Social Networks 32 (2010) 245–251

Contents lists available at ScienceDirect

Social Networks

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

Node centrality in weighted networks: Generalizing degree and shortest paths

Tore Opsahl[a,*], Filip Agneessens[b], John Skvoretz[c]

[a] Imperial College Business School, Imperial College London, London SW7 2AZ, UK
[b] Department of Organization Sciences, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands
[c] College of Arts & Sciences, University of South Florida, Tampa, FL 33620, USA

ARTICLE INFO

Degree
Closeness
Betweenness
Weighted networks

ABSTRACT

Ties often have a strength naturally associated with them that differentiate them from each other. Tie strength has been operationalized as weights. A few network measures have been proposed for weighted networks, including three common measures of node centrality: degree, closeness, and betweenness. However, these generalizations have solely focused on tie weights, and not on the number of ties, which was the central component of the original measures. This paper proposes generalizations that combine both these aspects. We illustrate the benefits of this approach by applying one of them to Freeman's EIES dataset.

1. Introduction

Social network scholars are increasingly interested in trying to capture more complex relational states between nodes. One of these avenues of research has focused on the issue of tie strength, and a number of studies from a wide range of fields have begun to explore this issue (Barrat et al., 2004; Brandes, 2001; Doreian et al., 2005; Freeman et al., 1991; Granovetter, 1973; Newman, 2001; Opsahl and Panzarasa, 2009; Yang and Knoke, 2001). Whether the nodes represent individuals, organizations, or even countries, and the ties refer to communication, cooperation, friendship, or trade, ties can be differentiated in most settings. These differences can be analyzed by defining a weighted network, in which ties are not just either present or absent, but have some form of weight attached to them. In a social network, the weight of a tie is generally a function of duration, emotional intensity, intimacy, and exchange of services (Granovetter, 1973). For non-social networks, the weight often quantifies the capacity or capability of the tie (e.g., the number of seats among airports; Colizza et al., 2007; Opsahl et al., 2008) or the number of synapses and gap junctions in a neural network (Watts and Strogatz, 1998). Nevertheless, most social network measures are solely defined for binary situations and, thus, unable to deal with weighted networks directly (Freeman, 2004; Wasserman and Faust, 1994). By dichotomizing the network, much of the information contained in a weighted network datasets is lost, and consequently, the complexity of the network topology cannot be described to the same extent or as richly. As a result, there has been a growing need for network measures that directly account for tie weights.

The centrality of nodes, or the identification of which nodes are more "central" than others, has been a key issue in network analysis (Freeman, 1978; Bonacich, 1987; Borgatti, 2005; Borgatti et al., 2006). Freeman (1978) argued that central nodes were those "in the thick of things" or focal points. To exemplify his idea, he used a network consisting of 5 nodes (see Fig. 1). The middle node has three advantages over the other nodes: it has more ties, it can reach all the others more quickly, and it controls the flow between the others. Based on these three features, Freeman (1978) formalized three different measures of node centrality: degree, closeness, and betweenness. Degree is the number of nodes that a focal node is connected to, and measures the involvement of the node in the network. Its simplicity is an advantage: only the local structure around a node must be known for it to be calculated (e.g., when using data from the General Social Survey; McPherson et al., 2001). However, there are limitations: the measure does not take into consideration the global structure of the network. For example, although a node might be connected to many others, it might not be in a position to reach others quickly to access resources, such as information or knowledge (Borgatti, 2005; Brass, 1984). To capture this feature, closeness centrality was defined as the inverse sum of shortest distances to all other nodes from a focal node. A main limitation of closeness is the lack of applicability to networks with disconnected components: two nodes that belong to different components do not have a finite distance between them. Thus, closeness is generally restricted to nodes within the largest component of a network[1].

* Corresponding author. Tel.: +44 20 7594 3035.
E-mail addresses: t.opsahl@imperial.ac.uk (T. Opsahl), f.agneessens@fsw.vu.nl (F. Agneessens), skvoretz@cas.usf.edu (J. Skvoretz).

[1] A possible method for overcoming this limitation is to sum the inversed distances instead of the inverse sum of distances as the limit of 1 over infinity is 0.

0378-8733/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.socnet.2010.03.006

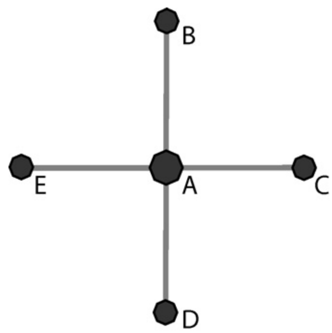

Fig. 1. A star network with 5 nodes and 4 edges. The size of the nodes corresponds to the nodes' degree. Adapted from Freeman (1978).

The last of the three measures, betweenness, assess the degree to which a node lies on the shortest path between two other nodes, and are able to funnel the flow in the network. In so doing, a node can assert control over the flow. Although this measure takes the global network structure into consideration and can be applied to networks with disconnected components, it is not without limitations. For example, a great proportion of nodes in a network generally does not lie on a shortest path between any two other nodes, and therefore receives the same score of 0.

Freeman's (1978) measures are only designed for binary networks. There have been a number of attempts to generalize Freeman's (1978) three node centrality measures to weighted networks (Barrat et al., 2004; Brandes, 2001; Newman, 2001). However, all these attempts have solely focused on tie weights, and not on the number of ties, which formed the basis of the original measures. First, degree was extended to weighted networks by Barrat et al. (2004) and defined as the sum of the weights attached to the ties connected to a node. An outcome of 10 could either be a result of 10 ties with a weight of 1, 1 tie with a weight of 10, or a combination between those two extremes. Second, the extensions of the closeness and betweenness centrality measures by Newman (2001) and Brandes (2001), respectively, rely on Dijkstra's (1959) shortest path algorithm. This algorithm defines the shortest path between two nodes as the least costly path. Brandes' (2001) and Newman's (2001) implementations suggest costs are only based on tie weights. In so doing, these three generalizations do not take into account a key feature, which the original measures were defined around, the number of ties (Freeman, 1978).

This raises a crucial question about the relative importance of tie weights to the number of ties in weighted networks. One can view the number of ties as more important than the weights, so that the presence of many ties with any weight might be considered more important than the total sum of tie weights. However, ties with large weights might be considered to have a much greater impact than ties with only small weights. This trade-off is the main motivation for this paper and drives the need for defining novel measures that enable researchers to set the relative importance between the number of ties and tie weights.

The rest of the paper is organized as follows. We start by proposing a generalization of degree centrality for weighted networks where the outcome is a combination of the number of ties and the tie weights. Then, in order to extend the closeness and betweenness centrality measures, we propose a generalization of shortest distances for weighted network that takes into account both the number of intermediary nodes and the tie weights. Subsequently, we suggest how the closeness and betweenness measures can take advantage of this generalized shortest distance algorithm. In Section 4, we evaluate the benefits of the proposed measures and explore the trade-off further by applying the degree measure to the well-known EIES dataset (Freeman and Freeman, 1979). In particular, we conduct a sensitivity analysis of the relative importance between the number of ties and the tie weights. Finally, we conclude with a discussion on the measures and various levels of the tuning parameter.

2. Degree

Freeman (1978) asserted that the degree of a focal node is the number of adjacencies in a network, i.e. the number of nodes that the focal node is connected to. Degree is a basic indicator and often used as a first step when studying networks (Freeman, 2004; McPherson et al., 2001; Wasserman and Faust, 1994). To formally describe this measure and ease the comparison among the different measures introduced in this paper, this measure can be formalized as:

$$k_i = C_D(i) = \sum_{j}^{N} x_{ij} \tag{1}$$

where i is the focal node, j represents all other nodes, N is the total number of nodes, and x is the adjacency matrix, in which the cell x_{ij} is defined as 1 if node i is connected to node j, and 0 otherwise.

Degree has generally been extended to the sum of weights when analyzing weighted networks (Barrat et al., 2004; Newman, 2004; Opsahl et al., 2008), and labeled *node* strength. This measure has been formalized as follows:

$$s_i = C_D^w(i) = \sum_{j}^{N} w_{ij} \tag{2}$$

where w is the weighted adjacency matrix, in which w_{ij} is greater than 0 if the node i is connected to node j, and the value represents the weight of the tie. This is equal to the definition of degree if the network is binary, i.e. each tie has a weight of 1. Conversely, in weighted networks, the outcomes of these two measures are different. Since node strength takes into consideration the weights of ties, this has been the preferred measure for analyzing weighted networks (e.g., Barrat et al., 2004; Opsahl et al., 2008). However, node strength is a blunt measure as it only takes into consideration a node's total level of involvement in the network, and not the number of other nodes to which it connected. To exemplify this, node A and node B have the same strength in Fig. 2, but node B is connected to twice as many nodes as node A, and is therefore, involved in more parts of the network. Since degree and strength can be both indicators of the level of involvement of a node in the surrounding network, it is important to incorporate both these measures when studying the centrality of a node.

In an attempt to combine both degree and strength, we use a tuning parameter, α, which determines the relative importance of

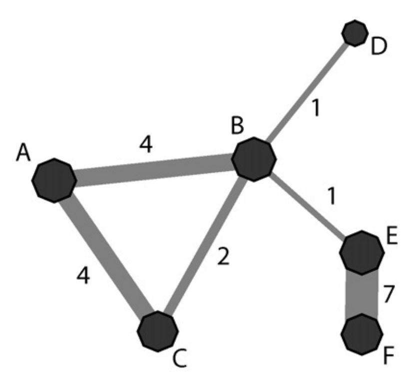

Fig. 2. A network with 6 nodes and 6 weighted edges. The size of the nodes correspond to the nodes' strength.