

ESTUDOS SOBRE NOVA MEDIDA DE CENTRALIDADE EM GRAFOS: A CENTRALIDADE EM CAMADAS

Luana Carneiro Brandão

Mestrado em Engenharia de Produção – Universidade Federal Fluminense
Rua Passo da Pátria, 156, São Domingos, 24210-240, Niterói, RJ
luanabrandao@id.uff.br

Renata Raposo Del-Vecchio

Departamento de Análise do Instituto de Matemática – Universidade Federal Fluminense
Rua Mário Santos Braga, Campus Valonguinho, Centro, 24020-140, Niterói, RJ
renata@vm.uff.br

João Carlos Correia Baptista Soares de Mello

Departamento de Engenharia de Produção – Universidade Federal Fluminense
Rua Passo da Pátria, 156, São Domingos, 24210-240, Niterói, RJ
jcsmello@pesquisador.cnpq.br

Resumo

O presente trabalho busca estudar e desenvolver a medida de centralidade em grafos, preliminarmente introduzida em 2011, denominada centralidade em camadas. Essa nova centralidade ordena os elementos de uma rede, com base na quantidade de conexões, de cada elemento, com vizinhos relativamente influentes. Tal medida apresenta, assim, semelhanças com as centralidades tradicionais de informação (*degree centrality*) e de autovetor (*eigenvector centrality*). Em comparação com a primeira, a nova centralidade apresenta-se mais completa, por considerar o nível de influência dos vizinhos, ao passo em que contorna as dificuldades associadas à aplicação da centralidade de autovetor. Neste estudo, são apresentados um exemplo numérico e um estudo de caso, para ilustrar tanto a nova metodologia, quanto as comparações entre as três medidas de centralidade em grafos.

Palavras-chave: Centralidade em grafos; Centralidade em camadas; Análise comparativa de centralidade.

Abstract

The present work studies and develops the graph centrality measurement called layer centrality, which was preliminarily introduced in 2011. This new centrality ranks the elements of a given network, based on the amount of connections between each element and their relatively influential neighbours. Hence, the new methodology presents similarities with traditional centrality measurements, namely, degree and eigenvector centralities. With regard to the first, the layer centrality is more complete, since it accounts for the relative influence of neighbours, whilst overcoming the difficulties associated with the application of eigenvector centrality. In this study, we present a numerical example and a case study to illustrate the new methodology, as well as the comparative analysis between the centrality measurements.

Keywords: Graph centrality; Layer centrality; Comparative analysis of centrality.

1. INTRODUÇÃO

O estudo de redes complexas em diversas áreas, como ciências sociais, biologia, economia, logística, etc., pela técnica de Pesquisas em Redes, cresceu exponencialmente nas duas últimas décadas. Tal crescimento se deve ao fato de que essa técnica permite a estruturação de diversos tipos de sistemas, os quais tendem a evoluir sem planejamento ou controle centralizados, auxiliando, ainda, a compreensão de suas propriedades, vulnerabilidades e modos de falhas (Ercsey-Ravasz e Toroczkai, 2010).

Em tais estudos, as medidas de centralidade em grafos tornam-se fundamentais, pois descrevem a importância de cada elemento em relação à rede, com base em suas conexões com outros elementos (Ercsey-Ravasz *et al.*, 2012). Existem, contudo, diversas formas de calcular essa importância e, conseqüentemente, diversas medidas de centralidade, cuja escolha, no caso concreto, deve ser pautada pelo tipo de rede e pela forma em que se processa o seu fluxo (Borgatti e Everett, 2006).

Uma medida bastante aplicada em análises variadas é a centralidade de autovetor (*eigenvector centrality*), e.g., Smith e Timberlake (2001), Choi *et al.* (2006), Ma e Timberlake (2008), Mahutga *et al.* (2010), Meghanathan (2014). Proposta por Bonacich (1972), essa medida considera, no cálculo da centralidade de determinado elemento, a importância dos elementos a ele conectados. Entretanto, tal centralidade depende de cálculos complexos, além de se mostrar inadequada para diversos tipos de redes, e.g., redes grandes ou com vínculos fortes (Neal, 2011; Taylor *et al.*, 2002).

Para contornar as supramencionadas dificuldades, o presente trabalho desenvolve uma nova centralidade, preliminarmente introduzida por Bergiante *et al.* (2011), denominada centralidade em camadas, a qual avalia a importância de cada elemento da rede, com base na sua quantidade de conexões com elementos relativamente influentes. Tal centralidade apresenta certa semelhança com a centralidade de autovetor, porém sem as dificuldades matemáticas ou restrições à aplicação a ela associadas. Além disso, por ser mais simples, o método também se torna mais natural e intuitivo.

De outro modo, a centralidade em camadas também apresenta certas características semelhantes à centralidade de informação (*degree centrality*), proposta por Freeman (1979), a qual avalia simplesmente a quantidade de ligações diretas de cada elemento da rede. Dessa forma, o presente trabalho estuda a nova centralidade, comparando-a com as medidas tradicionais de referência, i.e., a centralidade de informação e de autovetor. Este estudo apresenta, ainda, um exemplo numérico e um estudo de caso, para ilustrar tanto a nova metodologia, quanto as análises comparativas com as demais medidas de centralidade.

A próxima seção apresenta o referencial teórico, mais especificamente, definições básicas em Teoria de Grafos, assim como breves explicações das centralidades tradicionais que atuam como referência à nova metodologia. A terceira seção exibe uma revisão de algumas aplicações, existentes na literatura, com medidas de centralidades. A quarta seção explica e desenvolve a centralidade em camadas, analisando, ainda, um exemplo numérico ilustrativo. A quinta seção consiste em uma aplicação das medidas de centralidade deste trabalho para analisar a importância de portos no sistema portuário brasileiro. Por fim, a última seção apresenta as conclusões do presente.

2. REFERENCIAL TEÓRICO

Nesta seção, serão explicadas brevemente as medidas de centralidade tradicionais que servem de referência para o presente estudo, ou seja, as centralidades de informação (*degree centrality*) e de autovetor (*eigenvector centrality*). Entretanto, como o estudo de centralidades, é

fundamentado na Teoria dos Grafos (Freeman *et al.*, 1991), serão apresentadas, primeiramente, definições básicas da Teoria dos Grafos.

2.1. DEFINIÇÕES BÁSICAS EM TEORIA DOS GRAFOS

Um grafo é identificado por um par $G = (V, E)$, em que V é o conjunto de n vértices v_i e E é o conjunto de m arestas (do inglês, *edges*) $[v_i, v_j]$, sendo cada aresta formada por um par de vértices de V , ou seja, $v_i, v_j \in V$ (Bondy e Murty, 2008). Acrescente-se que os vértices conectados por arestas podem ser denominados vizinhos.

O presente trabalho utiliza, no exemplo numérico proposto, um grafo simples, i.e., que não apresenta mais de uma aresta entre dois vértices (arestas paralelas) ou arestas com vértices coincidentes (denominadas laços), ou seja, $[v_i, v_i]$. O grafo no exemplo apresentado também é não orientado, pois não há direcionamento entre dois vértices conectados por uma aresta. Outra característica do grafo em tela é a sua não valoração, i.e., todas as arestas apresentam valor unitário.

Em resumo, o grafo utilizado no presente estudo é formado por vértices, entre os quais existe, ou não, uma única aresta, de valor unitário e sem direção. Entretanto, não há, a princípio, impedimento para a aplicação da nova metodologia em outros tipos de grafos, e.g., com arestas paralelas, valoradas e/ou orientadas.

Ademais, utiliza-se apenas grafos conexos, em que existe pelo menos um caminho que liga quaisquer dois vértices. Matematicamente, define-se um caminho como um grafo $P = (V, E)$ em que $V = \{v_1, v_2, \dots, v_n\}$ e $[v_i, v_{i+1}] \in E$ para $1 \leq i < n$ (Bondy e Murty, 2008). Em palavras, um caminho é uma sequência de vértices, em que cada vértice é ligado ao seguinte por uma aresta.

A matriz de ordem n , $A = A(G) = [a_{ij}]$, para a qual $a_{ij} = 1$, se $[v_i, v_j] \in E$, e $a_{ij} = 0$, se $[v_i, v_j] \notin E$, é denominada matriz de adjacência de G . Para grafos não direcionados, $a_{ij} = a_{ji}$, sendo a matriz, portanto, simétrica. O polinômio característico associado à matriz de adjacência é dado por $p_G = \det(A(G) - \lambda I)$, cujas raízes $\lambda_1, \lambda_2, \dots, \lambda_n$ são os autovalores de G . O maior autovalor deste polinômio é o raio espectral da matriz $A(G)$, sendo denominado índice do grafo G (Del-Vecchio *et al.*, 2009).

2.2. MEDIDAS TRADICIONAIS DE CENTRALIDADE

A centralidade de informação incorpora a definição de centralidade proposta por Freeman (1979). Essa medida de centralidade corresponde ao número de interações ou conexões diretas que cada elemento estabelece com os demais, o que equivale ao grau do vértice em questão. Trata-se de um aspecto importante referente à posição estrutural do vértice (Ruhnau, 2000). Em redes de informações, por exemplo, tal medida traduz a capacidade de receber diretamente uma informação (Del-Vecchio *et al.*, 2009).

Utilizando-se a matriz de adjacência do grafo, obtém-se a centralidade de informação C_D de um vértice v_i , somando-se os valores da sua linha (ou coluna), como em (1).

$$C_D(v_i) = d(v_i) = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji}, \text{ onde } v_i \in V \quad (1)$$

Por outro lado, a importância de um elemento, em geral, depende não só da quantidade de conexões diretas, mas também do nível de influência de seus vizinhos. Esse conceito é incorporado pela centralidade de autovetor, proposta por Bonacich (1972), em que se adiciona, à centralidade de um elemento, a centralidade dos elementos com ele conectados. Tal medida corresponde, para cada elemento, à combinação linear das centralidades de seus vizinhos (Bonacich e Lloyd, 2001).

Utilizando-se a matriz de adjacência do grafo, a centralidade de autovetor de um vértice v_i é dada por x_i , que satisfaça à equação (2), ou $Ax = \lambda x$, em notação matricial.

$$\lambda x_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \quad (2)$$

Em (2), as soluções de λ correspondem aos autovalores da matriz de adjacência, enquanto as soluções de (x_1, \dots, x_n) correspondem aos autovetores da matriz. Para calcular a centralidade dos vértices do grafo, utilizam-se as coordenadas do autovetor positivo unitário, associadas ao maior autovalor da matriz (Bonacich, 1987).

Existem, ainda, outras medidas tradicionais de centralidade, algumas mais difundidas que outras. A centralidade de intermediação (*betweenness centrality*), proposta por Freeman (1977), por exemplo, é definida como a parcela de todos os caminhos geodésicos (menores caminhos) que passa por cada vértice (ou aresta, ou ainda por determinado subgrafo). Tal medida é implementada em todos os principais *softwares* de análise de redes e aplicada em diversos estudos (Brandes, 2008), sobretudo em transportes. Já a centralidade de proximidade (*closeness centrality*), proposta por Freeman (1971), também tradicional, porém menos popular, mede a distância geodésica total de um vértice a todos os outros na rede.

Apesar da grande quantidade de medidas de centralidade existentes na literatura, o presente trabalho utiliza como referência para o estudo da centralidade em camadas, apenas as medidas tradicionais que apresentam características em comum com a nova metodologia, ou seja, as centralidades de informação e de autovetor.

3. REVISÃO DA LITERATURA

Medidas de centralidade são amplamente utilizadas para analisar diversos tipos de rede, em áreas de conhecimento distintas. Hussain (2007) usou centralidade em grafos, dentre elas, a centralidade de informação, para desestabilizar redes terroristas, através da localização dos elementos-chave na rede e da previsão de um caminho que demonstre a vulnerabilidade da rede. Jiang *et al.* (2014) aplicaram centralidade de autovetor em esportes, no intuito de avaliar o desempenho de cada time, com base na dificuldade dos jogos vencidos, a qual dependeria da força dos seus oponentes. O objetivo maior do trabalho foi determinar os melhores treinadores universitários dos últimos séculos, para basquete, futebol e baseball masculinos.

Del-Vecchio *et al.* (2009) aplicaram medidas tradicionais de centralidade, incluindo a de informação e de autovetor, para extrair informações a respeito da estrutura interna do mercado de valores brasileiro, identificando ainda os agentes líderes no mercado acionário. Gonçalves *et al.* (2009) aplicaram medidas de centralidade, dentre as quais também constavam as de informação e de autovetor, para criar um ambiente urbano mais harmonioso e propiciar a mobilidade sustentável. O estudo se baseou em bairros do município do Rio de Janeiro, mas poderia ser aplicado em diversas metrópoles do Brasil.

Ademais, diversos autores propuseram novas medidas de centralidade, para que o conceito de importância em rede se adequasse ao objeto de estudo. Nesse sentido, Kirkland (2009) explorou uma noção diferente de centralidade em vértices, aplicando-a a redes de alimentos. Page *et al.* (1999) propuseram uma medida de centralidade, denominada *Page Rank*, para diversas aplicações relacionadas à internet, sendo, inclusive, utilizada pelo Google para ordenar os resultados de buscas. Joyce *et al.* (2010) desenvolveram uma nova medida de centralidade para redes cerebrais, denominada *leverage centrality*, que avalia o grau de dependência dos vizinhos a cada elemento. Assim como no presente estudo, aqueles autores utilizaram, como referência, medidas tradicionais de centralidade, dentre elas, as de informação e de autovetor, comparando-as à nova medida proposta.

Considerando as diversas medidas de centralidade existentes, Borgatti (2005) classificou fluxos de redes, baseando-se nos diferentes tipos de trajetória e métodos de distribuição, para identificar as medidas de centralidade que melhor se aplicam a cada tipo de fluxo de rede. O autor constata que as medidas pré-definidas de centralidade devem ser aplicadas somente para o tipo específico de fluxo para o qual foram projetadas e que, quando aplicadas a outros tipos de fluxo, levam a resultados incoerentes. Conclui-se, ainda, que as medidas de centralidade mais comuns não são apropriadas aos tipos de fluxos em que normalmente se está interessado.

4. CENTRALIDADE EM CAMADAS

Embora seja fundamentada em um conceito de importância, apropriado para diversos tipos de redes, não obstante, a centralidade de autovetor apresenta diversas dificuldades de aplicação (Neal, 2011). Além da complexidade dos cálculos, tal medida não pode ser aplicada a redes com agrupamentos ou cliques (Derudder *et al.*, 2003; Derudder e Taylor, 2005), e também não fornece resultados úteis quando as redes são grandes ou apresentam vínculos fortes (Taylor *et al.*, 2002). Além disso, para grafos direcionados, a aplicação dessa medida de centralidade é ainda mais restrita (Bonacich e Lloyd, 2001), haja vista a existência, nesses casos, de autovalores complexos, associados a autovetores com coordenadas também complexas (Wei *et al.*, 2011).

Dessa forma, o presente trabalho desenvolve uma nova metodologia de centralidade, preliminarmente introduzida por Bergiante *et al.* (2011), em que se avalia a importância de cada aeroporto para uma empresa do setor aéreo brasileiro. A medida proposta leva em consideração o nível relativo de influência dos vizinhos, no cálculo da centralidade de cada elemento, podendo ser aplicada aos mesmos tipos de redes avaliadas pelas centralidades de informação e de autovetor. Em relação à primeira, a nova medida apresenta-se mais completa, por levar em conta o nível de influência dos vizinhos de cada elemento, ao passo em que supera as dificuldades associadas à centralidade de autovetor, tornando-se ainda mais natural e intuitiva que esta última.

O novo método foi inspirado na técnica de Análise Envoltória de Dados (DEA – *Data Envelopment Analysis*) em Camadas, que ordena as unidades sob avaliação com base em suas influências nas camadas da fronteira de produção. A técnica de DEA em camadas foi proposta por Barr *et al.* (2000) e aplicada, por exemplo, por Gomes *et al.* (2009), que avaliam a sustentabilidade na agricultura da Amazônia Brasileira, para agrupar os fazendeiros em categorias de sustentabilidade distintas.

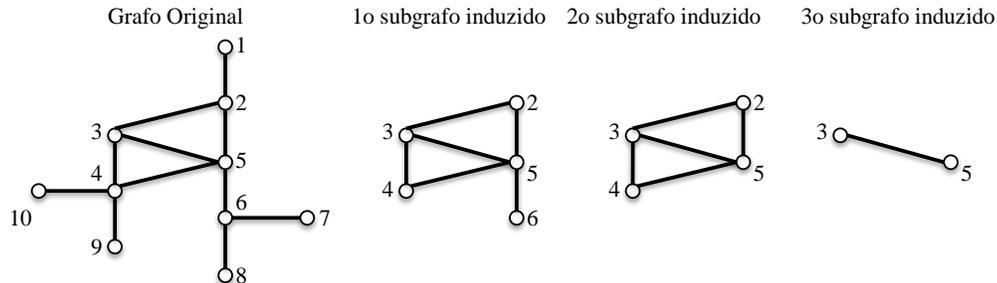
Quanto à metodologia propriamente dita, o primeiro passo da centralidade em camadas é identificar, na matriz de adjacências, o elemento com a menor soma das coordenadas na linha, e ordená-lo em último lugar, considerando-o, portanto, como o vértice de menor centralidade. Em seguida, excluem-se da matriz a linha e a coluna desse elemento. Com isso, elimina-se o vértice da matriz, e assim, o mesmo não pode ser ordenado outra vez, mas também se eliminam as conexões dos demais vértices com tal elemento.

Todo esse processo deve ser repetido, ordenando os vértices em ordem crescente de centralidade e excluindo as respectivas linhas e colunas, até o último vértice da matriz. Caso mais de um elemento apresente a mesma soma das coordenadas em suas linhas, ambos devem ser ordenados na mesma posição e ter as suas linhas e colunas excluídas da matriz simultaneamente. Ou seja, esses vértices ficam na mesma camada.

Em termos de grafos, essa metodologia consiste em identificar o(s) vértice(s) de menor grau e ordená-lo(s) em último lugar. Em seguida, cria-se um subgrafo induzido, em que todos os vértices, exceto aquele(s) com o menor grau, estão presentes. Na iteração seguinte, calculam-se os novos graus para o subgrafo formado, identificando-se, novamente, o(s) vértice(s) de menor grau, e removendo-o(s) para a formação de um novo subgrafo induzido. Esse processo é, assim, repetido até não existirem mais vértices a serem removidos.

No intuito de facilitar a compreensão da metodologia, apresenta-se, em seguida, um exemplo de grafo não direcionado e não valorado $G = (V, E)$, em que $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ e $E = \{[1,2], [2,3], [2,5], [3,4], [4,5], [4,9], [4,10], [5,6], [6,7], [6,8]\}$, o qual foi definido arbitrariamente. Na Figura 1, constam a representação desse grafo original e os subgrafos induzidos subsequentes, formados no decorrer da metodologia descrita anteriormente.

Figura 1 –Exemplo para ilustrar a metodologia da centralidade em camadas



No grafo original, os vértices de menor grau $d_0(i)$ são os vértices $i = 1, 7, 8, 9$ e 10 , com $d_0(i) = 1$, os quais são removidos na primeira iteração, sendo formado o primeiro subgrafo induzido. Neste, o vértice $i = 6$ apresenta o menor grau $d_1(i) = 1$, sendo, portanto, removido na segunda iteração, para a formação do segundo grafo induzido. Em seguida, os vértices $i = 2$ e 4 apresentam o menor grau $d_2(i) = 2$ e, por isso, são removidos na última iteração, para a formação do terceiro e último subgrafo induzido. A Tabela 1 apresenta a ordenação final para esse exemplo ilustrativo.

Tabela 1 –Ordenação final para o exemplo ilustrativo

| Ordem Final |
|-------------|
| 3,5 |
| 2,4 |
| 6 |
| 1,7,8,9,10 |

Os subgrafos formados em cada iteração podem ser associados às diferentes fronteiras de produção da técnica de DEA em camadas. Os subgrafos representam redes, formadas com parte dos elementos presentes na iteração anterior, com níveis diferentes de concentração, de modo semelhante às fronteiras de produção, formadas com parte dos elementos presentes na iteração anterior, com níveis diferentes de avanço tecnológico. Entretanto, as fronteiras de produção de técnica de DEA em camadas são formadas em ordem decrescente de avanço tecnológico, sendo a ordenação das unidades produtivas obtida de cima para baixo, enquanto os subgrafos da presente metodologia são formados em ordem crescente de concentração, sendo, portanto, a ordenação dos vértices obtida de baixo para cima.

A centralidade em camadas pode ser considerada intermediária entre a centralidade de informação e de autovetor, por se basear nos graus dos vértices, como ocorre na centralidade de informação, mas também considerar o nível de influência dos vizinhos, como na centralidade de autovetor. Mais especificamente, o primeiro vértice a ser eliminado, ou seja, o último elemento ordenado, depende somente dos graus dos vértices. Assim, a parte inferior da ordenação em camadas se assemelha bastante à ordenação da centralidade de informação.

Entretanto, os elementos conectados aos vértices eliminados perdem essas conexões na iteração subsequente. Assim, os elementos que apresentam graus elevados, por estarem conectados com elementos periféricos, perdem tais conexões durante as iterações iniciais e, portanto, perdem posições na ordenação em camadas, quando comparadas à ordenação de informação. Nessa situação, interpreta-se que tais elementos não são importantes para a propagação na rede.

Dessa forma, as posições iniciais da ordenação em camadas tendem a divergir da ordenação da centralidade de informação e, em certos casos, ainda não completamente estudados, a nova ordenação se assemelha à ordenação da centralidade de autovetor. Essa situação ocorre para o exemplo numérico fornecido anteriormente, e ilustrado na Figura 1, conforme será visto adiante.

Por outro lado, vale ressaltar que as centralidades em camadas e de autovetor são medidas distintas, pois avaliam o nível de influência dos vizinhos de maneira diferente. Enquanto a centralidade de autovetor calcula a combinação linear dos vizinhos de cada elemento (Bonacich e Lloyd, 2001), a centralidade em camadas mede o grau dos elementos com os vizinhos que ainda não foram eliminados do grafo por serem relativamente periféricos. Assim, apesar de suas ordenações serem semelhantes, em certos casos, para as posições iniciais, essa situação não é garantida.

Ademais, na centralidade em camadas, interpreta-se que, para elementos de graus menores, o nível de influência dos vizinhos é pouco relevante, justamente por serem comparativamente pouco numerosos. Entretanto, para elementos de graus mais elevados, o nível de influência dos vizinhos torna-se mais relevante. Essa distinção é feita, ainda, de maneira gradual, ou seja, o nível de influência dos vizinhos se torna cada vez mais relevante, conforme o aumento da centralidade em camadas dos elementos, i.e., conforme se desloca em direção ao início do *ranking*.

No intuito de ilustrar as características da centralidade em camadas, em comparação com as duas medidas tradicionais, foram calculadas, para o grafo apresentado anteriormente e ilustrado na Figura 1, as centralidades de informação e de autovetor; esta última, com o auxílio do programa UCINET (Borgatti *et al.*, 2002). Os resultados são apresentados na Tabela 2.

Tabela 2 –Medidas tradicionais de centralidade para o exemplo ilustrativo

| Vértice | Medidas de Centralidades | |
|---------|--------------------------|-----------|
| | Informação | Autovetor |
| 1 | 1 | 0.134 |
| 2 | 3 | 0.391 |
| 3 | 3 | 0.472 |
| 4 | 4 | 0.451 |
| 5 | 4 | 0.533 |
| 6 | 3 | 0.239 |
| 7 | 1 | 0.082 |
| 8 | 1 | 0.082 |
| 9 | 1 | 0.155 |
| 10 | 1 | 0.155 |

Assim, é possível comparar as ordenações obtidas pelas três medidas de centralidade, conforme a Tabela 3, a qual apresenta as posições de cada vértice, nos rankings das centralidades de informação, autovetor e camadas, para o grafo em tela.

Tabela 3 –Posições dos vértices nos rankings das centralidades para o exemplo ilustrativo

| Vértice | Posições nos Rankings das Centralidades | | |
|---------|---|-----------|---------|
| | Informação | Autovetor | Camadas |
| 1 | 6 | 8 | 6 |
| 2 | 3 | 4 | 3 |
| 3 | 3 | 2 | 1 |
| 4 | 1 | 3 | 3 |
| 5 | 1 | 1 | 1 |
| 6 | 3 | 5 | 5 |
| 7 | 6 | 9 | 6 |
| 8 | 6 | 9 | 6 |
| 9 | 6 | 6 | 6 |
| 10 | 6 | 6 | 6 |

Vale ressaltar que a centralidade em camadas não fornece uma medida propriamente dita, como as centralidades tradicionais. Entretanto, o decisor, em geral, está interessado com a ordenação dos elementos, ao invés de suas medidas e, portanto, essa desvantagem não costuma ser relevante para o decisor.

Observa-se que os cinco últimos vértices apresentam o mesmo posicionamento nas ordenações das centralidades de informação e em camadas, ou seja, os vértices 1, 7, 8, 9 e 10 ocupam o último lugar em ambos os *rankings*. Por outro lado, a ordenação dos demais vértices para a centralidade em camadas se aproxima mais à ordenação do autovetor. De fato, os dois primeiros colocados, os dois terceiros colocados e o quinto colocado são os mesmos para ambos os *rankings*, e distintos para o ranking de informação. A diferença entre os rankings de autovetor e em camadas para tais elementos ocorre porque os dois primeiros colocados estão empatados nesta última centralidade, mas encontram-se em posições distintas, na centralidade de autovetor. O mesmo ocorre com os dois terceiros colocados. Assim, esse exemplo confirma a ideia de a centralidade em camadas ser intermediária entre as duas centralidades tradicionais.

Além disso, vale comparar o vértice 7 com os vértices 9 e 10, no grafo original, ilustrado na Figura 1. Observa-se que os vértices 9 e 10 estão ligados ao vértice 4, muito mais influente, em qualquer das medidas de centralidade calculadas, do que o vértice 6, ao qual o vértice 7 está ligado. Essa distinção é feita pela centralidade de autovetor, sendo, porém, ignorada pela centralidade em camadas, como se pode observar pela Tabela 3. Dessa forma, confirma-se também a interpretação de que o nível de influência dos vizinhos para vértices com graus menores é menos relevante para a centralidade em camadas, em relação ao nível de influência dos vizinhos para vértices com graus mais elevados.

Acrescenta-se, ainda, que os resultados da centralidade em camadas podem ser interpretados de maneira mais natural e intuitiva do que os resultados da centralidade de autovetor. Em relação à primeira medida, basta uma análise, relativamente simples, da sequência de eliminação dos elementos, para compreender a ordenação final. Já em relação à medida tradicional, tal compreensão pode não ser trivial, como ocorre para o exemplo estudado, apesar da pequena quantidade de elementos. Afinal, não é fácil prever como a centralidade de autovetor ordena, por exemplo, os vértices 3, 4 e 5, antes do cálculo das suas medidas.

5. APLICAÇÃO NO SISTEMA PORTUÁRIO BRASILEIRO

As medidas de centralidade contempladas neste estudo podem ser aplicadas na avaliação de unidades portuárias brasileiras, incluindo Terminais de Uso Privativo (TUPs) e portos públicos, no intuito de avaliar a importância de cada unidade para o sistema portuário, no cenário atual. Para isso, considera-se o total da carga não containerizada transportada entre tais unidades, no período de janeiro a maio de 2013, disponível no Sistema de Informações Gerenciais de Acesso Público (SIG) da Agência Nacional de Transportes Aquaviários (ANTAQ).

Vale esclarecer que a carga containerizada não é considerada, por apresentar unidade de medida padrão diferente dos demais tipos de carga, não sendo trivial a junção de ambas. A carga containerizada é comumente medida em TEU (de *Twenty-Foot Equivalent Unit*, em inglês, ou unidade equivalente a 20 pés), enquanto os demais tipos de carga, quais sejam, carga solta, granel líquido e granel sólido, são medidos em toneladas.

A Tabela 4 apresenta os rankings das centralidades de informação, de autovetor e em camadas para algumas das 94 unidades portuárias que realizaram movimentação de carga não containerizada no período considerado, ordenadas pela centralidade de informação. Vale esclarecer que 48 unidades apresentaram centralidade de autovetor nula e por isso ficaram todos em último lugar (posição 47) no ranking do autovetor. Além disso, 10 unidades apresentaram a segunda menor medida de autovetor (0.001) e, por isso, ficaram todas na posição 37.

Tabela 4 – Rankings de centralidade para unidades portuárias brasileiras

| Unidade Portuária | UF | Informação | Autovetor | Camadas |
|--------------------------------------|----|------------|-----------|---------|
| Santos | SP | 1 | 3 | 5 |
| Vila do Conde | PA | 2 | 1 | 1 |
| Itaqui | MA | 3 | 6 | 7 |
| TUP Porto Trombetas | PA | 4 | 2 | 1 |
| Rio Grande | RS | 5 | 5 | 18 |
| Porto Velho | RO | 6 | 37 | 3 |
| Suape | PE | 7 | 8 | 6 |
| Fortaleza | CE | 8 | 12 | 8 |
| Rio de Janeiro | RJ | 9 | 7 | 11 |
| Paranaguá | PR | 10 | 4 | 10 |
| Vitória | ES | 11 | 9 | 15 |
| TUP Almirante Barroso | SP | 12 | 10 | 9 |
| Santarém | PA | 13 | 47 | 16 |
| TUP Hermasa Graneleiro | AM | 14 | 47 | 4 |
| TUP Madre de Deus | BA | 15 | 27 | 14 |
| Salvador | BA | 16 | 12 | 12 |
| Manaus | AM | 17 | 27 | 13 |
| São Francisco do Sul | SC | 18 | 34 | 25 |
| Belém | PA | 19 | 12 | 17 |
| Aratu | BA | 20 | 12 | 19 |
| TUP Almirante Tamandaré | RJ | 21 | 19 | 21 |
| Maceió | AL | 22 | 23 | 20 |
| Areia Branca | RN | 23 | 12 | 28 |
| Porto Alegre | RS | 24 | 19 | 27 |
| São Sebastião | SP | 25 | 23 | 23 |
| Terminais Interiores | - | 26 | 11 | 34 |
| ... | | | | |
| TUP Pecém | CE | 71 | 47 | 71 |
| TUP Terminal Aquaviário B. do Riacho | ES | 72 | 47 | 72 |
| TUP Porto Murucupi | PA | 73 | 47 | 73 |
| Salvaterra | PA | 74 | 47 | 74 |
| TUP GE Oil & Gás do Brasil | RJ | 75 | 47 | 75 |
| TUP Caulim da Amazônia | PA | 76 | 47 | 76 |
| TUP Ilha Redonda | RJ | 77 | 47 | 77 |
| Forno | RJ | 78 | 47 | 78 |
| TUP Bertolini Belém | PA | 79 | 47 | 79 |
| TUP de GNL da Baía da Guanabara | RJ | 80 | 47 | 80 |
| TUP Bianchini | RS | 81 | 47 | 81 |
| Fernando de Noronha | PE | 82 | 47 | 82 |
| Pelotas | RS | 83 | 47 | 83 |
| Monte Alegre | PA | 84 | 47 | 84 |
| Itaituba | PA | 85 | 37 | 85 |
| TUP Terminal Marítimo Inácio Barbosa | SE | 86 | 47 | 86 |
| Icoaraci | PA | 87 | 47 | 87 |
| TUP Navecunha | AM | 88 | 47 | 88 |
| TUP Teporti | SC | 89 | 47 | 89 |
| TUP Porto Itapoá | SC | 90 | 47 | 90 |
| TUP Ilha do Governador | RJ | 91 | 47 | 91 |
| TUP Tergasul | RS | 92 | 47 | 92 |
| TUP Brasfels | RJ | 93 | 47 | 93 |
| TUP Sanave | AM | 94 | 47 | 94 |

De maneira geral, os três rankings são muito semelhantes do centro para o final, mais precisamente para as últimas 55 unidades. Isso ocorre porque os portos com as menores medidas de centralidade de informação também apresentam baixa centralidade de autovetor. A principal diferença é que a centralidade de autovetor não discrimina essas unidades, o que não ocorre com as outras duas medidas de centralidade.

As duas primeiras posições nos rankings das centralidades em camadas e de autovetor são idênticas, e diferentes do ranking da centralidade de informação. Em outras palavras, esse estudo de caso confirma que a nova medida considera completamente a centralidade dos vizinhos para as primeiras unidades.

Ademais, as centralidades de autovetor explicam as diferenças entre os rankings das centralidades de informação e em camadas para diversas unidades. Particularmente, Santos, Itaqui, São Francisco do Sul e Plataforma Marítima perderam posições no ranking em camadas devido ao seu baixo valor relativo de autovetor, enquanto TUP Almirante Barroso ganhou posições no ranking em camadas devido à sua relativamente alta medida de autovetor.

Por outro lado, a centralidade de autovetor não explica as diferenças entre os dois outros rankings para unidades como Rio Grande, TUP Hermasa Graneleiro e Terminais Interiores. Rio Grande e Terminais Interiores perderam posições no ranking em camadas, apesar dos altos valores relativos de autovetor, enquanto TUP Hermasa Graneleiro subiu no ranking em camadas, apesar dos baixos valores relativos de autovetor.

Vale observar que Rio Grande ocupa a 5^a posição nos rankings de informação e de autovetor, mas ocupa o 18^o lugar no ranking em camadas. Analisando os dados originais, constata-se que 56% das conexões de Rio Grande ocorrem com unidades muito periféricas, quase que ligadas exclusivamente com esse porto, quais sejam, Estrela, Itajaí, TUP CPMC Guarba, TUP Mita, TUP Niterói, TUP Oleoplan, and TUP Santa Clara, e 80% com unidades que foram eliminadas comparativamente cedo na metodologia em camadas, o que justifica o seu mal posicionamento nesse ranking, relativamente aos demais rankings.

De maneira similar, Terminais Interiores ocupam a 26^a posição no ranking de informação, a 11^a posição no ranking de autovetor, mas a 34^a posição no ranking em camadas. Com base nos dados originais, observa-se que 37% de suas conexões ocorrem com TUP Almirante Soares Dutra, na 35^a posição do ranking em camadas, ou seja, com uma unidade que acabara de ser retirado da matriz de adjacências.

Por outro lado, TUP Hermasa Graneleiro encontra-se na 14^a posição do ranking de informação, porém apenas com uma unidade, o que justifica a centralidade de autovetor nula. Entretanto, essa unidade é Porto Velho, que ocupa a 3^a posição no ranking em camadas, o que explica a 4^a posição do TUP Hermasa Graneleiro no ranking em camadas.

Em outras palavras, a centralidade em camadas considera de fato a centralidade dos vizinhos, assim como a de autovetor. Porém, nem sempre essa nova medida comporta-se como um intermediário entre as duas medidas tradicionais. Particularmente, unidades que apresentam ligações com muitas unidades medianas tendem a ocupar posições inferiores no ranking em camadas, em relação ao ranking de autovetor. Por outro lado, unidades ligadas com poucas unidades bem posicionadas tendem a ocupar posições superiores no ranking em camadas, em relação ao ranking de autovetor.

Em relação aos resultados da Tabela 4 propriamente ditos, observa-se que Santos movimenta a maior quantidade de carga não containerizada diretamente com outras unidade nacionais. Entretanto, a maior parte dessa movimentação ocorre com unidades periféricas da rede portuária, e por isso Santos perde posições nos demais rankings. Mais precisamente, apenas 4% de suas movimentações ocorrem com um porto bem colocado no ranking (Vila do Conde), sendo as demais com portos colocados na nona posição em diante.

Apesar de o resultado de Santos nos rankings das centralidades de autovetor e em camadas parecer incoerente, uma vez que Santos é reconhecido como o maior porto brasileiro (ALESP, 2011; Previdelli, 2012), esclareça-se que os principais estudos na área reconhecem Santos como o maior porto brasileiro de contêiner (*Container Management*, 2007; *Containerisation International*, 2011, 2012, 2013, 2014). Como se analisa apenas carga não conteneurizada, o resultado da Tabela 4 não é incoerente com estudos da área, e sim complementa-os.

De maneira geral, nota-se que as unidades mais influentes no transporte nacional de cargas são portos públicos, e não TUPs. Além disso, pela análise dos dados originais, verifica-se que dentre as últimas 13 unidades, 12 (excetuando-se o TUP Teporti) não realizaram qualquer movimentação com portos estrangeiros, apresentando, efetivamente, pequena importância para o sistema portuário brasileiro.

6. CONCLUSÕES

O presente desenvolveu a centralidade em camadas, preliminarmente introduzida por Bergiante *et al.* (2011), a qual avalia a quantidade de conexões de cada elemento com outros elementos relativamente influentes. A nova metodologia apresenta características semelhantes à centralidade de informação (*degree centrality*) e à de autovetor (*eigenvector centrality*), podendo ser aplicada aos mesmos tipos de redes avaliadas por essas medidas tradicionais. Em comparação com a primeira medida, a nova centralidade apresenta-se mais completa, por levar em conta o nível de influência dos vizinhos, ao passo em que supera as dificuldades matemáticas e restrições à aplicação associadas à centralidade de autovetor, sendo, ainda, mais natural e intuitiva que esta última.

Foi apresentado um exemplo numérico para ilustrar a aplicação da nova metodologia, bem como possibilitar análises comparativas entre as três medidas de centralidade. Em tal exemplo, a centralidade em camadas comportou-se perfeitamente como intermediária entre as duas medidas tradicionais, ordenando a primeira metade das unidades como a centralidade de autovetor e a segunda metade, como a centralidade de informação. Observou-se apenas que em relação à primeira metade da *ranking*, a nova centralidade empatou unidades ordenadas em posições subsequentes pela centralidade de autovetor.

Observou-se, ainda, que a centralidade em camadas considera o nível de influência dos vizinhos apenas para elementos relativamente bem conectados e, ainda, de maneira gradual. Em outras palavras, as unidades periféricas são ordenadas de modo semelhante à centralidade de grau porque o nível de influência dos seus vizinhos, por serem pouco numerosos, não é relevante para a centralidade em camadas. À medida em que são avaliados elementos com mais conexões com vizinhos, o nível de influência de tais vizinhos torna-se mais relevante para a nova centralidade. Por isso, a extremidade superior do *ranking* em camadas tende a apresentar semelhanças com o *ranking* de autovetor.

Por outro lado, vale ressaltar que as centralidades em camadas e de autovetor são medidas diferentes, as quais avaliam distintamente o nível de influência dos vizinhos. Assim, não há garantia de que as posições iniciais das ordenações referentes às centralidades em camadas e de autovetor serão semelhantes.

As medidas de centralidade foram, ainda, aplicadas em um estudo de caso para analisar a importância de portos no sistema portuário brasileiro, com base na movimentação de carga não conteneurizada no período de janeiro a maio de 2013. Verificou-se que a centralidade em camadas comportou-se como intermediária entre as centralidades tradicionais na maioria dos casos, porém de maneira extrema para unidades em certas condições identificadas.

De maneira geral, o resultado das três medidas de centralidade no estudo de caso foi semelhante, sendo possível verificar que a centralidade em camadas considera a centralidade dos vizinhos para ordenar as unidades, utilizando, porém, um método mais natural e intuitivo do que o autovetor. Além disso, a nova medida discriminou melhor os portos no estudo de caso, podendo se tornar, assim, uma boa ferramenta de apoio à decisão.

Trabalhos futuros podem desenvolver o estudo das propriedades da centralidade em camadas, de maneira geral. Pode-se, ainda, aplicar a nova centralidade a outros casos concretos, para analisar a sua capacidade de prover informações relevantes a respeito dos elementos das redes estudadas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ALESP (2011). Santos abriga o maior porto da América Latina. 28 de Julho, <http://www.al.sp.gov.br/noticia/?id=260023>, acesso em 08/03/2015.
- [2] Barr, R. S.; Durchholz, M. L.; Seiford, L. M. (2000). *Peeling the DEA onion: Layering and rankordering DMUs using tiered DEA*. Southern Methodist University.
- [3] Bergiante, N. C. R.; Soares de Mello, J.C.C.B.; Nunes, M.V.R.; Paschoalino F.F. (2011). Aplicação de uma proposta de medida de centralidade para avaliação de malha aérea de uma empresa do setor de transporte aéreo brasileiro. *Journal of Transport Literature*, 5(4), 119-135.
- [4] Bonacich, P. (1972). Factoring and Weighting Approaches to Status Scores and Clique Identification. *Journal of Mathematical Sociology*, 2, 113-120.
- [5] Bonacich, P. (1987). Power and Centrality: A Family of Measures. *The American Journal of Sociology*, 92(5), 1170-1182.
- [6] Bonacich, P.; Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23, 191-201.
- [7] Bondy, A.; Murty, U.S.R. (2008). *Graph Theory*. Springer, 2008, XII, 651 p.
- [8] Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55-71.
- [9] Borgatti, S. P.; Everett, M.G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, 28(4), 466-484.
- [10] Borgatti, S. P.; Everett, M.G.; Freeman, L.C. (2002). *Ucinet 6 for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- [11] Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2), 136-145.
- [12] Choi, J. H.; Barnett, G.A.; Chon, B. (2006). Comparing world city networks: a network analysis of Internet backbone and air transport intercity linkages. *Global Networks*, 6(1), 81-99.
- [13] Container Management (2007) World Top Container Ports. July/August, <http://www.iss-shipping.com/Attachments/Top%20120%20Container%20ports.pdf>, acesso em 08/03/2015.
- [14] Containerisation International (2011) Top 100 Container Ports 2011. http://europe.nxtbook.com/nxteu/informa/ci_top100ports2011/, acesso em 08/03/2015.
- [15] Containerisation International (2012) Top 100 Container Ports 2012. http://europe.nxtbook.com/nxteu/informa/ci_top100ports2012/, acesso em 08/03/2015.
- [16] Containerisation International (2013) Top 100 Container Ports 2013. http://europe.nxtbook.com/nxteu/informa/ci_top100ports2013/, acesso em 08/03/2015.
- [17] Containerisation International (2014) One Hundred Ports – the World’s Biggest Container Terminals. http://europe.nxtbook.com/nxteu/informa/ci_top100ports2014/, acesso em 08/03/2015.

- [18] Del-Vecchio, R. R.; Galvão, D. J. C.; Silva, L.; de Lima, R. F. V. L. (2009). Medidas de Centralidade da Teoria dos Grafos aplicada a Fundos de Ações no Brasil. In: *XLI Simpósio Brasileiro de Pesquisa Operacional*, Porto Seguro, Brasil.
- [19] Derudder, B.; Taylor, P. J. (2005). The cliquishness of world cities. *Global Networks*, 5(1), 71-91.
- [20] Derudder, B.; Taylor, P. J.; Witlox, F.; Catalano, G. (2003). Hierarchical tendencies and regional patterns in the world city network: a global urban analysis of 234 cities. *Regional Studies*, 37(9), 875-886.
- [21] Ercsey-Ravasz, M.; Lichtenwalter, R.; Chawla, N. V.; Toroczkai, Z. (2012). Range-limited Centrality Measures in Complex Networks. *Physical Review E*, 85(6), 066103.
- [22] Ercsey-Ravasz, M.; Toroczkai, Z. (2010). Centrality Scaling in Large Networks. *Physical Review Letters*, 105(3), 038701.
- [23] Freeman, L. C. (1977). A set of measurements of centrality based on betweenness. *Sociometry*, 40(1), 35-41.
- [24] Freeman, L. C. (1979) Centrality in networks: I Conceptual clarification. *Social Networks*, 1(3), 215-239.
- [25] Freeman, L. C.; Borgatti, S. P.; White, D.R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, 13(2), 141-154.
- [26] Gomes, E.G.; Soares de Mello, J. C. C. B.; Souza, G. S.; Angulo Meza, L.; Mangabeira, J. A. C. (2009). Efficiency and sustainability assessment for a group of farmers in the Brazilian Amazon. *Annals of Operations Research*, 169(1), 167-181.
- [27] Gonçalves, J. A. M.; Portugal, L. S.; Nassi, C. D. (2009). Centrality indicators as an instrument to evaluate the integration of urban equipment in the area of influence of a rail corridor. *Transportation Research Part A*, 43(1), 13-25.
- [28] Hussain, D. M. A. (2007). Destabilization of Terrorist Networks through Argument Driven Hypothesis Model. *Journal of Software*, 2(6), 22-29.
- [29] Jiang, T.S., Polizzi, Z.T.; Yuan, C.Q. (2014). A Networks and Machine Learning Approach to Determine the Best College Coaches of the 20th-21st Centuries. *CoRR*, abs/1404.2885.
- [30] Joyce, K. E.; Laurienti, P. J.; Burdette, J. H.; Hayasaka, S. S. (2010). A New Measure of Centrality for Brain Networks. *PLoS ONE*, 5(8), e12200.
- [31] Kirkland, S. (2009), Algebraic connectivity for vertex-deleted subgraphs, and a notion of vertex centrality. *Discrete Mathematics*, 310(4), 911-921.
- [32] Mahutga, M. C.; Ma, X.; Smith, D. A.; Timberlake, M. (2010). Economic globalization and the structure of the world city system: the case of airline passenger data. *Urban Studies*, 47(9), 1925-1947.
- [33] Ma, X.; Timberlake, M. (2008). Identifying China's leading world city: a network approach. *GeoJournal*, 71(1), 19-35.
- [34] Meghanathan, N. (2014). Centrality-Based Connected Dominating Sets for Complex Network Graphs. *International Journal of Interdisciplinary Telecommunications and Networking*, 6(2), 1-24.
- [35] Neal, Z. P. (2011). Differentiating Centrality and Power in the World City Network. *Urban Studies*, 48(13), 2733-2748.
- [36] Page, L.; Brin, S.; Motwani, R; Winograd, T. (1999). *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report, Stanford InfoLab.
- [37] Previdelli, A. (2012). Os 12 melhores portos públicos brasileiros. Exame.com, 1º de Novembro, <http://exame.abril.com.br/brasil/noticias/os-12-melhores-portos-publicos-brasileiros>, acesso em 08/03/2015.
- [38] Ruhnau, B. (2000). Eigenvector-centrality – a Node-centrality?. *Social Networks*, 22(4), 357-365.

- [39] Smith, D. A.; Timberlake, M. F. (2001). World city networks and hierarchies: 1977–1997, *American Behavioral Scientist*, 44(10), 1656-1678.
- [40] Taylor, P. J.; Catalano, G.; Walker, D. R. F. (2002). Measurement of the world city network. *Urban Studies*, 39(13), 2367-2376.
- [41] Wei, W.; Pfeffer, J.; Reminga, J.; Carley, K. M. (2011). *Handling Weighted, Asymmetric, Self-Looped, and Disconnected Networks in ORA*. Technical Report, Pittsburgh: Carnegie Mellon University.