Physics Letters A 374 (2009) 126–130

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Measuring the flow of information among cities using the diffusion power

B.A. Mello [a,*], L.H. Batistuta [b,c], R. Boueri [b,c], D.O. Cajueiro [d,e]

[a] *Institute of Physics, University of Brasilia, DF, 70910-900, Brazil*
[b] *Institute of Applied Economic Research (IPEA), Brasilia, DF, Brazil*
[c] *Department of Economics, Catholic University of Brasilia, DF, Brazil*
[d] *Department of Economics, University of Brasilia, DF, 70910-900, Brazil*
[e] *National Institute of Science and Technology for Complex Systems, Brazil*

ARTICLE INFO

Article history:
Received 17 July 2009
Received in revised form 22 October 2009
Accepted 23 October 2009
Available online 28 October 2009
Communicated by C.R. Doering

PACS:
89.75.-k
89.65.-s
89.70.Hj

Keywords:
Centrality measures
Complex networks
Diffusion power
Domination power
Page rank

ABSTRACT

In this Letter, we define here the so-called *diffusion power* – an extension of the dominance power, which considers the interaction between neighbors of higher orders. Using this measure, we analyze the centrality of cities in two networks of the flow of information among these cities, namely a network of calls among the cities and a network of radio stations. Finally, we explain the centralities of the cities evaluated using the diffusion power in terms of the specific characteristics of the cities that belong to the network.

1. Introduction

Main stream research in the statistical physics field has been to characterize the dynamics of systems such as airports [1], financial institutions [2–4], web pages [5], flow of information networks [6,7], information city networks [8–11], social networks [12–14] and chaotic interacting functions [15] that may be described by complex weblike structures. Several comprehensive reviews on this subject are now available [16–18].

In the literature of complex networks, one of the most important concepts called *centrality* is used to define the relative importance of a node in a network. This concept can be defined in several different ways. For a typical undirected unweighed network, the simplest definition is the so-called degree of a node.[1] Other popular definitions such as *closeness centrality* [19], which is the inverse of the average distance from one node to all other nodes, *graph centrality* [20], which is the inverse of the maximal distance from one node to all other nodes and *efficiency* [21], which is the average of the inverse of the distance from one node to all other nodes, are based on the characteristic path length of a node. Another important concept in this context is the so-called *betweenness centrality* [22,23] that counts[2] the number of times that a node lies in the path between the others. Some of these measures built for undirected and unweighted networks may be generalized for the case of directed and weighted networks [25,4]. For instance, if one can associate weights to the edges of a network, the degree of a node can be generalized to be the *in-strength* and *out-strength* of a node. While the former is the sum of the weights of the edges that arrive to a given node, the later is the sum of the weights of the edges that leave a given node. Furthermore, if one is dealing with a geographic (spatial) network where the distances between the nodes are previously defined, closeness centrality, graph centrality and efficiency may be trivially extended for the case of directed and weighted networks.

* Corresponding author.
E-mail addresses: bernardo@fis.unb.br (B.A. Mello), rogerio.boueri@ipea.gov.br (R. Boueri), danielcajueiro@unb.br (D.O. Cajueiro).

[1] This concept is a kind of first order approximation to the centrality of a node, since it considers only the neighbors of first order.

[2] The way that the number of times that a path lies in the path between the others is counted may vary depending on the definition. For instance, one may consider any path or only the shortest paths between two nodes. A discussion about this topic may be found in [24].

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2009.10.062

An interesting revision of measures of centrality may be found in [26].

A constraint that usually arises is when the distances between the nodes are not previously defined, but only the weights of the edges of a given network. In this situation, the issue is how to define the distance from the previously defined weights. A path to circumvent this problem is to use, for instance, in-strength or out-strength measures defined above or the so-called *dominance power* [27] of a node, which is a measure of the influence of a node in all other nodes of the network relative to the influence of all other nodes (this measure will be precisely defined in Section 2). However, as in the case of the strength measure, this measure also only considers the neighbors of the first order.

In this Letter, we analyze the centrality of cities in two networks of the flow of information among these cities, namely a network of phone calls among the cities and a network of radio stations. Since we do not have a previously defined distance among the nodes of this network, we define here the so-called *diffusion power* – an extension of the dominance power, which considers the interaction between neighbors of higher orders. Furthermore, we explain the centralities of the cities evaluated using the diffusion power in terms of the specific characteristics of the cities that belong to the network.

The subject has some resemblance with the topic of Opinion Formation, since, in that field, there is information flow among agents [28] and it may depend on an influence network [29].

The reminder of this Letter is structured in the following way. Section 2 revises the measure known as the dominance power and defines the concept of diffusion power. Section 3 describes the process to construct the networks. Section 4 presents the results of the application of the diffusion power in the evaluation of the centrality of the cities that belong to the networks. Finally, Section 5 presents the main conclusions of this Letter.

2. Diffusion power

In order to define the diffusion power, we first revise the measure of centrality known as dominance power introduced by [27]. For a directed weighted network, the dominance power measures the influence of a node i on its neighbors normalized by the influence of all other nodes of the network on these neighbors as given by equation

$$\beta(i) = \sum_{j \neq i} \frac{w_{ij}}{s_{in}(j)} \tag{1}$$

where w_{ij} is the weight of the edge that comes from i and goes to j and $s_{in}(j)$ is the in-strength of node j given by

$$s_{in}(j) = \sum_k w_{kj}. \tag{2}$$

As already stressed, since this measure defines centrality considering only the neighbors of first order, it does not measure the effect of the propagation of the influence of a node by its neighbors.

Therefore, we extend this concept accounting for the propagation of the higher order neighbors as given by equation

$$D(i) = \sum_{j \neq i} \frac{[1 + f D(j)] w_{ij}}{s_{in}(j)} \tag{3}$$

where f is a free parameter defined in the interval $[0, 1)$ that measures the effects of higher order interaction among the nodes of a network. Furthermore, f defined in this interval ensures that the solution of the linear system defined in Eq. (1) may be found iteratively and it will converge to a fixed point [30]. Finally, if $f = 0$, the diffusion power is equal to the dominance power.

The idea proposed here to evaluate the diffusion power is quite similar to the idea of the Google founders [31] used to calculate the page rank:

$$P(i) = (1 - d) + d \sum_{j \neq i} w_{ji} \frac{P(j)}{s_{out}(j)}. \tag{4}$$

In the page rank expression above, w_{ji} is 1 if there is a link pointing from page j to page i, and zero otherwise.

The damping factor (d) used in page rank represents the probability that a surfer who reaches a given page clicks in a link on that page, moving to the next page. In our model, the damping factor is the probability that an information received by a node is forwarded to another node. There is not a consensus about the value that should be used for the damping factor, but the original paper from Brin and Page [31] suggested using 0.85.

A detailed discussion of the page rank may also be found in [26]. It is also worth remarking that the page rank has been extended in several directions. A revision of these attempts may be found for instance in [32].

3. Description of the networks data

In this Letter, by measuring the centrality among the cities, we are trying to find the most influential cities. In particular, we consider here two different networks of Brazilian cities, namely an undirected weighted network where the weight is the number of phone calls from one city to the other and a directed network where the weight is the number of radio stations of a city that reaches the other city.

3.1. The network of calls among the cities

Although there are few examples of networks of calls considered in the literature such as [33,34], we believe that the idea of using a network of call in order to identify the most influential cities is new.

The network considered here is based on a data file provided by the Brazilian Agency of Telecommunications with information about the number of calls between two cities and the total number of minutes of the calls between two cities. The data was assembled from all phone calls between wired phones located in one of the 5555 Brazilian municipalities which have access to this service during the month of September of 2008. As one may note, this data allowed to build a directed network. However, we decided to build an undirected network since in a phone call the information flows in both directions. Furthermore, instead of building the weights of the network based on the number of minutes, we decided to consider the number of calls, since most relevant information is given in few minutes.

3.2. The network of radio stations

As far as we know, a network based on radio stations is a totally new idea. In Brazil, mainly in the countryside, radio stations still have a fundamental role in the dissemination of information between close cities.

The network of radios was built here based on a data file provided by the Brazilian Agency of Telecommunications and the Ministry of Telecommunications. Two cities are connected in this directed and weighted network if at least one radio station from one city reaches the other city. The network is directed and the weight of the connection is used to count the number of the stations of the source city that reaches the target city.