)$. Thus, the two distance vectors also in this case are not equal, which completes the proof of sufficiency.

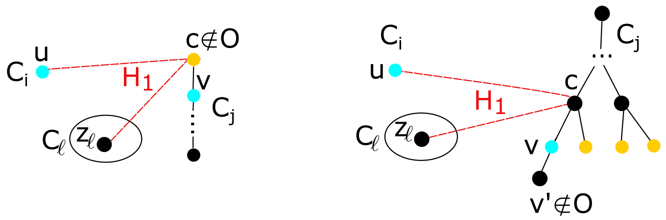
Now, we prove the claim of necessity. Let O be an arbitrary extended resolving set. We will show that O has to be at least of the size given by the sufficient condition.

Case I: Let C_i and C_j be two components with at least 2 vertices, such that both have a leaf which is not included in O. Let u be such a leaf in component C_i with neighbor u' and v be a leaf in C_j with neighbor v', such that $u, v \notin O$. We claim that u and v are indistinguishable, as illustrated in Fig. 2a. We can construct H_1 by connecting u with v', and u with some vertex z_ℓ of every other component C_ℓ (if there are more than 2 components). H_2 is then constructed by connecting v with u' and v with the same vertex z_ℓ for every C_ℓ with $\ell \notin \{i, j\}$ as in H_1; the other newly added edges are the same in H_1 and H_2 (and not involving either C_i nor C_j). Now, we have $\boldsymbol{d}_{H_1}(u, O) = \boldsymbol{d}_{H_2}(v, O)$, as follows. For any vertex $r \in C_i \setminus \{u\}$, we have $d_{H_1}(u, r) = 1 + d_{H_1}(u', r)$, and $d_{H_2}(v, r) = d_{H_2}(u', r) + 1 = d_{H_1}(u', r) + 1$. For any vertex $r \in C_j \setminus \{v\}$, we have $d_{H_1}(u, r) = d_{H_1}(v', r) + 1$, and $d_{H_2}(v, r) = d_{H_2}(v', r) + 1 = d_{H_1}(v', r) + 1$. Finally, for a vertex $r \in C_\ell$, $\ell \neq i, j$, we have



(a) Both components have a leaf which is not included in O. (b) All leaves of C_i are in O and C_j is not a path. (c) All leaves of C_i are in O and C_j is a path.

Fig. 2. Case I in the Proof of Theorem 1.1: Constructing trees H_1 and H_2 when both components C_i and C_j have at least two vertices.



(a) C_j is a path. (b) C_j is not a path.

Fig. 3. Case II in the Proof of Theorem 1.1: Constructing trees H_1 and H_2 when component C_i has only one vertex.

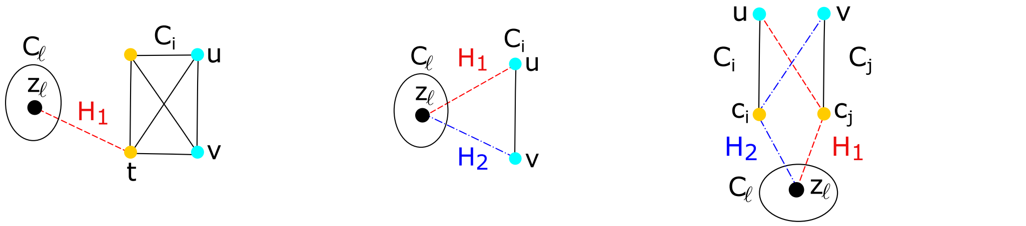
$d_{H_1}(u,r) = 1 + d_{H_1}(z_\ell, r) = 1 + d_{H_2}(z_\ell, r) = d_{H_2}(v,r)$. Thus the vertices u and v are indistinguishable, and the claim holds. Hence, either all the leaves of component C_i or component C_j have to be included in O. Without loss of generality, let us assume that all the leaves of C_i are included in O. Now we assume that only $|S_j| - 1$ vertices are selected from the component C_j. In the first sub-case, when C_j is not a path, from [2], we have $|S_j| = |L(C_j)| - |K(C_j)|$. If only $|S_j| - 1$ vertices were taken from C_j, then there exists a vertex c in $K(C_j)$ with two associated leaves u and v, such that no vertex from the paths $c - u$ nor $c - v$ is included in O. But then there exist a vertex u' on the path $c - u$, and a vertex v' on the path $c - v$, such that $d_{C_j}(u', c) = d_{C_j}(v', c)$. Note that u' might coincide with u, and v' might coincide with v. The vertices u' and v' are indistinguishable from each other in C_j. Constructing a tree H_1 by connecting any vertex z_ℓ from every other component C_ℓ, $\ell \neq j$, with any fixed vertex in $K(C_j)$, we see that u' and v' still are indistinguishable by vertices in O, as shown in Fig. 2b. In the second sub-case, when C_j is a path with leaves u and v, S_j comprises only one leaf. If no vertex from C_j is in O, H_1 can be constructed by connecting one of its leaves u with some vertex z_ℓ of every other component C_ℓ, while H_2 is constructed by connecting z_ℓ to the other leaf v, and vertices u and v are indistinguishable, as Fig. 2c shows. Thus, at least $|S_j|$ vertices have to be taken from C_j.

Case II: C_i consists of only one vertex, u, and C_j has at least 2 vertices. By the same arguments as in Case I, it can be seen that at least $|S_j|$ vertices from component C_j have to be included in O. We will show now that u has to be included in O as well. In the first sub-case, when C_j is a path, H_1 can be constructed by connecting u with the leaf c of C_j where $c \in O$, and then connecting c to a vertex z_ℓ of every other component C_ℓ, $\ell \neq i, j$. Let v be the vertex in C_j which is the neighbor of c. If u is not chosen, u is indistinguishable within H_1 from v, as can be seen in Fig. 3a. As for the second sub-case, when C_j is not a path, let c be a vertex in $K(C_j)$ such that the path to its associated leaf v' contains no vertices from O. Then H_1 is constructed by connecting u with c, and then connecting c to a vertex z_ℓ of every other component C_ℓ, $\ell \neq i, j$. Let v be the neighbor of c in C_j which lies on the path $c - v'$. Note that v can coincide with v'. Then u is indistinguishable within H_1 from v, as shown in Fig. 3b. Hence, u must also be included in O.

Case III: Both C_i and C_j contain only one vertex. Call these u and v, respectively. At least one of them has to be included in O: otherwise, we can construct H_1 by connecting both u and v to some vertex z_ℓ from every other component C_ℓ, $\ell \neq i, j$, and then u and v are indistinguishable within H_1. If there are only two components, each with one vertex, H_1 is constructed by connecting them. Clearly, if neither vertex is included in O, the set O is empty, and the two vertices are indistinguishable within H_1.

Therefore, for any pair of components C_i and C_j, an extended resolving set O has to include all leaves from one component and a resolving set from the other, unless both have size 1, in which case only 1 vertex is enough. Hence, if there exists at least one component which has 2 or more vertices, from all but one component all the leaves have to be taken, and from the remaining component, at least a resolving set. If all k components have only one vertex, the set O has to contain $k - 1$ vertices. □

Proof of Theorem 1.2. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. First we need to show that the set O consisting of all but one vertex from each component is a set of smallest cardinality for which $\mathbf{d}_{H_1}(u, O) \neq \mathbf{d}_{H_2}(v, O)$ holds when all the components are complete graphs with at least 3 vertices.



(a) C_i is of size at least 3, with two vertices that are not included in O.

(b) C_i is of size 2 and neither vertex is included in O.

(c) Both components, C_i and C_j, are of size 2 and each have only one node included in O.

Fig. 4. Proof of Theorem 1.2: Constructing trees H_1 and H_2 when both components C_i and C_j are complete graphs.

First, we prove the claim of sufficiency. Let us denote the set of all but one vertex on component C_i by O_i. If u and v are in the same component, they are distinguishable, since each O_i is a resolving set of component C_i [9]. Hence, let us assume that vertex $u \in V(C_i)$ is not included in O_i, and that vertex $v \in V(C_j)$ is not included in O_j, and $i \neq j$. Let $p \in V(C_i)$ and $q \in V(C_j)$, such that $p - q$ is contained in every shortest path between vertices of components C_i and C_j in H_2. Note again that $d_{H_2}(p,q) \geq 1$. We prove the claim by contradiction and assume that the following relations hold:

$$d_{H_1}(u,r) = 1 = d_{H_2}(v,r) = d_{H_2}(v,q) + d_{H_2}(q,p) + d_{H_2}(p,r), \tag{1}$$

for every $r \in O_i$. Then for $d_{H_2}(p,q) = 1$, the condition $d_{H_2}(p,r) = 0$ would have to hold, while for $d_{H_2}(p,q) > 1$, the condition $d_{H_2}(p,r) < 0$ would have to hold for all $r \in O_i$. In either case, this is not possible, and the claim is proved.

To prove the claim of necessity, let O be an arbitrary extended resolving set. Assume that in one component C_i there are two vertices, u and v, that are not included in O. We construct H_1 by adding the edges between a fixed vertex $t \in V(C_i) \setminus \{u, v\}$ and some fixed vertex z_ℓ of some other component C_ℓ. Then we have $d_{H_1}(u,r) = d_{H_1}(v,r)$ for all $r \in O$, and this completes the proof.

Now, we consider the case when there is at least one component of size 1 or 2. We need to show that a minimum cardinality extended resolving set O should contain all but one vertex of each component of size at least 3 and all but one vertex from the components of sizes 1 or 2. First, we prove the claim of sufficiency. If either u or v is included in O, w.l.o.g. say u, then u and v are clearly distinguishable, as, $d_{H_1}(u,u) = 0 \neq d_{H_2}(u,v)$, for any $H_1, H_2 \in (F)$. Therefore, let $u \notin O$ be a vertex belonging to component C_i, and let $v \notin O$ be a vertex belonging to component C_j. Let $p \in V(C_i)$ and $q \in V(C_j)$, such that the path $p - q$ is contained in every shortest path between vertices from C_i and C_j in H_2. Note that $d_{H_2}(p,q) \geq 1$. Since only one vertex from a component of sizes 1 or 2 is not included in O, let us assume u is in a component of size at least 3. Then by the same argument as in (1) and the following discussion, we must have $d_{H_1}(u,r) \neq d_{H_2}(v,r)$ for some $r \in O$.

To prove the claim of necessity, let O be an arbitrary extended resolving set. We will show that O has to be at least of the size given by the sufficient condition.

Case I: Let C_i be a component of size at least 3, with two vertices, u and v, that are not included in O. This is exactly the case discussed when there are only components of size at least 3 and it is shown in Fig. 4a. Thus, all but one vertex from each component of size at least 3 has to be included in O.

Case II: Let C_i be a component of size 2, where neither vertex is included in O. Let u, v be the two vertices from the component C_i. We can construct H_1 by connecting u with any vertex $z_\ell \in V(C_\ell)$, and H_2 by connecting v again to the same vertex $z_\ell \in V(C_\ell)$, for any $\ell \neq i$. In both H_1 and H_2 all other connections between components are the same and not including C_i. Then, we have that u and v cannot be distinguished, as illustrated in Fig. 4b. Hence at least one vertex from a component of size 2 has to be included in O.

Case III: Let C_i be a component of size 2 with one vertex, u, that is not included in O. First, let us consider the sub-case when C_j is a component of size 2, and vertex $v \in V(C_j)$ is a vertex not included in O. Then, H_1 can be constructed by connecting u with the vertex $c_j \in C_j$, where $c_j \in O$, and then connecting c_j to some vertex z_ℓ of every other component C_ℓ. H_2 can be constructed by connecting v with the vertex $c_i \in C_i$, where $c_i \in O$, and then connecting c_i to the same vertex z_ℓ of every other component C_ℓ. All the other connections in H_1 and H_2 are the same and not including either C_i or C_j. Now we have that $d_{H_1}(u,r) = d_{H_2}(v,r)$ for all $r \in O$, and u and v are indistinguishable, as Fig. 4c shows. Next, we consider the sub-case when C_j is an isolated vertex, not included in O. We can construct H_1 by connecting $c_i \in V(C_i)$, where $c_i \in O$ to the isolated vertex from C_j. Then, u and the vertex from C_j are indistinguishable. Hence, all but one vertex from the components of size 2 have to be included in O.

Case IV: Let both C_i and C_j be isolated vertices. Note that this is equivalent to Case III of the Proof of Theorem 1.1. Now all but one vertex from the components of sizes 1 have to be included in O.

Hence, the set O should contain all but one vertex on each component of at least size 3 and all but one vertex from the components of sizes 1 or 2. □

Proof of Theorem 1.3. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. We need to show that the set O comprising three corner vertices from $k-1$ components and a resolving set of the k-th component is a set of smallest cardinality for which $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ holds when all the components are grids.

Let us denote the size of the grid C_i as $x_i \times y_i$. We assume that each vertex $l \in V(C_i)$ has assigned to it a position vector (x_l, y_l) which represents its location on the integer lattice C_i, with the first selected corner vertex r_1^i at position $(0,0)$, r_2^i at $(x_i, 0)$ and r_3^i at $(0, y_i)$. First, let us prove the claim of sufficiency. If u and v are in the same component, they are distinguishable, since any two corner vertices having the same value in one coordinate form a resolving set of a grid [9]. Hence, let us assume that $u \in V(C_i)$ and $v \in V(C_j)$, for $i \neq j$ and $i < k$. Let p be the vertex in C_i and q the vertex in C_j, such that any path from a vertex in C_i to any vertex in C_j in H_2 contains a subpath $p - q$, with $d_{H_2}(p,q) \geq 1$. If $u = p$, then for all $r \in O_i$ we have $d_{H_2}(v,r) = d_{H_2}(r,p) + d_{H_2}(p,q) + d_{H_2}(q,v) > d_{H_2}(r,p) = d_{H_1}(r,u)$. Therefore u and v are distinguishable. For $u \neq p$, let us prove the claim by contradiction. Assuming $\boldsymbol{d}_{H_1}(u, O_i) = \boldsymbol{d}_{H_2}(v, O_i)$, we obtain the following equations:

$$
\begin{aligned}
d_{H_1}(u, r_1^i) &= x_u + y_u \\
&= d_{H_2}(v, r_1^i) = x_p + y_p + d_{H_2}(p,q) + d_{H_2}(q,v) \\
d_{H_1}(u, r_2^i) &= x_i - x_u + y_u \\
&= d_{H_2}(v, r_2^i) = x_i - x_p + y_p + d_{H_2}(p,q) + d_{H_2}(q,v) \\
d_{H_1}(u, r_3^i) &= x_u + y_i - y_u \\
&= d_{H_2}(v, r_3^i) = x_p + y_i - y_p + d_{H_2}(p,q) + d_{H_2}(q,v).
\end{aligned}
\tag{2}
$$

The system of equations (2) can be rewritten in matrix form

$$A\alpha = b,$$

where

$$A = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix}, \quad \alpha = \begin{bmatrix} x_u \\ y_u \\ d_{H_2}(p,q) + d_{H_2}(q,v) \end{bmatrix}, \quad b = \begin{bmatrix} x_p + y_p \\ -x_p + y_p \\ x_p - y_p \end{bmatrix}.$$

A is a matrix of full rank, implying that the system of equations (2) has a unique solution, given by $A^{-1}b$. The only solution is $x_u = x_p$, $y_u = y_p$, and $d_{H_2}(p,q) + d_{H_2}(q,v) = 0$, contradicting $d_{H_2}(p,q) \geq 1$. The set $\cup_{i=1}^{k-1} O_i \cup S_k$ is a set of cardinality $3k-1$, and this completes the sufficiency claim.

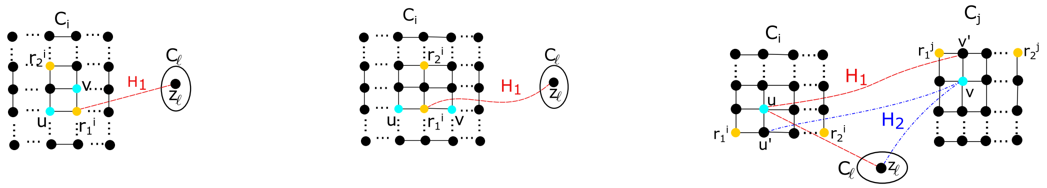
For the claim of necessity, let us assume that there exist two components C_i and C_j, such that from each of them, only two vertices are chosen. Let $\{r_1^i, r_2^i\}$ be the set of two vertices from C_i and let $\{r_1^j, r_2^j\}$ be the set of two vertices from C_j that are included in O.

Case I: In at least one component, the vertices included in O are not two corner vertices with one identical coordinate. Let us assume that this is the case with C_i. We claim that there exist two vertices u and v in C_i which are indistinguishable by r_1^i and r_2^i. Denote by $(x_{r_1^i}, y_{r_1^i})$ and by $(x_{r_2^i}, y_{r_2^i})$ the positions at which r_1^i and r_2^i are located in the grid.

First, let us consider the sub-case, when r_1^i and r_2^i differ in both coordinates, as shown in Fig. 5a. Without loss of generality, let us assume that $y_{r_1^i} < y_{r_2^i}$. Let $\Delta_x = |x_{r_2^i} - x_{r_1^i}|$ and $\Delta_y = y_{r_2^i} - y_{r_1^i}$. Let u be the vertex at $(x_{r_2^i}, y_{r_1^i})$. For $\Delta_x \leq \Delta_y$, let v be the vertex at $(x_{r_1^i}, y_{r_1^i} + \Delta_x)$, while for $\Delta_x > \Delta_y$, we consider two possible cases. For $x_{r_1^i} < x_{r_2^i}$, let v be the vertex at $(x_{r_1^i} + \Delta_x - \Delta_y, y_{r_2^i})$, and for $x_{r_1^i} > x_{r_2^i}$, let v be the vertex at $(x_{r_2^i} + \Delta_y, y_{r_2^i})$. Constructing a tree H_1 by connecting some vertex z_ℓ from every other component C_ℓ, $\ell \neq i$, with either r_1^i or r_2^i, we have $d_{H_1}(u, r_1^i) = \Delta_x = d_{H_1}(v, r_1^i)$ and $d_{H_1}(u, r_2^i) = \Delta_y = d_{H_1}(v, r_2^i)$. Hence the vertices u and v are indistinguishable by any vertex in O.

In the second sub-case, r_1^i and r_2^i differ in only one coordinate, as Fig. 5b illustrates. Then, let u and v be two neighbors of r_1^i, which are not on the shortest path $r_1^i - r_2^i$. These two vertices exist, as all vertices on the grid, except the corner vertices, have at least 3 neighbors. Now, we have $d_{C_i}(u, r_1^i) = 1 = d_{C_i}(v, r_1^i)$ and $d_{C_i}(u, r_2^i) = 1 + d_{C_i}(r_1^i, r_2^i) = d_{C_i}(v, r_2^i)$. Therefore, there always exist two vertices u and v, such that they are not distinguishable by any two vertices of C_i which are not two corner vertices with one identical coordinate. Constructing a tree H_1 by connecting some vertex z_ℓ from every other component C_ℓ, $\ell \neq i$, with either r_1^i or r_2^i, we see that u and v still are indistinguishable by any vertex in O.

Case II: From both components C_i and C_j, two corner vertices with one identical coordinate are included in O. Let u' be a vertex on C_i that is a neighbor of r_1^i such that it shares one coordinate with both r_1^i and r_2^i. Then let u be a neighbor of u' such that it does not share any coordinates with r_1^i. Similarly, let v' be a vertex in C_j that is a neighbor of r_1^j such that it shares one coordinate with both r_1^j and r_2^j. Then let v be a neighbor of v' such that it does not share any coordinates with r_1^j. We can construct H_1 by connecting u with v' and u with some vertex z_ℓ of every other component C_ℓ (if there



(a) The vertices r_1^i and r_2^i are not corner vertices and they differ in both coordinates.

(b) The vertices r_1^i and r_2^i are not corner vertices and they differ in one coordinate.

(c) The vertices r_1^i and r_2^i are two corner vertices with one identical coordinate.

Fig. 5. Proof of Theorem 1.3: Constructing H_1 and H_2 when the components are grids.

are more than 2 components). Then H_2 is constructed by connecting v with u' and v with the same vertex z_ℓ as in H_1, as shown in Fig. 5c. The distances of u and v from the vertices in O are

$$d_{H_1}(u, r_1^i) = d_{H_2}(v, r_1^i) = 2$$
$$d_{H_1}(u, r_2^i) = d_{H_2}(v, r_2^i) = 1 + d_{H_1}(u', r_2^i)$$
$$d_{H_1}(u, r_1^j) = d_{H_2}(v, r_1^j) = 2$$
$$d_{H_1}(u, r_2^j) = d_{H_2}(v, r_2^j) = 1 + d_{H_2}(v', r_2^j)$$
$$d_{H_1}(u, r) = d_{H_2}(v, r) = 1 + d_{H_1}(z_\ell, r),$$

for $r \in C_\ell$, $\ell \neq i, \ell \neq j$. Hence the vertices u and v are indistinguishable.

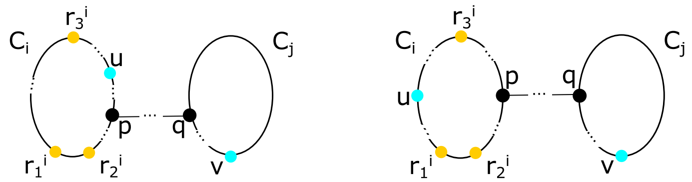
Therefore, at least 3 vertices of component C_i or component C_j have to be included in O. Without loss of generality, let us assume that 3 vertices in C_i are included in O. Now we assume that only $|S_j| - 1 = 1$ vertex is selected from C_j. Then there exist two vertices u and v in component C_j, which are at the same distance from the only vertex r included from S_j. We construct H_1 by connecting some vertex z_ℓ from every other component C_ℓ to vertex r in component C_j. Observe that the vertices u and v are still not distinguishable within H_1, and hence at least $|S_j| = 2$ vertices have to be included from component C_j. In conclusion, for any two components, at least 3 vertices from one and 2 vertices from the other one have to be included in O, and thus $|O| \geq 3(k-1) + 2 = 3k - 1$. □

Proof of Theorem 1.4. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. We need to show that the set O as defined in the statement of the theorem with cardinality equal to $2k + k_e - 1$ if the number of components with an even number of vertices, k_e is non-zero, or $2k$, otherwise, is a set of smallest cardinality for which $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ holds when all components are cycles.

First, let us prove the claim of sufficiency. As in Theorem 1.3, let us assume that vertex u is located in component C_i and vertex v is in component C_j (when u and v belong to the same component, they are clearly distinguishable, as any two neighboring vertices of an even cycle and any two vertices at distance $(n_i - 1)/2$ in the case of an odd cycle C_i form a resolving set of a cycle). Let $u \in V(C_i)$, $v \in V(C_j)$, with $i \neq j$ and $i < k$. Let p be the vertex in C_i and q the vertex in C_j, such that any path from a vertex in C_i to any vertex in C_j in H_2 contains a subpath $p - q$, with $d_{H_2}(p, q) \geq 1$. If the vertices u and v are not distinguishable by O_i, then $d_{H_1}(u, r) = d_{H_2}(v, r) = d_{H_1}(p, r) + d_{H_2}(p, q) + d_{H_2}(q, v)$ holds for some H_1 and H_2 and all $r \in O_i$. Therefore, the following must hold

$$d_{H_1}(u, r) > d_{H_1}(p, r). \tag{3}$$

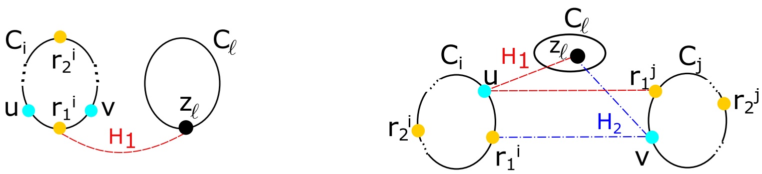
Case I: Both components C_i and C_j have an even number of vertices. For a component C_i, due to the placement of r_3^i, the distance $d_H(r_1^i, r_3^i) \in \left\{\frac{n_i - 2}{2}, \frac{n_i}{2}\right\}$, for any $H \in \mathcal{H}(F)$. The same holds for $d_H(r_2^i, r_3^i)$. Let us first consider the sub-case where both p and u lie in the same half of the cycle, i.e., both lie either on the shorter path $r_2^i - r_3^i$ or on the shorter path $r_1^i - r_3^i$, as shown in Fig. 6a. Suppose without loss of generality that they both lie on the shorter path $r_2^i - r_3^i$. As one of the vertices out of $\{u, p\}$ is closer to r_3^i and the other one is closer to r_2^i, (3) cannot hold simultaneously for both r_2^i and r_3^i. The other sub-case that needs to be considered is when u and p lie in different semi-cycles, one on the shorter path $r_2^i - r_3^i$, and the other on the shorter path $r_1^i - r_3^i$, as illustrated in Fig. 6b. Notice that for any vertex w that is on the shorter path $r_1^i - r_3^i$, the distance $d_H(w, r_2^i) = \min\{d_H(w, r_1^i) + 1, d_H(w, r_3^i) + d_H(r_3^i, r_2^i)\}$. We have $d_H(w, r_1^i) + 1 \leq \frac{n_i}{2} - 1 + 1 = \frac{n_i}{2}$, and also $d_H(w, r_3^i) + d_H(r_3^i, r_2^i) \geq 1 + \frac{n_i - 2}{2} = \frac{n_i}{2}$. Therefore, the first term of the minimum can never be larger than the second term, hence we can write $d_H(w, r_2^i) = d_H(w, r_1^i) + 1$. The same reasoning holds when w is on the shorter path $r_2^i - r_3^i$, and then we can write $d_H(w, r_1^i) = d_H(w, r_2^i) + 1$. Now, for vertices u and p that lie in different semi-cycles, we either have $d_{H_1}(u, r_1^i) = d_{H_1}(u, r_2^i) + 1$ and $d_{H_1}(p, r_1^i) = d_{H_1}(p, r_2^i) - 1$, or $d_{H_1}(u, r_1^i) = d_{H_1}(u, r_2^i) - 1$ and $d_{H_1}(p, r_1^i) = d_{H_1}(p, r_2^i) + 1$. If



(a) The vertices p and u lie in the same semi-cycle.

(b) The vertices p and u lie in different semi-cycles.

Fig. 6. Case I in the Proof of Theorem 1.4: Both cycle components have an even number of vertices.



(a) The vertices r_1^i and r_2^i are exactly at distance $\frac{n_i}{2}$.

(b) The vertices r_1^i and r_2^i are not at distance $\frac{n_i}{2}$.

Fig. 7. Proof of Theorem 1.4: Constructing H_1 and H_2 when the components are cycles.

now (3) holds and u and v are indistinguishable by r_1^i and r_2^i, then we would have to have

$$d_{H_2}(p,q)+d_{H_2}(q,v)=d_{H_1}(u,r_1^i)-d_{H_1}(p,r_1^i)=d_{H_1}(u,r_2^i)-d_{H_1}(p,r_2^i).$$

If now $d_{H_1}(u,r_1^i)=d_{H_1}(u,r_2^i)+1$ and $d_{H_1}(p,r_1^i)=d_{H_1}(p,r_2^i)-1$ holds, then we would have to have

$$d_{H_1}(u,r_1^i)-d_{H_1}(p,r_1^i)=d_{H_1}(u,r_2^i)-1-d_{H_1}(p,r_2^i)-1,$$

which is clearly not possible. If $d_{H_1}(u,r_1^i)=d_{H_1}(u,r_2^i)-1$ and $d_{H_1}(p,r_1^i)=d_{H_1}(p,r_2^i)+1$ holds, the analogous contradiction appears, and hence in both cases (3) cannot hold.

Case II: At least one of the components C_i or C_j has an odd number of vertices. Let us assume that this is the case with C_i. Similarly, as in Case I, let us first consider the sub-case where both p and u lie in the same half of the cycle, i.e. both on the shorter path $r_1^i - r_2^i$ or both on the longer path $r_1^i - r_2^i$. As before, one of the vertices out of $\{u,p\}$ is closer to r_1^i, and the other is closer to r_2^i, and thus (3) cannot hold simultaneously for both r_1^i and r_2^i. The other sub-case that needs to be considered is when u and p lie in different semi-cycles, one on the shorter path $r_1^i - r_2^i$, of length $\frac{n_i-1}{2}$, and the other on the longer path $r_1^i - r_2^i$, of length $\frac{n_i+1}{2}$. Then either we have $d_{H_1}(u,r_2^i)=\frac{n_i-1}{2}-d_{H_1}(u,r_1^i)$ and $d_{H_1}(p,r_2^i)=\frac{n_i+1}{2}-d_{H_1}(p,r_1^i)$, or $d_{H_1}(u,r_2^i)=\frac{n_i+1}{2}-d_{H_1}(u,r_1^i)$ and $d_{H_1}(p,r_2^i)=\frac{n_i-1}{2}-d_{H_1}(p,r_1^i)$. From $d_{H_1}(u,r_2^i)>d_{H_1}(p,r_2^i)$ as given by Condition (3), we obtain $d_{H_1}(p,r_1^i)>d_{H_1}(u,r_1^i)+1$ or $d_{H_1}(p,r_1^i)>d_{H_1}(u,r_1^i)-1$. In either case, we get that (3) cannot hold for both $r=r_1^i$ and $r=r_2^i$.

Note that when comparing components C_i and C_j with $i\neq j$, only vertices of the extended resolving set coming from component C_i were used to distinguish between any two vertices from components C_i and C_j. Hence, for one component, say, C_k, it is enough to choose a resolving set, that is, a set that distinguishes all vertices within C_k (a minimum cardinality resolving set is always of size 2). Hence, if $k_e>0$, we may assume that C_k is an even cycle. Thus only 2 vertices are chosen from C_k, and from all other even cycles 3 vertices are chosen. Thus, in this case $2k+k_e-1$ vertices are enough. If $k_e=0$, then 2 vertices are chosen from each component, giving the bound $2k$ in this case.

Now, we prove the claim of necessity. Observe first that clearly at least 2 vertices of each cycle have to be chosen, as otherwise the two neighbors of the chosen vertex r cannot be distinguished; one can construct a graph H_1 by connecting r with one fixed vertex of each other component, and the two neighbors of r are indistinguishable.

Let us first assume that there exist two components C_i and C_j both containing an even number of vertices, and from each component, only two vertices are included in O. Denote by r_1^i, r_2^i the vertices chosen from C_i and by r_1^j, r_2^j the vertices chosen from C_j. If in at least one component, say C_i, the two selected vertices r_1^i and r_2^i are at distance exactly $\frac{n_i}{2}$ from each other, let u and v be two neighbors of r_1^i. Note that u and v are equidistant from both r_1^i and r_2^i. Constructing H_1 by connecting some vertex z_ℓ from every other component C_ℓ to r_1^i, the vertices u and v are still not distinguishable within H_1, as shown in Fig. 7a. Otherwise, let us assume that in both components C_i and C_j the vertices selected in O are not at distance exactly $\frac{n_i}{2}$ ($\frac{n_j}{2}$, respectively) from each other. Let u then be a neighbor of r_1^i in C_i that is on the longer path $r_1^i - r_2^i$, and let v be a neighbor of r_1^j in C_j that is on the longer path $r_1^j - r_2^j$. We can construct H_1 by connecting u with r_1^j

and u with some vertex z_ℓ of every component C_ℓ (if there are more than 2 components). H_2 is constructed by connecting v with r_1^i and v with the same vertex z_ℓ (for every other component C_ℓ) as in H_1, as shown in Fig. 7b. The distances of the vertices u, v from the vertices in O are

$$d_{H_1}(u, r_1^i) = d_{H_2}(v, r_1^i) = 1$$
$$d_{H_1}(u, r_2^i) = d_{H_2}(v, r_2^i) = 1 + d_{H_1}(r_1^i, r_2^i)$$
$$d_{H_1}(u, r_1^j) = d_{H_2}(v, r_1^j) = 1$$
$$d_{H_1}(u, r_2^j) = d_{H_2}(v, r_2^j) = 1 + d_{H_2}(r_1^j, r_2^j)$$
$$d_{H_1}(u, r) = d_{H_2}(v, r) = 1 + d_{H_1}(z_\ell, r),$$

for $r \in O_l$, $l \neq i, j$. Hence the vertices u and v are indistinguishable.

Therefore, if both C_i and C_j have an even number of vertices, at least 3 vertices of component C_i or 3 vertices of component C_j have to be included in O. Hence, from all but one component with an even number of vertices, 3 vertices have to be chosen, and from the remaining ones, at least 2. This completes the proof. □

3. Proofs of results for general graph classes

We start with the following easy observation.

Observation 3.1. *Let G be a connected graph. Consider any two vertices r and u of G, and consider a shortest path $r - u$. Either u is a boundary vertex for r, or there exists some vertex u' such that the shortest path $r - u$ can be extended to a shortest path $r - u'$, with u' being a boundary vertex for r.*

Proof. If u is not a boundary vertex for r, then by definition there exists a neighbor w of u such that $d_G(w, r) > d_G(u, r)$. Thus, $d_G(w, r) \geq d_G(u, r) + 1$, and in particular, a shortest path $r - u$ can be extended to w such that along this extended path, the lower bound can be attained, and thus $d_G(w, r) = d_G(u, r) + 1$. Hence, the path $r - w$ going through u is also a shortest path $r - w$. If w is then a boundary vertex for r, we are done, and otherwise we iteratively apply the same argument with w playing the role of u. The claim follows. □

We are now ready to show our results in terms of boundary vertices.

Proof of Theorem 1.5. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. We need to show that for the set $O = \cup_{i=1}^{k-1} \partial(C_i) \cup S_k$ the condition $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ holds for an arbitrary graph.

Since the boundary is a resolving set, any two vertices belonging to the same component are distinguishable by a set that contains the boundaries of $k - 1$ components and a resolving set of the k-th component. As before, let $u \in V(C_i)$, $v \in V(C_j)$, let $p \in V(C_i)$ and $q \in V(C_j)$ such that any path from a vertex in C_i to any vertex in C_j in H_2 contains a subpath $p - q$, and let $i < k$. If u is a boundary vertex, it is distinguishable from v, since $0 = d_{H_1}(u, u) < d_{H_2}(u, v)$. If u is not a boundary vertex, and $u = p$, then for any boundary vertex $r \in \partial(C_i)$, $d_{H_2}(r, v) = d_{H_2}(r, p) + d_{H_2}(p, q) + d_{H_2}(q, v) \geq d_{H_1}(r, p) + d_{H_2}(p, q) > d_{H_1}(r, p)$. Thus, the two distance vectors are not equal either. Now, we consider the case when $u \neq p$. If u is a boundary vertex for p, let $u' = u$. Otherwise, the shortest path between p and u in component C_i can be extended to a shortest path $p - u'$ by Observation 3.1, such that u' is a boundary vertex of p. For a fixed shortest path $p - u'$ we have $d_{H_2}(u', v) = d_{H_2}(u', p) + d_{H_2}(p, q) + d_{H_2}(q, v) = d_{H_1}(u', u) + d_{H_1}(u, p) + d_{H_2}(p, q) + d_{H_2}(q, v) > d_{H_1}(u', u)$, which completes the proof. □

Proof of Theorem 1.6. Let $u, v \in V(G)$ be any two different vertices, and let H_1, H_2 be any two graphs from the set of possible graphs $\mathcal{H}(F)$. We need to show that for the set $O = \cup_{i=1}^{k-1} O_i \cup S_k$, where $O_i = S_i \cup \partial(S_i)$, the condition $\boldsymbol{d}_{H_1}(u, O) \neq \boldsymbol{d}_{H_2}(v, O)$ holds for an arbitrary graph.

Let $r \in S_i$ be a vertex from a resolving set of a component C_i. Once more, let $u \in V(C_i)$, $v \in V(C_j)$, with $i < k$. Let p be the vertex in C_i and q the vertex in C_j, such that any path from a vertex in C_i to any vertex in C_j in H_2 contains a subpath $p - q$, with $d_{H_2}(p, q) \geq 1$. As in the proof of Theorem 1.5, if u is a boundary vertex for r, let $u = u'$. Otherwise, by Observation 3.1, the shortest path between r and u in component C_i can be extended to a shortest path $r - u'$, with u' being a boundary vertex for r. We need to show that $d_{H_1}(u, u') \neq d_{H_2}(v, u')$, for any vertex v belonging to some other component C_j (as in the previous theorems, if u and v are in the same component, they are distinguishable by the resolving set of that component). If u is a boundary vertex itself, then we clearly have $d_{H_1}(u, u') = 0 \neq d_{H_2}(v, u')$, so we may assume $u \neq u'$. If r does not distinguish u and v, then $d_{H_1}(u, r) = d_{H_2}(v, r) = d_{H_1}(p, r) + d_{H_2}(p, q) + d_{H_2}(q, v)$ and

$$d_{H_1}(u, r) > d_{H_1}(p, r), \tag{4}$$

holds, since $d_{H_2}(p, q) \geq 1$.

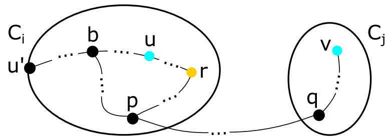


Fig. 8. Proof of Theorem 1.6: Extending the shortest path $r - u$ to a shortest path $r - u'$.

Case I: There exists a shortest path from u' to p in component C_i that passes through u. Hence there exists a shortest path from u' to v in H_2 that passes through u. Then we have $d_{H_2}(u', v) = d_{H_1}(u', u) + d_{H_1}(u, p) + d_{H_2}(p, v) > d_{H_1}(u', u)$. Thus u and v have different distances to u', and they are distinguishable.

Case II: All shortest paths from u' to p in component C_i do not pass through u. Hence no shortest path from u' to v in H_2 passes through u. Let b be the vertex closest to u on this path, such that the path $b - u'$ is common to both shortest paths $p - u'$ and $r - u'$, as illustrated in Fig. 8. Note that b might coincide with u', but not with u. Also observe that since b is on the extension of a shortest path $r - u$ to a shortest path $r - u'$, at least one shortest path $r - b$ passes through u. Therefore, we have

$$d_{H_1}(r, u) + d_{H_1}(u, b) \leq d_{H_1}(r, p) + d_{H_1}(p, b). \tag{5}$$

Since (4) holds, from (5) it follows that

$$d_{H_1}(u, b) < d_{H_1}(p, b). \tag{6}$$

Now, from (6) and the fact that $d_{H_2}(v, p) = d_{H_2}(v, q) + d_{H_2}(q, p) \geq 1$ we obtain

$$\begin{aligned} d_{H_2}(v, u') &= d_{H_2}(v, p) + d_{H_2}(p, b) + d_{H_2}(b, u') \\ &> d_{H_2}(v, p) + d_{H_2}(u, b) + d_{H_2}(b, u') \\ &> d_{H_1}(u, u'). \end{aligned}$$

Therefore, u and v have different distances to the boundary vertex u', and they are thus distinguishable by a boundary vertex of a vertex belonging to the resolving set. The theorem follows. □

Inspecting the proofs of Theorems 1.1, 1.3, 1.4, we see that when comparing two vertices from C_i and C_j, in fact only the structure of C_i and its resolving set matters. Therefore, whenever one of the components of the observed disconnected graph F is a tree (cycle, or grid, respectively), then instead of including a resolving set and its boundary vertices, it is sufficient to choose all leaves in the case the component is a tree (two neighboring vertices together with a vertex at distance at least $\frac{n-2}{2}$ from both of them in the case of the even cycle on n vertices, two vertices at distance $\frac{n-1}{2}$ from each other in the case of an odd cycle on n vertices, and three corner vertices in the case of the grid, respectively). Note that this might be better than the bound claimed by Theorem 1.6, which for example in the case of the grid requires all four corner points to be chosen. Also, note that applying Theorem 1.6 to the case when all components are trees can yield exactly the results of Theorem 1.1. When a subset of leaves is selected as a resolving set of a tree component, then the resolving set and its boundary is precisely the set of all the leaves, hence Theorem 1.6 constructs a minimum cardinality extended resolving set. However, for some classes of graphs we conjecture that the bound given in Theorem 1.6 can be arbitrarily bad, and it is future work to make this argument rigorous.

4. Concluding remarks

We have introduced and analyzed the concept of an extended metric dimension for different graph classes. The proposed metric enables the introduction of uncertainty in graph topology in problems modeled with metric dimension. One such problem is to find the minimum number of observed nodes needed for identification of the source node of network diffusion, in the settings where knowing the full network topology is not feasible.

We have given exact answers on this extended metric dimension for trees, cycles, grids, and complete graphs, and have given general upper bounds for arbitrary graphs in terms of their boundary. Needless to say, it would be interesting to determine this number exactly for other graph classes, such as bipartite graphs, or to find tighter bounds. Additionally, in practical scenarios involving network diffusion, links connecting the vertices of the network represent stochastic propagation times of some rumor or a virus. Hence, it would be of practical interest to analyze a suitably defined stochastic version of both the standard and extended metric dimension problems.

Acknowledgements

This research was partially supported by Fundação para a Ciência e a Tecnologia (project FCT [UID/EEA/50009/2013], a Ph.D. grant from the Carnegie Mellon-Portugal program, number SFRH/BD/51567/2011) and EU-H2020 WiMUST project (grant agreement No. 645141). We would also like to thank the anonymous referees for careful reading.

References

[1] B. Bollobás, D. Mitsche, P. Prałat, Metric dimension for random graphs, Electron. J. Combin. 20 (2013) P1.
[2] 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.
[3] G. Chartrand, D. Erwin, G.L. Johns, P. Zhang, Boundary vertices in graphs, Discrete Math. 263 (2003) 25–34.
[4] X. Chen, X. Hu, C. Wang, Approximability of the minimum weighted doubly resolving set problem, in: Computing and Combinatorics, Proceedings of 20th International Conference, COCOON, 2014.
[5] J. Díaz, O. Pottonen, M.J. Serna, E.J. van Leeuwen, On the complexity of metric dimension, in: Algorithms, ESA 2012, in: Lecture Notes in Computer Science, vol. 7501, Springer, Berlin, Heidelberg, 2012, pp. 419–430.
[6] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman, 1979.
[7] F. Harary, R.A. Melter, The metric dimension of a graph, Ars Combin. 2 (1976) 191–195.
[8] C. Hernando, M. Mora, I.M. Pelayo, C. Seara, Some structural, metric and convex properties on the boundary of a graph, Electron. Notes Discrete Math. 24 (2006) 203–209.
[9] S. Khuller, B. Raghavachari, A. Rosenfeld, Landmarks in graphs, Discrete Appl. Math. 70 (1996) 217–229.
[10] P. Slater, Leaves of trees, Congr. Numer. 14 (1975) 549–559.
[11] S. Zejnilovic, J. Gomes, B. Sinopoli, Network observability and localization of the source of diffusion based on a subset of nodes, in: 51st Annual Allerton Conference on Communication, Control, and Computing, 2013, pp. 847–852.
[12] S. Zejnilovic, D. Mitsche, J. Gomes, B. Sinopoli, Network observability for source localization in graphs with unobserved edges, in: IEEE Global Conference on Signal and Information Processing, GlobalSIP, 2014.