Discrete Applied Mathematics 236 (2018) 347–386

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

The resistance perturbation distance: A metric for the analysis of dynamic networks

Nathan D. Monnig [a], François G. Meyer [b,*]

[a] Numerica Corporation, Fort Collins, CO 080528, United States
[b] University of Colorado at Boulder, Boulder, CO 80305, United States

A R T I C L E   I N F O

Article history:
Received 23 April 2016
Received in revised form 29 July 2017
Accepted 12 October 2017
Available online 22 November 2017

Keywords:
Graph distance
Dynamic graph
Effective resistance
Commute time
Laplacian

A B S T R A C T

To quantify the fundamental evolution of time-varying networks, and detect abnormal behavior, one needs a notion of temporal difference that captures significant organizational changes between two successive instants. In this work, we propose a family of distances that can be tuned to quantify structural changes occurring on a graph at different scales: from the local scale formed by the neighbors of each vertex, to the largest scale that quantifies the connections between clusters, or communities. Our approach results in the definition of a true distance, and not merely a notion of similarity. We propose fast (linear in the number of edges) randomized algorithms that can quickly compute an approximation to the graph metric. The third contribution involves a fast algorithm to increase the robustness of a network by optimally decreasing the Kirchhoff index. Finally, we conduct several experiments on synthetic graphs and real networks, and we demonstrate that we can detect configurational changes that are directly related to the hidden variables governing the evolution of dynamic networks.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Many complex systems are well represented as graphs or networks, with the agents represented as vertices and edges symbolizing relationships or similarities between them. In many instances, the relationships between vertices evolve as a function of time: edges may appear and disappear, the weights along the edges may change. The study of such *dynamic graphs* often involves the identification of patterns that couple changes in the network topology with the latent dynamical processes that drive the evolution of the connectivity of the network [2,24,30,31,40].

To quantify the temporal and structural evolution of time-varying networks, and detect abnormal behavior, one needs a notion of temporal difference that captures significant configurational changes between two successive instants. The design of similarity measures for the pairwise comparison of graphs [43] is therefore of fundamental importance.

Because we are interested in detecting changes between two successive instants, we focus on the problem of measuring the distance between two graphs on the same vertex set, with known vertex correspondence (see Fig. 1). We note that determining whether two graphs are isomorphic under a permutation of the vertex labels is a combinatorially hard problem (e.g., [35], and references therein, but see the recent results [3]). Several notion of similarities (e.g., [8,36,29], and references therein) have been proposed. Unlike a true metric, a similarity merely provides a notion of resemblance. Most approaches rely on the construction of a feature vector that provides a signature of the graph characteristics; the respective feature

* Corresponding author.
E-mail addresses: nathan.monnig@colorado.edu (N.D. Monnig), fmeyer@colorado.edu (F.G. Meyer).

https://doi.org/10.1016/j.dam.2017.10.007
0166-218X/© 2017 Elsevier B.V. All rights reserved.

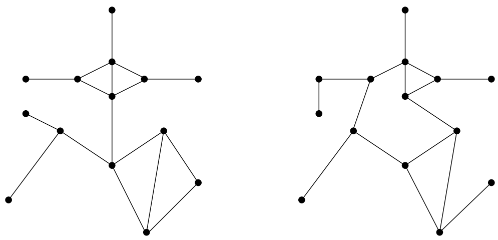

Fig. 1. Dynamic graph $G^{(t)}$ at time t (left) and $t+1$ (right).

vectors of the two graphs are then compared using some norm, or distance. A similarity function is typically not injective (two graphs can be perfectly similar without being the same), and rarely satisfies the triangular inequality.

Instead of comparing two feature vectors, several researchers (e.g., [1,9,11,4,19,41,50] and references therein) have proposed to use a kernel function. This approach offers the same advantage as the computation of a similarity: the isomorphism problem need not be solved. Unfortunately, the kernels do not define proper metrics, and we are left with a weaker notion of similarity.

Several distances between two graphs with the same size have been proposed (e.g., [7,11], and references therein). As detailed in Section 3.2, we argue that existing distances either fail to capture a notion of structural similarity, or lead to algorithms that have a high computational complexity.

1.1. Contribution and organization of the paper

The contributions of this work are threefold. First, we propose a family of distances that can be tuned to quantify configurational changes that occur on a graph at different scales: from the local scale formed by the local neighbors of each vertex, to the largest scale that quantifies the connections between clusters, or communities. Our approach results in the definition of a true distance, and not merely a notion of similarity. The second contribution encompasses fast computational algorithms to evaluate the metrics developed in the first part. We developed fast (linear in the number of edges) randomized algorithms that can quickly compute an approximation to the graph metric. The third contribution involves fast algorithms to increase the robustness of a network by optimally decreasing the Kirchhoff index. Finally, we conduct several experiments on synthetic and real dynamic networks, and we demonstrate that the resistance perturbation distance can detect the significant changes in the hidden latent variables that control the network dynamics.

The remainder of this paper is organized as follows. In the next section we introduce the main mathematical concepts and corresponding nomenclature. In Section 3 we formally define the problem and review the existing literature. In Section 4 we propose a novel framework for constructing graph distances; we focus the rest of the paper on the **resistance perturbation distance**, which is defined in Section 5. In Section 6, we study simple perturbations of several prototypical graphs for which the resistance perturbation distance can be computed analytically. Fast randomized algorithms are described in Section 7. The optimization of the robustness of a network, based on optimally decreasing the Kirchhoff index, is described in Section 8. In Section 9, we use the resistance perturbation metric to detect significant changes in synthetic and real dynamic networks. We conclude in Section 10 with a discussion on future work. Some technical details and proofs are left aside in Appendix. A list of the main notations used in the paper is provided in Section 11.

2. Preliminaries and notation

We introduce in this section the main concepts and associated nomenclature.

We denote by $\boldsymbol{e}_i$ the ith vector of the canonical basis in $\mathbb{R}^n$. The space of matrices of size $n \times m$ with entries in $\mathbb{R}$ is denoted by $\boldsymbol{M}_{n\times m}$; to alleviate notations we write $\boldsymbol{M}_n$ to denote $\boldsymbol{M}_{n\times n}$.

We denote by $G = (V, E, w)$ an undirected weighted graph with a vertex set $V = \{1, \ldots, n\}$, an edge set E, and a symmetric weight function w that quantifies the similarity between any two vertices i and j. In this work, we use the terms graph and network exchangeably.

The weighted adjacency matrix, $\boldsymbol{A} \in \boldsymbol{M}_n$, is given by

$$A_{ij} = A_{ji} = \begin{cases} w_e & \text{if the edge } e = [i, j] \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

For simplicity, we will always assume G is connected and does not contain any self-loops. We further define the combinatorial Laplacian matrix,

$$\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}, \tag{2}$$

Download English Version:

https://daneshyari.com/en/article/6871582

Download Persian Version:

https://daneshyari.com/article/6871582

Daneshyari.com