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

ARTICLE INFO

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

ABSTRACT

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 $1 \leq \beta(G) \leq n - 1$, 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]).

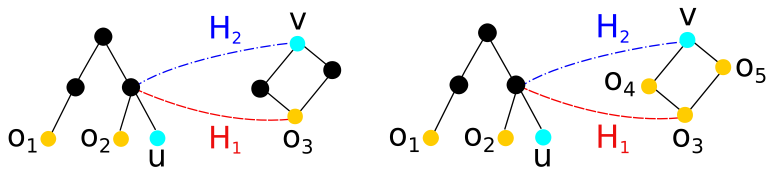
In this paper[1] we extend the concept of metric dimension to graphs with some edges missing: suppose we are given a finite, simple graph $F = (V, E)$ with $|V| = n$ consisting of $k \geq 2$ connected components, denoted by C_i, for $i = 1, \ldots k$. Denote the class $\mathcal{H}(F)$ to be the class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