Available online at www.sciencedirect.com
ScienceDirect
Journal of the Franklin Institute 000 (2017) 1–22
www.elsevier.com/locate/jfranklin

Fast distributed algebraic connectivity estimation in large scale networks

Eduardo Montijano[a,*], Juan I. Montijano[b], Carlos Sagues[a]

[a] *Instituto de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza, Spain*
[b] *Instituto Universitario de Matemáticas y Aplicaciones (IUMA), Universidad de Zaragoza, Spain*

Received 3 August 2015; received in revised form 15 January 2017; accepted 9 May 2017
Available online xxx

Abstract

This paper presents a distributed method to estimate the algebraic connectivity of fixed undirected communication graphs. The proposed algorithm uses bisection to estimate the second smallest eigenvalue of the Laplacian matrix associated to the graph. In order to decide the sub-interval in which the eigenvalue is contained, a distributed averaging algorithm based on Chebyshev polynomials is considered together with a max consensus algorithm. The information exchanged by neighbors in the graph each communication round is constant and independent of the size of the network, making it scalable to large networks. Besides, exploiting the convergence properties of Chebyshev polynomials we provide a direct estimation of the algebraic connectivity so that, instead of the midpoint of the bisection interval, the new approximation can be used. Finally, our algorithm also provides upper and lower bounds on the algebraic connectivity and an estimation of the Fiedler eigenvector associated to it. Simulations in large networks demonstrate the scalability and accuracy of the algorithm.

1. Introduction

Knowledge of the properties of the network topology can play a fundamental role when facing the problem of designing distributed algorithms to be run in ad hoc networks. The research community, aware of this issue, has devoted a lot of effort in the last decade to propose distributed solutions that allow a network to determine most of its properties. A parameter of

* Corresponding author.
E-mail address: emonti@unizar.es (E. Montijano).

http://dx.doi.org/10.1016/j.jfranklin.2017.05.021
0016-0032/

particular importance in communication networks is the Algebraic Connectivity (AC). This parameter is crucial for example in the convergence speed of linear iterations in distributed averaging [1] and provides information about whether the graph topology is connected or not [2,3]. However, since the algebraic connectivity depends on all the communication links present in the network its computation is not trivial in a distributed fashion.

Some works use the Fourier Transform to estimate the entire spectrum of the Laplacian matrix, instead of just the algebraic connectivity [4,5]. The second algorithm allows for an easy clusterization of the graph [6]. On the other hand, the computation of the Fourier Transform requires all the values of the measured signal along time, therefore needing to store large amounts of data in each node of the network.

A distributed algorithm to compute an orthogonal decomposition that contains all the eigenvalues of the graph is given in [7]. Distributed powers of a deflated matrix are computed using event-triggered consensus in [8], providing an estimation with upper and lower bounds of the algebraic connectivity [9]. In [10] they introduce a distributed algorithm to compute the coefficients of the minimal polynomial, which is equivalent to compute all the eigenvalues of the weight matrix. A similar approach is given in [11], where the coefficients of the minimal polynomial are computed by solving an optimization problem in a distributed manner. The number and the size of the messages the nodes need to exchange in these kind of approaches, [7–11], grows linearly with the size of the network, which implies that for large networks a lot of communications will be required.

The mathematical relationship between the eigenvalues of the Laplacian matrix, and their spectral moments, and local structural features, like the number of triangles in the network topology, is analyzed in [12,13], also giving distributed algorithms to compute and control these local features [14]. While these features can be used for more than spectral information, its distributed computation requires, besides distributed averaging linear iterations, information from two-hop neighbors.

Finally, several recent works exploit a power iteration algorithm [15] to distributively estimate the AC, along with the associated Fiedler vector. A mean correction algorithm together with the power iteration is introduced in [16] to avoid numerical problems of the latter. This algorithm is used to adapt the weights of the Laplacian in [17]. Similarly, the power iteration is used to estimate the AC and then optimize a distributed averaging procedure in [18]. A stochastic power iteration with nested consensus operations is used in [19] considering random graphs, and directed graph topologies are analyzed in [20]. The main advantage of using the power iteration is that the amount of information that nodes exchange at each communication round with their neighbors remains constant independently of the size of the network. On the other hand, the limitation of the power iteration is that it requires a large increasing number of communication rounds to estimate the AC with sufficient accuracy as the number of nodes grows, making it hardly applicable in large scale networks. This is caused by the convergence rate of the algorithm, which is determined by the quotient between the second and third largest eigenvalues of the weight matrix associated to the Laplacian. While this quotient is low for small networks, for large scale networks, with hundreds of nodes, it is close to the unity, implying that the number of communication rounds required to compute a good estimation of the algebraic connectivity will be quite large. Moreover, the nested iterations presented in methods like [19] increase even more the number of communication rounds required to estimate the AC with sufficient accuracy.

To overcome this limitation, our contribution is a novel fast distributed algorithm, able to compute the AC and the Fiedler vector in large scale networks. Our algorithm builds