Discrete Applied Mathematics 186 (2015) 147–157

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Group centralization of network indices

Matjaž Krnc [a,*], Riste Škrekovski [b]

[a] *Institute of Mathematics, Physics and Mechanics, Jadranska 21, 1000 Ljubljana, Slovenia*

[b] *Department of Mathematics, University of Ljubljana, 1000 Ljubljana, Slovenia*

ARTICLE INFO

Article history:
Received 31 October 2013
Received in revised form 22 August 2014
Accepted 6 January 2015
Available online 7 February 2015

Keywords:
Freeman centralization
Group centrality
Betweenness
Eccentricity
Degree

ABSTRACT

In social network analysis various centrality indices are introduced to quantify importance of nodes in networks. Group centrality indices, introduced in 1999 by Borgatti and Everett, measure the importance of groups of nodes in networks. While centrality measures compare the importance of different nodes within a graph, the associated notion of *centralization*, as introduced in 1979 by Freeman allows us to compare the relative importance of nodes within their respective graphs. In this paper, we study the notion of group centralization with respect to eccentricity, degree and betweenness centrality measures. For groups of size 2, we determine the maximum achieved value of group eccentricity and group betweenness centralization and describe the corresponding extremal graphs. For group degree centralization we do the same with arbitrary size of group. We conclude with posing few open problems.

1. Introduction

For many decades in social science research, social networks have been the subject of study. With the rapid growth of Internet and World Wide Web in recent years, many large-scale online-based social networks appeared (including Facebook, Linkedin), and many large-scale social network data, such as co-authorship networks, become easily available online for analysis [17,18,10].

A social network is typically represented as a graph, where individual persons or nodes are represented as vertices, and the relationships between pairs of individuals as edges. In the paper, we will therefore freely interchange terms vertex/node and graph/network, without any meaningful difference. *Centrality* is an important concept in studying social networks [19,14]. We can imagine centrality as a measure of how central is the position of an individual (or a small group) within a network.

Various vertex-based measures of the centrality have been proposed to determine the relative importance of a vertex within the graph. Among measures of centrality, some of widely used in network analysis are: degree centrality, betweenness centrality, closeness centrality, eccentricity centrality, Google PageRank, eigenvector centrality, Katz centrality, Alpha centrality, and others. For detailed definitions and discussion on various centrality indices, we refer the reader to [8,1,2,16,20,22].

In his study, Freeman [14] realized that despite all defined vertex-centrality indices, there was a need for *graph centrality* measure based on differences in point centrality. He defined a centralization index that can be used in combination with any vertex-centrality to determine to what extent some vertex in network stands out from others in terms of given centrality index. Furthermore he used this approach to compare different graphs, depending on their highest centralization scores.

* Corresponding author.
E-mail addresses: matjaz.krnc@gmail.com (M. Krnc), skrekovski@gmail.com (R. Škrekovski).

http://dx.doi.org/10.1016/j.dam.2015.01.007
0166-218X/© 2015 Published by Elsevier B.V.

In the same article Freeman remarked that the centralizations for degree centrality, betweenness centrality and closeness centrality obtain its maximum if and only if G is a star. The statement was later proved in detail by Everett, Sinclair and Dankelmann [13]. In this article we determine some graphs that maximize group centralization with respect to eccentricity, degree and betweenness centrality measures.

To find extremal graph and/or maximize subset inside from algorithmical point of view can be a time consuming problem. In 2001 Brandes [6] improved algorithm for calculating betweenness centrality to $O(nm)$ and later in 2008 [7] extended his algorithm to group betweenness and other similar centralities. There are also some efficient heuristics and greedy approaches that can find vertices or groups that are sub-optimal in terms of various centrality measures, see Puzis, Dolev et al. [21,9].

In 1999, Everett and Borgatti [5] introduced the concept of *group centrality* which enables researchers to answer questions such as "how central is the engineering department in the informal influence network of this company?" or "among middle managers in a given organization, which are more central, the men or the women?" With these measures we can also solve the inverse problem: given the network of ties among organization members, how can we form a team that is maximally central? In [5], the authors introduced group centrality for measures of degree, closeness and betweenness centrality, which we use in this paper. In 2006, Borgatti introduced important group centrality measure (usually called *KPP*) that is motivated by *key players problem* (see [4]). In his paper he focused on finding a set of vertices for the purpose of optimally diffusing something through the network by using selected vertices as seeds, or for maximally fragmenting the network by removing the key nodes. Interestingly, Borgatti claims that previously mentioned group closeness and betweenness are not proper tools to define KPP centrality. He therefore used tools like graph fragmentation and information entropy to define KPP centrality.

Several more concepts of vertex centrality with respect to some subset of vertices have been introduced throughout last decade. In 2003, Smith and White [23] introduced a measure called *personalization* that shows, how central an individual is according to given subset R (group of important people) in given social network. In 2005, *subgraph centrality* has been introduced by Estrada and Rodríguez-Velázquez [11], and characterizes the participation of each node in all subgraphs in a network, which is calculated from the spectra of the adjacency matrix of the network. In the same year, Everett and Borgatti in [12], introduced another measure (i.e. *core centrality*), where they evaluate the extent to which a network revolves around a core group of nodes. Finally, very recently Bell [3] introduced the concept called *subgroup centrality*, where centrality (of one vertex) is calculated only on restricted set of vertices. Let us remark that all four mentioned centralities in principle measure importance of an individual vertex (with respect to some conditions) and are different from group centrality, proposed in [5].

Knowing all those group centrality measures it is natural to ask how much some choice of central group stands out from all other groups of the same cardinality (with respect to given group centrality index). Following Freeman's approach, we define group centralization notion in Section 2 and discuss it further in later sections.

In the sequel, we will use the following notion. Denote by $\mathcal{G}_n$ the family of non-isomorphic connected graphs on n vertices. Notice that when we consider a graph G, we usually assume $G \in \mathcal{G}_n$. A *star graph* S_n is a tree on $n+1$ vertices, with one vertex of degree n and n leaves. We will use $N(v)$ as a set of vertices in the neighborhood of v. As we deal with group centralization, by $C \subseteq V(G)$ we always denote the group we consider, and in addition we assume c is the size of C, i.e. $c = |C|$. Since $C = V(G)$ always trivially produces zero centrality (and therefore centralization), we will always assume $c < n$. At last, the *distance* from a vertex $x \in V(G)$ to a set of vertices $C \subseteq V(G)$ is defined by $d(x, C) = \min_{x \in C} \{d(x, c)\}$.

The paper is structured as follows. In Section 2 we introduce group centralization notion for arbitrary centrality index, and briefly describe its origin. In Section 3, we consider degree group centralization, and characterize extremal-pairs for graph family $\mathcal{G}_n$. In Section 4 we deal with eccentricity group centralization in the same graph family for groups of size 2 and describe the corresponding extremal graphs. In Section 5, we then do similar for betweenness group centralization. We conclude with posing few open problems in Section 6.

2. Group centralization

In many real life networks, it is intuitively clear, that some nodes are more important than others. Also some graphs are more depending on the most central vertices than others. While centrality measures compare the importance of a node within graph, the associated notion of *centralization*, as introduced by Freeman [14] allows us to compare the relative importance of nodes within their respective graphs. He proposed a very general approach with which the centralization of a graph G can be calculated. A clique where every vertex is connected to every other vertex is clearly not very centralized; on the other hand, the star topology, in which only one vertex v is connected to all others and all other vertices are only connected to v is a centralized graph. Thus, one would expect a star to have greater centralization than clique. In a network G, given a centrality index $X : V(G) \to \mathbb{R}$, the *centralization* of a node v is given by

$$X_1(G, v) = \sum_{u \in V(G)} (X(v) - X(u)) . \tag{1}$$

Following Freeman's idea, group centralization can be naturally generalized as a measure of how central its most central set of size c is in relation to how central all the other sets of the same cardinality are. Now we state this formally.