Physica A 392 (2013) 3833–3845

Contents lists available at SciVerse ScienceDirect

Physica A

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

Geometry of complex networks and topological centrality

Gyan Ranjan *, Zhi-Li Zhang

Department of Computer Science, University of Minnesota, Twin Cities, USA

HIGHLIGHTS

- A geometric approach to centrality using the Moore–Penrose pseudo-inverse of the Laplacian.
- Topological centrality of a node is determined in terms of the position vector and the robustness of the overall network in terms of the Kirchhoff index.
- Interpretations provided in terms of detour overheads, recurrence probabilities and connectedness in bi-partitions.
- Empirical analysis shows how these indices reflect structural roles of nodes in the network and their sensitivity to perturbations.

ARTICLE INFO

Article history:
Received 9 July 2011
Received in revised form 29 October 2012
Available online 22 April 2013

Keywords:
Moore–Penrose pseudo-inverse of the Laplacian
Topological centrality
Kirchhoff index
Expected hitting and commute times
Equivalent electrical network
Connected bi-partitions of a graph

ABSTRACT

We explore the geometry of complex networks in terms of an n-dimensional Euclidean embedding represented by the Moore–Penrose pseudo-inverse of the graph Laplacian ($\mathbf{L}^+$). The squared distance of a node i to the origin in this n-dimensional space (l^+_{ii}), yields a topological centrality index, defined as $\mathcal{C}^*(i) = 1/l^+_{ii}$. In turn, the sum of reciprocals of individual node centralities, $\sum_i 1/\mathcal{C}^*(i) = \sum_i l^+_{ii}$, or the trace of $\mathbf{L}^+$, yields the well-known Kirchhoff index ($\mathcal{K}$), an overall structural descriptor for the network. To put into context this geometric definition of centrality, we provide alternative interpretations of the proposed indices that connect them to meaningful topological characteristics — first, as forced detour overheads and frequency of recurrences in random walks that has an interesting analogy to voltage distributions in the equivalent electrical network; and then as the average connectedness of i in all the bi-partitions of the graph. These interpretations respectively help establish the topological centrality ($\mathcal{C}^*(i)$) of node i as a measure of its overall *position* as well as its overall *connectedness* in the network; thus reflecting the robustness of i to random multiple edge failures. Through empirical evaluations using synthetic and real world networks, we demonstrate how the topological centrality is better able to distinguish nodes in terms of their structural roles in the network and, along with Kirchhoff index, is appropriately sensitive to perturbations/re-wirings in the network.

1. Introduction

Unlike traditional studies on network robustness, that typically treat networks as combinatoric objects and rely exclusively on classical graph-theoretic concepts (such as degree distributions, geodesics and minimum cuts), we explore a geometric approach as an alternative. To do so, we embed the network into an n-dimensional Euclidean space (n being the number of nodes in the network) represented by the Moore–Penrose pseudo-inverse of its graph Laplacian, denoted henceforth by $\mathbf{L}^+$. The diagonal entries of $\mathbf{L}^+$, denoted as l^+_{ii} for the node i, represent the squared distance of node i to the origin in this n-dimensional space and provide a measure of the node's *topological centrality*, given as $\mathcal{C}^*(i) = 1/l^+_{ii}$. Closer the node i is to the origin in this space, or equivalently lower the l^+_{ii}, more *topologically central* i is. Similarly, the trace of

* Corresponding author. Tel.: +1 6128405155.
E-mail addresses: granjan@cs.umn.edu (G. Ranjan), zhzhang@cs.umn.edu (Z.-L. Zhang).

0378-4371/$ – see front matter
http://dx.doi.org/10.1016/j.physa.2013.04.013

$\mathbf{L}^+$, $\mathrm{Tr}(\mathbf{L}^+) = \sum_i 1/\mathcal{C}^*(i)$, determines the overall *volume* of the embedding and yields the well-known *Kirchhoff index* ($\mathcal{K}$), a structural descriptor for the network as a whole. Once again, lower the value of $\mathcal{K}$ for a network (from amongst all possible networks with the same number of nodes and edges), more compact the embedding, and more structurally robust the overall network is. In short, topological centrality defines a ranking of the nodes of a given network, whereas the Kirchhoff index provides a geometric measure to rank different networks of comparable sizes.

In order to illustrate how the two geometric quantities defined above actually reflect the structural properties of the underlying network, and in particular to structural robustness against multiple failures, we provide three alternative interpretations for them in terms of: (a) detour overheads in random walks, (b) voltage distributions and the phenomenon of recurrence when the network is treated as an electrical circuit, and (c) the average connectedness of nodes when the network breaks into two, thereby making global communication untenable. We describe each of these in detail below.

First, we show how topological centrality of a node reveals its overall *position* in the network. By equating topological centrality of a node *i*, i.e. $\mathcal{C}^*(i) = 1/l^+_{ii}$, to the (reciprocal of) average detour overhead incurred when a random walk between any source–destination pair is forced to go through *i*, we get a measure of the node's position. Intuitively, the average overhead incurred in such forced detours (measured in terms of the number of steps in the random walk) is lower if node *i* is *centrally* positioned in the network (higher $\mathcal{C}^*(i)$ and lower l^+_{ii}) and higher if *i* is *peripheral*. Secondly, we show how $\mathcal{C}^*(i)$ captures voltage distribution when the network is transformed into an equivalent electrical network (EEN). This, in turn, is related to the probability with which a random detour through *i* returns to the source node; referred to as the phenomenon of *recurrence* in random walk literature. To be precise, higher $\mathcal{C}^*(i)$ implies that a random detour through node *i* forces the random walk between any source–destination pair to return to the source node with lower probability, thereby incurring lower detour overhead. Both of these interpretations, namely average detour overhead and probability of recurrence, therefore, demonstrate how $\mathcal{C}^*(i)$ quantifies the overall position of node *i* in the network. Finally, we establish how $\mathcal{C}^*(i)$ captures the overall connectedness of node *i*. To do so, we equate it to the number of nodes that *i* can communicate with when a subset of edges in the network fail in such a way that the network is partitioned into two connected sub-networks. As connected bi-partitions represent a regressed state of the network when not all pairs of nodes can maintain communication, a higher value of $\mathcal{C}^*(i)$, implies that *i* is present in the larger of the two sub-networks, on an average, in such bi-partitions. Thus, $\mathcal{C}^*(i)$ reflects the immunity/vulnerability of node *i* towards multiple edge failures in the network, a distinct topological characteristic.

Through numerical simulations using synthetic and realistic network topologies, we demonstrate that our new indices better characterize robustness of nodes in network, both in terms of position as well as connectedness, as compared to other existing metrics (e.g. node centrality measured based on degree, shortest paths, etc.). A rank-order of nodes by their topological centralities ($\mathcal{C}^*(i)$) helps distinguish them in terms of their structural roles (such as core, gateway, etc.). Also, topological centrality and Kirchhoff index, are both appropriately sensitive to local perturbations in the network, a desirable property not displayed by some of the other popular centrality indices in literature (as shown later in this paper).

The rest of the paper is organized as follows: we begin by providing a brief overview of several structural indices, characterizing node centrality as well as overall descriptors for networks, found in literature in Section 2. Section 3 introduces a geometric embedding of the network using the eigenspace of $\mathbf{L}^+$, topological centrality and Kirchhoff index as measures of robustness. Section 4 demonstrates how topological centrality of a node reflects the average detour overhead in random walks through a particular node in question followed by its equivalence to the probability of recurrence. In Section 5 we show how topological centrality captures the average connectedness of nodes in the bi-partitions of a network. Section 6 presents comparative empirical analysis with simulated as well as real world networks while in Section 7 we analyze the computational complexity of the proposed metrics with respect to others popular in literature. Finally, the paper is concluded with a discussion of future work in Section 8.

2. Related work

Robustness of nodes to failures in complex networks is dependent on their overall *position* and *connectedness* in the network. Several centralities, that characterize position and/or connectedness of nodes in complex networks in different ways, have therefore been proposed in literature. Perhaps the simplest of all is degree—the number of edges incident on a node. Degree is essentially a *local* measure i.e. a first order/one-hop connectedness index. A second-order variant called *joint-degree*, given by the product of degrees of a pair of nodes that are connected by an edge in the network, is also in vogue. However, except in *scale free* networks that display the so called *rich club connectivity* [1–3], neither degree nor joint-degree determine the overall position or the connectedness of nodes.

A class of structural indices called *betweennesses*, namely shortest path/geodesic (*GB*) [4,5], flow (*FB*) [6] and random-walk (*RB*) [7] betweenness respectively quantify the positions of nodes, with respect to source–destination pairs in the network. The set of betweennesses, therefore, reflect the role played by a node in the communication between other node-pairs in the network and are not the measures of a node's own connectedness.

Another popular centrality measure is geodesic closeness (*GC*) [4,5]. It is defined as the (reciprocal of) average shortest-path distance of a node from all other nodes in the network. Clearly, geodesic closeness is a *p*th-order measure of connectedness where $p = \{1, 2, \ldots, \delta\}$, δ being the geodesic diameter of the graph, and is better suited for characterizing global connectedness properties than the aforementioned indices. However, communication in networks is not always confined to shortest paths alone and *GC* being geodesic based, ignores other alternative paths between nodes, however