



## MEDIDAS DE CENTRALIDADE EM GRAFOS

Leandro Quintanilha de Freitas

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Produção, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Produção.

Orientadores: Nair Maria Maia de Abreu

Renata Raposo Del-Vecchio

Rio de Janeiro  
Dezembro de 2010

# MEDIDAS DE CENTRALIDADE EM GRAFOS

Leandro Quintanilha de Freitas

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE PRODUÇÃO.

Examinada por:

---

Prof. Nair Maria Maia de Abreu, D. Sc.

---

Prof. Renata Raposo Del-Vecchio, D. Sc.

---

Prof. Lilian Markenzon, Ph. D.

---

Prof. Samuel Jurkiewicz, D. Sc.

RIO DE JANEIRO, RJ - BRASIL

DEZEMBRO DE 2010

Freitas, Leandro Quintanilha

Medidas de Centralidade em Grafos/ Leandro Quintanilha  
de Freitas. - Rio de Janeiro:UFRJ/COPPE, 2010.

VIII, 103 p.:il; 29,7 cm.

Orientadores: Nair Maria Maia de Abreu

Renata Raposo Del-Vecchio.

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de  
Engenharia de Produção, 2010.

Referências Bibliográficas: p. 93-101.

1. Grafos I. Abreu, N.M.M, Del-Vecchio, R.R.. II.  
Universidade Federal do Rio de Janeiro, COPPE, Programa  
de Engenharia de Produção. III. Título.

Aos meus queridos pais Edson e Rosângela  
e minha noiva Juliana.

## Agradecimentos

À Deus, criador do universo, pelo caminho, pela sabedoria, pela vida.

À minha orientadora, professora Nair Maria Maia de Abreu, pelo apoio, amizade, atenção e dedicação demonstrada durante todo o período de orientação.

À professora Renata Del-Vecchio pela co-orientação, pela oportunidade e pelo carinho ao longo desta caminhada.

Aos professores Samuel Jurkiewicz e Lilian Markenzon participantes da banca examinadora.

Aos meus pais e familiares, pelo apoio e por toda a educação que me deram ao longo de minha vida.

À minha querida noiva, Juliana Mariá, que esteve comigo em todos os momentos.

Aos meus amigos que muito contribuíram para encarar com bom humor tantas horas de trabalho e estudo. Em particular, à Cristiano Souza e Leonardo Macchiarulo.

Aos colegas do grupo de Teoria Espectral de Grafos pelo apoio, amizade e pelos ensinamentos.

Ao PEP, pelos professores, pelos funcionários, pelos amigos, pela infra-estrutura e ambiente favoráveis ao desenvolvimento da pesquisa.

Finalmente, à CAPES, pelo incentivo financeiro.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M. Sc.)

## MEDIDAS DE CENTRALIDADE EM GRAFOS

Leandro Quintanilha de Freitas

Dezembro/2010

Orientadores: Nair Maria Maia de Abreu

Renata Raposo Del-Vecchio

Programa: Engenharia de Produção

O estudo de redes é de grande interesse na área científica, dada a capacidade de uma dada rede poder representar por meio de modelagem diversos problemas de natureza real. Em grafos como modelos para redes, as medidas de centralidade buscam medir a variação da importância dos vértices, em função de alguns invariantes do grafo. Nesta dissertação estão reunidas medidas de centralidade e os principais resultados a elas relacionados encontrados na literatura visando padronizar a notação e a nomenclatura para disponibilizá-los aos interessados nesta área atualmente tão estudada. Ao final são apresentadas aplicações em grafos que modelam redes reais. Uma análise envolvendo a comparação dos vértices quanto à centralidade e adequação das medidas em determinados tipos de redes.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M. Sc.)

## CENTRALITY MEASURES IN GRAPHS

Leandro Quintanilha de Freitas

December/2010

Advisors: Nair Maria Maia de Abreu

Renata Raposo Del-Vecchio

Department: Production Engineering

The study of networks it is of great interest in the scientific area, given the capacity of one network to be able to represent by means of modeling diverse problems of real nature. In graphs as models for networks, the measures of centrality search to measure the variation of the importance of the vertices, in function of some parameters of the graph. In this monograph are congregated measures of centrality and the main results to them related found in literature aiming to standardize the notation and the nomenclature to give its' access to the interested parties in this currently so studied area. To the end applications in graphs are presented that shape real networks. An analysis involving the comparison of vertices as for the centrality and adequacy of the measures in determined types of networks.

# Sumário

<b>Capítulo 1 – Introdução.....</b>	<b>1</b>
<b>Capítulo 2 – Definições e Notações Básicas.....</b>	<b>6</b>
2.1 Grafos.....	6
2.2 Grafos e Matrizes.....	9
<b>Capítulo 3 – Medidas de Centralidade Não-Espectrais.....</b>	<b>17</b>
3.1 Centralidade de Grau.....	18
3.2 Centralidade de Proximidade.....	26
3.2.1 Centralidade de Eficiência.....	22
3.3 Centralidade de Intermediação.....	26
3.3.1 Centralidade de Intermediação de Percursos Aleatórios.....	31
3.4 Centralidade de Grafos.....	36
<b>Capítulo 4 – Medidas de Centralidade Espectrais.....</b>	<b>40</b>
4.1 Centralidade de Autovetor.....	40
4.2 Coordenadas de um autovetor e suas relações com o grau.....	48
4.3 Centralidade de autovetor em grafos desconexos.....	57
4.4 Centralidade de um vértice via conectividade algébrica.....	60
<b>Capítulo 5 – Centralidade em Grafos Valorados.....</b>	<b>64</b>
5.1 Medidas de centralidade em grafos .....	64
5.2 Centralidade de autovetor em grafos valorados.....	65
5.3 Centralidade de Intermediação de Fluxo.....	67
<b>Capítulo 6 – Aplicações e Resultados.....</b>	<b>72</b>
6.1 Vértices automorficamente equivalentes e vértices universais.....	72
6.2 Aplicações.....	78
6.2.1 Rede de Co-autorias.....	78
6.2.2 Rede de AIDS.....	84
6.2.3 Outras Aplicações.....	88
<b>Capítulo 7 – Considerações Finais.....</b>	<b>91</b>
<b>Referências Bibliográficas.....</b>	<b>93</b>
<b>Índice Remissivo.....</b>	<b>102</b>



# Capítulo 1 - Introdução

O estudo de redes é de grande interesse na área científica, dada a capacidade de uma dada rede poder representar por meio de modelagem diversos problemas de natureza real. Há diversos tipos de redes, dentre as quais podemos destacar, as redes sociais, compostas de indivíduos que estão ligados por um ou mais tipos específicos de interdependência, como amizade, parentesco, relações comerciais, relações sexuais, poder, conhecimento ou prestígio.

Uma rede pode ser representada matematicamente por um objeto chamado *grafo*, utilizado para descrever sua estrutura topológica. Um grafo é constituído por um conjunto de pontos, denominados *nós* ou *vértices*, conectados por linhas que expressam a relação entre eles, denominadas *arestas*.

Em grafos como modelos para redes sociais, os vértices representam os atores e as arestas a relação entre eles. Atores importantes são tais que estão frequentemente envolvidos na relação com outros atores. Estes envolvimento os tornam mais visíveis, possuindo a maioria do acesso ou controle, sendo considerados mais centrais na rede [37]. É neste sentido que as medidas de centralidade tentam descrever as propriedades da localização de um ator na rede. Estas medidas levam em consideração as diferentes maneiras em que um ator interage e se comunica com o restante da rede, sendo mais importantes, ou centrais, aqueles vértices localizados em posições mais estratégicas na rede, dada em função de alguns invariantes do grafo.

A noção de centralidade em redes sociais foi introduzida por BAVELAS [4], em 1948, quando ele afirma que, num grupo de pessoas, um particular indivíduo que se encontra estrategicamente localizado num caminho mais curto de comunicação entre pares de indivíduos, está numa posição mais central da rede. Ele será responsável por

transmitir, modificar, ou reter a informação entre membros do grupo e será tão mais influente quanto mais central estiver posicionado na rede. A centralidade tem sido usada para investigar a influência dos vértices em redes inter-organizacionais, GRASSI, STEFANI e TORRIERO [26], redes de alimento, KIRKLAND [30], mercado financeiro, DEL-VECCHIO et al. [15], fluxo de informações, BORGATTI [11], combate a redes terroristas, HUSSAIN [51], busca em redes de internet, WILF [38], propagação de doenças, BORGATTI [10], entre outros.

Ao longo de décadas, pesquisadores em redes têm introduzido um grande número de medidas de centralidade, com o objetivo de medir a variação da importância dos vértices de acordo com critérios pré-estabelecidos. Entretanto, diversas medidas foram criadas com base apenas em idéias intuitivas, sem buscar a formalização necessária para o desenvolvimento da teoria e outras foram introduzidas de maneira tão complexa que é difícil até mesmo descobrir o que está sendo medido, trazendo poucos reflexos no aspecto qualitativo.

FREEMAN [21], em 1979, abordou o conceito de centralidade revisando um grande número de medidas até então publicadas e reduziu-as a três definições clássicas, a *centralidade de grau (Degree Centrality)*, *centralidade de proximidade (Closeness Centrality)* e *centralidade de intermediação (Betweenness Centrality)*. O grau é uma medida da influência direta que um vértice tem em relação a seus contatos, a proximidade está relacionada com o tempo que uma informação leva para ser compartilhada por todos os vértices na rede, e a intermediação de um vértice pode ser considerada como o controle da comunicação entre todos os demais pares de vértices da rede. As medidas de centralidade de proximidade e centralidade de intermediação são baseadas na suposição que a informação (ou qualquer conteúdo da ligação) é

transmitida somente ao longo dos possíveis caminhos mais curtos, denominados *geodésicas*.

BONACICH [6], em 1987, propôs uma medida denominada *centralidade de autovetor (Eigenvector Centrality)*, baseada em autovalores e autovetores de matrizes simétricas, cujo objetivo é medir a importância de um vértice em função da importância de seus vizinhos. Isto quer dizer que, mesmo se um vértice está conectado somente a alguns outros vértices da rede (tendo assim uma baixa centralidade de grau), estes vizinhos podem ser importantes e, conseqüentemente, o vértice será também importante, obtendo uma centralidade de autovetor elevada.

BORGATTI, FREEMAN e WHITE [24], em 1991, levantaram uma questão importante baseada no fato de que uma informação (ou qualquer conteúdo da ligação) que viaja através da rede, não necessariamente deve seguir caminhos mais curtos, e propuseram assim uma medida de centralidade de intermediação com base no fluxo máximo em redes, denominada *centralidade de intermediação de fluxo (Flow Betweenness Centrality)*, aplicável a uma variedade mais ampla e realística de um conjunto de dados da rede, que corresponde a grafos valorados.

HAGE e HARARY [29], em 1995, propuseram uma medida de centralidade baseada na excentricidade de um vértice, chamada *centralidade de eficiência (Efficiency Centrality)*, com o objetivo de determinar uma localização que minimize as distâncias a todas as demais localizações da rede. Os problemas de localização são clássicos em otimização combinatória com muitas aplicações na indústria e na economia. Trata-se de um problema bem conhecido em Pesquisa Operacional e que possui inúmeras aplicações também no estudo de redes sociais.

NEWMAN [33], em 2005, introduziu uma medida de centralidade de intermediação baseada em percursos aleatórios, chamada *centralidade de intermediação de percursos*

*aleatórios (Random Walks Betweenness Centrality)*, que é apropriada a uma rede em que um determinado fluxo percorre vértice a vértice de maneira aleatória até que encontre seu destino. A idéia de se usar esta medida partiu da centralidade baseada no fluxo de corrente elétrica (*Current Flow Centrality*) e mostra-se, em [33], que ambas as medidas são similares.

Recentemente, GRASSI, STEFANI e TORRIERO, [26], [27] e [28], em 2006, 2007 e 2009, vêm pesquisando a centralidade de autovetor, obtendo resultados e parâmetros de comparação entre suas coordenadas e relacionando à medida de centralidade de grau. Já KIRKLAND, [30], em 2009, introduziu uma medida de centralidade baseada em conectividade algébrica, que mede a importância de um vértice no que diz respeito à vulnerabilidade da rede, isto é, a propensão que a rede tem para ser destruída.

O conceito de centralidade pode ser então formalizado para que se possa aplicar a qualquer rede. Logo, é muito importante que saibamos medir a centralidade de cada componente de uma dada rede, o que equivale a medir a centralidade de cada vértice no grafo que a modela. A partir do resultado de cada medida, é possível ordenar os vértices da rede em função de sua importância relativa, que está associada ao significado que cada medida considera ser um vértice, mais central que outros.

O presente trabalho tem como objetivo reunir num mesmo documento as principais medidas de centralidade existentes na literatura, descrevendo a noção de centralidade em termos da teoria dos grafos. Apresentaremos ainda, os cálculos, análises e interpretações de cada medida de centralidade buscando parâmetros de comparações em determinadas classes específicas de grafos e investigando o comportamento das medidas de centralidade aplicadas em diferentes situações. Desta forma, a dissertação se desenvolve com a apresentação, no capítulo 2, dos conceitos básicos de Teoria de

Grafos relativos à centralidade em redes; com a descrição no capítulo 3, das medidas de centralidade não espectrais em grafos, juntamente com a apresentação dos seus principais resultados aplicados a exemplos. No capítulo 4, são tratadas as medidas de centralidade espectrais em grafos; no capítulo 5, as medidas de centralidade convencionais são aplicadas e analisadas em grafos valorados e no capítulo 6, expomos alguns resultados relacionando classes específicas de grafos e algumas aplicações, dentre elas, a rede de co-autorias do grupo TEG-RIO. Finalmente, no capítulo 7, algumas considerações são apresentadas.

## Capítulo 2 – Definições e Notações Básicas

Este capítulo se destina a apresentar conceitos básicos de Teoria dos Grafos necessários para que o texto desta dissertação seja auto-contido. Os conceitos aqui reescritos podem ser encontrados em livros clássicos de Teoria dos Grafos. Ao leitor interessado recomendamos as referências [5], [8], e [16], de onde nos baseamos para desenvolver a 1ª seção. Os conceitos da 2ª seção podem ser encontrados em [8], [16], e [18].

### 2.1 Grafos

Um *grafo*  $G = (V(G), E(G))$  é uma estrutura, consistindo de um conjunto finito e não vazio de vértices  $V = V(G)$  e um conjunto  $E = E(G)$  formado por pares de elementos de  $V$ , denominado arestas de  $G$ . O conjunto de arestas  $E$  induz uma relação binária em  $V$  que é chamada *relação de adjacência* dos vértices de  $G$ . Sejam  $v_i$  e  $v_j \in V$ , dizemos que  $v_i$  e  $v_j$  são *vértices adjacentes* ou *vizinhos* se a aresta  $e_{ij} = \{v_i, v_j\} \in E$ . Neste caso, dizemos que  $e_{ij}$  *incide* tanto em  $v_i$  quanto em  $v_j$ , sendo  $v_i$  e  $v_j$  os *extremos* de  $e_{ij}$ . O conjunto formado por todos os vértices adjacentes a  $v_i$  é denominado vizinhança de  $v_i$  (ou vizinhança de primeira ordem de  $v_i$ ) e denotado por  $\Gamma(v_i)$ . A vizinhança de segunda ordem de  $v_i$  é  $\bigcup_{v_j \in \Gamma(v_i)} \Gamma(v_j)$ . Similarmente, vizinhanças de ordens superiores podem ser definidas. O número de vértices  $n$  e o número de arestas  $m$  de  $G$  são respectivamente iguais a  $|V(G)| = n$  e  $|E(G)| = m$ . Uma aresta incidente a apenas um vértice é denominada *laço* e duas ou mais arestas incidentes ao mesmo par

de vértices são chamadas *arestas múltiplas*. Grafos sem laços ou arestas múltiplas são chamados de *grafos simples*. Um *grafo orientado* é definido quando um par de vértices  $v_i$  e  $v_j \in V$  conectados por uma aresta tem sentido definido de  $v_i$  para  $v_j$  ou vice-versa. Neste caso, denominamos por *arco* a aresta orientada, e denotamos por  $e_{ij} = (v_i, v_j)$ . Neste texto, a maioria dos problemas envolve grafos simples e não orientados e assim deverão ser entendidos a menos que especificado em contrário.

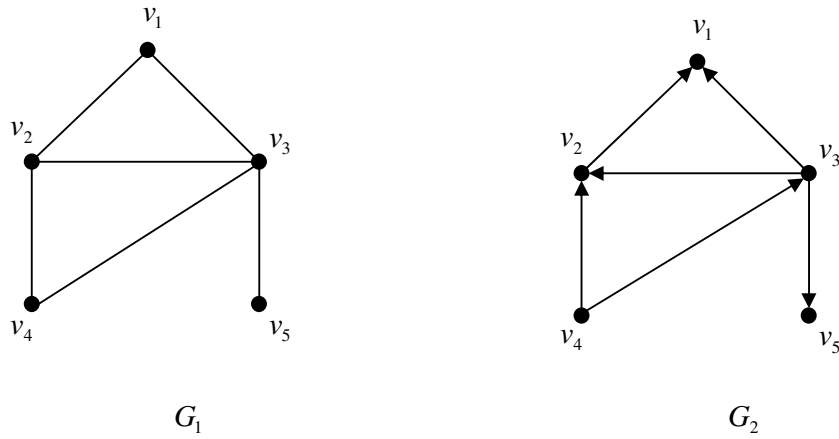


Figura 2.1:  $G_1$  é um grafo simples e  $G_2$  é um grafo orientado.

O *grau* de um vértice  $v_k$ , denotado por  $d(v_k)$  ou  $d_k$ , é o número de arestas incidentes em  $v_k$ . O *grau mínimo* de  $G$ , denotado por  $\delta(G)$ , é o menor grau dentre todos os outros e o *grau máximo*, denotado por  $\Delta(G)$ , é o maior grau dentre todos os graus dos vértices de  $G$ . O número  $\bar{d} = \frac{1}{|V|} \sum_{v_i \in V} d_i$  é chamado *grau médio* de  $G$ .

Dizemos que um vértice  $v_k \in V$  é um *vértice universal* de  $G$  se  $d_k = n - 1$ . Um grafo  $G$  é *k-regular*, ou simplesmente regular, se cada vértice tem grau  $k$ . Um grafo é chamado *completo* se quaisquer dois pares distintos de vértices são adjacentes.  $K_n$  denota o grafo completo de  $n$  vértices. O grafo  $G_1 = (V_1, E_1)$  é um *subgrafo* de  $G$  se

$V_1 \subseteq V$  e  $E_1 \subseteq E$ . Além disso, se  $G_1$  contém exatamente as arestas de  $G$  incidentes nos vértices de  $V_1$ ,  $G_1$  é um *subgrafo induzido* de  $G$ , e se contém todos os vértices de  $G$ ,  $G_1$  é um *subgrafo gerador* de  $G$ . Uma *cadeia*, ou *percurso*, é uma sequência finita e alternada de vértices e arestas. Uma *trilha* é uma cadeia sem aresta repetida. Um *caminho*, denotado por  $P_n$ , é uma cadeia onde todos os  $n$  vértices são distintos. Um grafo  $G$  é *conexo* se existe um caminho ligando cada par de vértices. Caso contrário,  $G$  é *desconexo*. Um grafo  $G$  desconexo é formado por pelo menos dois subgrafos induzidos conexos denominados *componentes conexas* de  $G$ . O *comprimento* de um caminho é o número de arestas que neles ocorrem. Um menor caminho entre dois vértices  $v_i$  e  $v_j$  é conhecido como *geodésica*. Seja  $G$  um grafo conexo, a *distância* entre dois vértices  $v_i$  e  $v_j$  de  $G$ , denotada por  $dist(v_i, v_j)$ , é o comprimento da geodésica entre  $v_i$  e  $v_j$ . O número de geodésicas entre  $v_i$  e  $v_j$  é denotado por  $g_{ij}$ . Seja  $v_k$  um vértice pertencente a uma geodésica de  $v_i$  a  $v_j$ ,  $g_{ij}(v_k)$  denota o número de geodésicas entre  $v_i$  e  $v_j$  que passam por  $v_k$ . Se  $v_i$  e  $v_j$  estão em componentes conexas distintas, definimos  $dist(v_i, v_j) = \infty$ . Se os vértices finais de uma trilha são os mesmos (uma trilha fechada) então nós referimos como um *circuito*. Ao circuito sem vértice repetido denominamos um *ciclo*. Em particular,  $C_n$  denota o ciclo de  $n$  vértices. Note que um ciclo é também um circuito, no entanto, um circuito não é necessariamente um ciclo. Dado um vértice  $v_i \in V$ , *afastamento*, ou *excentricidade*  $e(v_i)$  é a maior distância entre  $v_i$  e algum outro vértice em  $G$ . O *raio*  $r(G)$  de um grafo é a menor das excentricidades. *Diâmetro*  $diam(G)$  de um grafo é a maior das excentricidades existentes no grafo (logo, a maior distância).



Um *conjunto independente* em um grafo é um conjunto de vértices não adjacentes entre si. Um grafo é *bipartido* se seus vértices podem ser separados em dois conjuntos não vazios, disjuntos e independentes  $V_1$  e  $V_2$  de tal forma que toda aresta  $e_{ij} \in E$  tem uma extremidade em  $V_1$  e a outra em  $V_2$ . Quando  $|V_1| = r$ ,  $|V_2| = s$  e existe uma aresta incidindo em quaisquer pares de vértices  $u_i$  e  $v_i$  com  $u_i \in V_1$  e  $v_i \in V_2$ , dizemos que o grafo é *bipartido completo*, e denotamos por  $K_{r,s}$ . Um grafo  $G$  é chamado *árvore* quando  $G$  é um grafo conexo e sem ciclos. Toda árvore é um grafo bipartido. Um *grafo estrela*, denotado por  $S_n$ , é um grafo bipartido de  $n$  vértices que possui um conjunto independente com um único vértice e o outro, com  $n-1$  vértices. Note que a estrela e o caminho são exemplos de árvores.

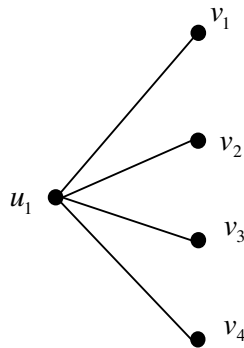


Figura 2.2: Grafo estrela  $S_5$ .

## 2.2 Grafos e Matrizes

Nesta seção, apresentaremos alguns conceitos de Teoria Espectral de Grafos. Vamos assumir, como pré-requisito, conhecimentos básicos de álgebra linear.

Seja  $G$  um grafo com  $n$  vértices. A *matriz de adjacência*  $A(G)$  de  $G$  é uma matriz de ordem  $n$  cujas entradas são:

$$a_{ij} = \begin{cases} 1, & \text{se } \{v_i, v_j\} \in E; \\ 0, & \text{caso contrário.} \end{cases} \quad (2.1)$$

Para um grafo não-orientado,  $A(G)$  é uma matriz binária e simétrica, isto é,  $a_{ij} = a_{ji}$ , para todo  $i, j$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq n$ , cujo traço é zero. Um exemplo de um grafo simples não orientado de 5 vértices com 6 arestas e sua respectiva matriz de adjacência  $A(G)$  é dado abaixo:

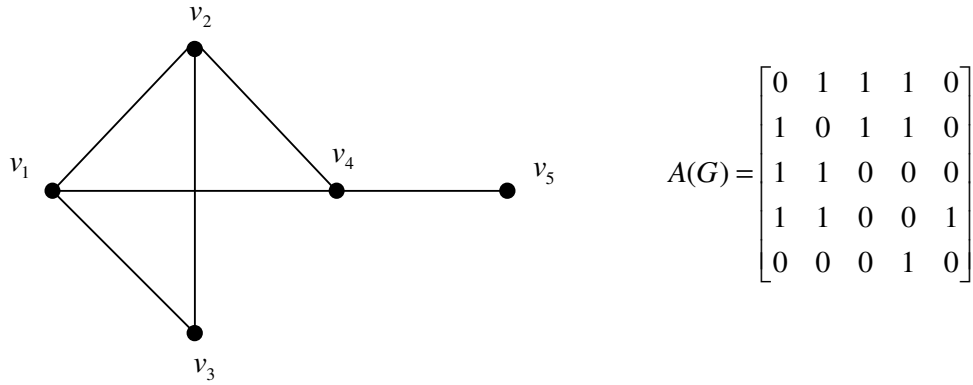


Figura 2.3: Grafo  $G$  e sua respectiva matriz de adjacência.

Note que a soma das entradas de cada linha da matriz de adjacência de um grafo é igual ao grau do vértice correspondente. A  $k$ -ésima potência da matriz de adjacência é relacionada ao número de cadeias de comprimento  $k$  no grafo. Em particular,  $(A(G)^k)_{ij}$  fornece o número de cadeias de comprimento  $k$  entre  $v_i$  e  $v_j$ .

O polinômio característico de  $A(G)$ , ou seja,  $\det(\lambda I - A(G))$ , é denominado *polinômio característico do grafo  $G$*  e denotado por  $p_G$ ;  $\lambda$  é dito ser um *autovalor do grafo  $G$*  quando é uma raiz de  $p_G$ . Se  $A(G)$  possui  $t$  autovalores distintos, o *espectro* de  $G$ , denotado por  $\text{spect}G$ , é definido como a matriz  $2 \times t$ , onde a primeira linha é constituída pelos autovalores distintos de  $A(G)$ , dispostos em ordem não-crescente e a segunda, pelas respectivas *multiplicidades algébricas*. Assim, se  $\lambda_1 > \dots > \lambda_t$  são os

autovalores distintos de  $A(G)$  e, para  $1 \leq i \leq t$ ,  $ma(\lambda_i)$  são suas multiplicidades algébricas, então o espectro de  $G$  é:

$$spect G = \begin{bmatrix} \lambda_1 & \cdots & \lambda_t \\ ma(\lambda_1) & \cdots & ma(\lambda_t) \end{bmatrix}. \quad (2.2)$$

O maior autovalor de  $G$  é denominado *índice de  $G$*  e denotado por  $ind(G)$ .

A Proposição 2.1, cuja prova está em [1], relaciona o índice de um grafo ao grau médio e ao grau máximo de  $G$ , invariantes estes já definidos na seção 2.1.

**Proposição 2.1.** Se  $G$  é um grafo cujo maior autovalor é  $\lambda_1$  então

$$\bar{d} \leq \lambda_1 \leq \Delta.$$

**Exemplo:**

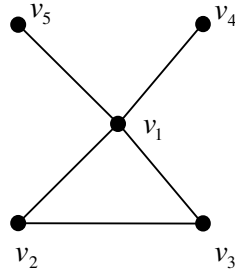


Figura 2.4: Grafo  $G$

Considere o grafo  $G$  da Figura 2.4. Seu polinômio característico é  $p_G(\lambda) = \lambda^5 - 5\lambda^3 - 2\lambda^2 + 2\lambda$  e seu espectro é dado por:

$$spect(G) = \begin{bmatrix} 2,342 & 0,470 & 0 & -1 & -1,813 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Portanto o índice do grafo  $G$  é  $ind(G) = \lambda_1 = 2,342$ . Além disso, os graus médio e máximo são respectivamente  $\bar{d} = 2$  e  $\Delta = 4$ . Como se vê na Proposição 2.1, temos  $\bar{d} \leq \lambda_1 \leq \Delta$ .

A *multiplicidade geométrica* do autovalor  $\lambda$ , denotado por  $mg(\lambda)$ , é a dimensão do autoespaço  $X_\lambda$  associado ao autovalor  $\lambda$ , isto é

$$mg(\lambda) = \dim X_\lambda.$$

O *raio espectral* de um grafo  $G$ , denotado por  $\rho(G)$ , é o número real não negativo  $\rho(G) = \max_i |\lambda_i|$ , onde  $\lambda_1, \lambda_2, \dots, \lambda_n$  são os autovalores de  $G$ . O raio espectral de  $G$  é o raio do menor círculo centrado na origem que contém o espectro de  $G$ . Esse círculo é chamado *círculo espectral de  $G$* .

Para um grafo  $G$  conexo, a matriz de adjacência  $A = A(G)$  é irredutível, isto é, não existe uma matriz de permutação  $P$  tal que:

$$P^t A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

com  $B$  e  $D$  matrizes quadradas. Sob tais condições, podemos enunciar o Teorema de Perron-Frobenius [18]:

**Teorema 2.1. (Teorema de Perron-Frobenius)** Seja  $A$  uma matriz  $n \times n$ , simétrica, não negativa, irredutível e com autovalores  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . Então

- (1)  $\lambda_1 > 0$  e, associado a esse autovalor, existe um autovetor cujas coordenadas são, todas elas, positivas;
- (2)  $\lambda_1$  é estritamente maior que  $\lambda_2$ , ou seja,  $\lambda_1 > \lambda_2$ ;
- (3)  $|\lambda_i| \leq \lambda_1$  para todo  $i \in \{1, 2, \dots, n\}$ .

**Prova:**

(1) Primeiro observe que como a matriz  $A$  é não negativa,  $tr(A) = \sum_i \lambda_i \geq 0$ ; e como, por hipótese,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ , segue-se que  $\lambda_1 \geq 0$ . O teorema de Raleigh-Ritz [1] garante que

$$\lambda_1 = \max_{\|y\|=1} y^t A y.$$

Seja  $y = (y_1, y_2, \dots, y_n)^t$  um autovetor de norma 1 associado a  $\lambda_1$ . Então temos que:

$$\lambda_1 = y^t A y = \sum_{i=1}^n \sum_{j=1}^n a_{ij} y_i y_j \leq \sum_{i=1}^n \sum_{j=1}^n a_{ij} |y_i| |y_j| \leq \lambda_1 = \max_{\|y\|=1} y^t A y.$$

Ou seja,  $x = (|y_1|, |y_2|, \dots, |y_n|)^t$  é também um autovetor de norma 1 associado a  $\lambda_1$ , sendo que as coordenadas são não-negativas. Suponha que existam coordenadas nulas em  $x$ . Seja  $\sigma$  uma permutação tal que  $|y_{\sigma(i)}| > 0$  para todo  $i \leq m$  e  $|y_{\sigma(i)}| = 0$  para todo  $i > m$ . Escrevendo a matriz  $A$  segundo a permutação  $\sigma$  temos:

$$A_\sigma = \begin{pmatrix} B_{m \times m} & C_{m \times (n-m)} \\ D_{(n-m) \times m} & E_{(n-m) \times (n-m)} \end{pmatrix} \text{ e } x_\sigma = \begin{pmatrix} y'_{m \times 1} \\ 0_{(n-m) \times 1} \end{pmatrix}$$

As matrizes  $A_\sigma$  e  $A$  têm os mesmos autovalores já que são matrizes semelhantes, sendo assim  $A_\sigma x_\sigma = \lambda_1 x_\sigma$ . Temos então que:

$$\begin{pmatrix} B & C \\ D & E \end{pmatrix} \begin{pmatrix} y' \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 y' \\ 0 \end{pmatrix}, \quad (2.3)$$

e devemos ter  $D = 0$  e, portanto,  $A$  é uma matriz redutível, o que contradiz a hipótese. Logo, o autovetor  $x$  possui todas as coordenadas positivas. Observe ainda que supondo  $\lambda_1 = 0$  em (2.3) temos, da mesma maneira, que  $D = 0$  o que contraria, novamente, o fato de  $A$  ser irredutível.

(2) Suponhamos por absurdo que  $mg(\lambda_1) \geq 2$ . Note que  $ma(\lambda_1) = mg(\lambda_1)$ , pois a matriz  $A$  é simétrica. Escolhemos  $y$  e  $x$  dois autovetores ortonormais associados a  $\lambda_1$  e definimos os autovetores associados a  $\lambda_1$  como sendo:

$$y' = \begin{pmatrix} y_1 + |y_1| \\ y_2 + |y_2| \\ \vdots \\ y_n + |y_n| \end{pmatrix} \text{ e } x' = \begin{pmatrix} x_1 + |x_1| \\ x_2 + |x_2| \\ \vdots \\ x_n + |x_n| \end{pmatrix}$$

onde  $y_i$  e  $x_i$  são as coordenadas de  $y$  e  $x$ , respectivamente.

É fácil ver que  $y_i + |y_i| \geq 0$  para todo  $i \in \{1, 2, \dots, n\}$ , portanto  $y_i + |y_i| > 0$ , caso contrário, usando a mesma estratégia usada na prova do item (1), concluiríamos que a matriz  $A$  não seria irredutível. Da mesma forma temos  $x_i + |x_i| > 0$  para todo  $i \in \{1, 2, \dots, n\}$ . Logo temos  $y_i, x_i > 0$ , para todo  $i \in \{1, 2, \dots, n\}$  e como, por hipótese, os vetores  $y$  e  $x$  são ortogonais, isto é,  $y \cdot x = 0$ , chegamos ao absurdo. Assim,  $mg(\lambda_1) = 1$  e  $\lambda_1 > \lambda_2$ .

(3) Basta provar que se  $x$  é um autovetor associado a  $\lambda_1$ , e cuja norma vale 1, então

$$|\lambda_1| = |x'Ax| = \left| \sum_{k=1}^n \sum_{l=1}^n a_{kl} x_k x_l \right| \leq \sum_{k=1}^n \sum_{l=1}^n a_{kl} |x_k| |x_l| \leq \lambda_1.$$

□

O Teorema de Perron-Frobenius garante que matrizes de adjacência associadas a grafos conexos possuem o maior autovalor positivo e, associado a este, um único autovetor unitário positivo. A prova aqui apresentada é a encontrada em [18]. O resultado vale para o caso em que apenas se exige que  $A$  seja não-negativa e irredutível. Ver [63].

Seja  $G$  um grafo de ordem  $n$ ,  $A(G)$  sua matriz de adjacência e  $D(G)$  a matriz diagonal de ordem  $n$ , cujos elementos da diagonal principal são os graus dos vértices

de  $G$ . A *matriz Laplaciana* de  $G$ , denotada por  $L(G)$ , é a matriz simétrica definida por:

$$L(G) = D(G) - A(G). \quad (2.4)$$

Analogamente ao que foi definido para matrizes de adjacência, podemos falar no polinômio característico e espectro da matriz Laplaciana.

Seja  $G$  um grafo e  $L(G)$  sua matriz Laplaciana. Então,

1. O polinômio característico associado à  $L(G)$ , denotado por  $p_{L(G)}(\mu)$ , é denominado *polinômio característico laplaciano* de  $G$ ;
2.  $\mu$  é um autovalor de  $L(G)$  se  $\mu$  é raiz de  $p_{L(G)}$ ;
3. O espectro de  $L(G)$ , denotado por  $\zeta(G)$ , é definido como sendo uma matriz cujos elementos são todos os autovalores de  $L(G)$ , ordenados em ordem não-crescente.

Assim, se  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$  são os autovalores de  $L(G)$ , então:

$$\zeta(G) = (\mu_1, \mu_2, \dots, \mu_n). \quad (2.5)$$

A matriz laplaciana  $L(G)$  é semi-definida positiva e, portanto, todos os seus autovalores são maiores ou iguais a zero. Além disso, como as colunas de  $L(G)$  formam um conjunto linearmente dependente, tem-se que pelo menos um autovalor de  $L(G)$  é nulo [1].

Sejam  $G$  um grafo de ordem  $n$  e  $\zeta(G) = (\mu_1, \dots, \mu_{n-1}, 0)$  o espectro da matriz Laplaciana de  $G$ . A *conectividade algébrica* de  $G$  é definida por  $\alpha(G) = \mu_{n-1}$ .

A Proposição 2.2, cuja prova pode ser encontrada em [1], relaciona a conexidade de um grafo  $G$  com a sua conectividade algébrica  $\alpha(G)$ .

**Proposição 2.2.** Seja  $\alpha(G)$  a conectividade algébrica de um grafo  $G$ . Então  $G$  é conexo se e somente se  $\alpha(G) \neq 0$ .

No próximo capítulo, serão apresentados os principais conceitos e definições, incluindo exemplos, referentes às medidas de centralidade não espectrais em grafos e no capítulo 4, apresentaremos então as medidas de centralidade decorrentes de invariantes espectrais.



## Capítulo 3 – Medidas de Centralidade Não-Espectrais

As medidas de centralidade surgiram da análise em redes sociais, que são um tipo específico de rede que se referem a conjunto de pessoas conectadas por relacionamentos sociais. Toda rede, social ou não, pode ser representada por um grafo. Cada posição na rede corresponde a um vértice no grafo e cada possível relação entre as posições corresponde a uma aresta que conecta um par de vértices.

Intuitivamente, numa rede, os vértices mais centrais são aqueles que a partir dos quais podemos atingir qualquer outro com mais facilidade ou rapidez.

Neste capítulo nós introduziremos diferentes medidas de centralidade que avaliam a importância dos nós em uma rede de acordo com sua posição estrutural, dadas em função de alguns invariantes do grafo. A *centralidade de grau*, assim chamada por ser uma função do grau de um vértice, conta o número de arestas incidentes a um vértice do grafo. A *centralidade de proximidade* está relacionada com a distância total de um vértice a todos os demais vértices do grafo. A *centralidade de eficiência* minimiza as distâncias de um vértice para alcançar qualquer outro vértice no grafo. A *centralidade de intermediação* mede quantas geodésicas entre todos os pares de vértices do grafo passam através de um determinado vértice. A *centralidade de intermediação de percursos aleatórios* considera todos os caminhos, geodésicos e não-geodésicos, passando através de um vértice possíveis de serem usados por um percurso aleatório. Além disso, considera as probabilidades de tais caminhos serem usados. Por fim, é apresentada uma medida de centralidade não do vértice, mas de todo o grafo  $G$ , que é dada em função das medidas de centralidade dos vértices de  $G$ , o que nos permite comparar grafos de mesma ordem e topologicamente diferentes. Tais medidas de centralidade traduzem diferentes aspectos da posição de um vértice em um grafo e,

conseqüentemente, a escolha da medida a ser aplicada com mais correção deverá depender da aplicação em particular considerada.

### 3.1 Centralidade de Grau

A concepção mais simples e intuitiva no que diz respeito à centralidade de um vértice é o número de contatos diretos que ele possui. Uma pessoa que se encontra em uma posição que permite o contato direto com muitos outros é vista pelos demais como um canal maior de informações, razão pela qual dizemos ser mais central. SHAW [58], em 1964, usou o grau do vértice como medida de centralidade e, em seguida, MACKENZIE [55] e BEAUCHAMP [57] aplicaram variantes do grau em problemas de redes sociais. Segundo FREEMAN [21], Nieminen, em 1974, foi quem chamou a medida usada por Shaw e os demais pesquisadores de *centralidade de grau*, e ainda afirma ser esta a medida mais adequada (pelo menos até aquela data, 1979). Assim chamada, a centralidade de grau nada mais é que a contagem do número de adjacências de um vértice  $v_k$ , ou seja, este parâmetro coincide com o próprio grau de  $v_k$ .

Formalmente, temos:

**Definição 3.1.** Seja  $G$  um grafo qualquer (conexo ou não) com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade de grau* de  $v_k$ , denotada por  $d_k$ , é o número de arestas incidentes a  $v_k$ . Isto é,

$$d_k = \sum_{j=1}^n a_{kj}, \quad (3.1)$$

onde  $a_{kj}$  são elementos da matriz de adjacência  $A(G)$ .

Considere, por exemplo, um vértice  $v_i$  num grafo  $G$  em que  $n=8$ . Suponha que  $d_i=6$ . Por outro lado, considere outro grafo  $H$ , com  $n=25$ , e um vértice  $v_j$  em  $H$  que tenha  $d_j=6$ . Ambos os vértices  $v_i$  e  $v_j$  tem o mesmo potencial de controle em termos absolutos com relação ao sistema de comunicação das respectivas redes modeladas por  $G$  e  $H$ . Contudo, enquanto  $v_i$  poderá dominar mais que a metade do sistema de comunicação em  $G$ ,  $v_j$  dominará apenas uma pequena porcentagem da rede  $H$ . Isto nos leva à seguinte definição:

**Definição 3.2.** Seja  $G$  um grafo qualquer (conexo ou não) com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade relativa de grau* de  $v_k$  é dada por:

$$c'_D(v_k) = \frac{d_k}{n-1}. \quad (3.2)$$

É fácil ver que o maior valor de  $d_k$  é  $n-1$ . Portanto, esta medida reflete a proporção dos vértices adjacentes à  $v_k$  em relação à ordem do grafo, ou seja, o número total de vértices do grafo  $G$ . Assim,  $0 \leq c'_D(v_k) \leq 1$ . É fácil notar que o limite superior é atingido por qualquer vértice universal e o inferior por vértices isolados.

**Exemplo:**

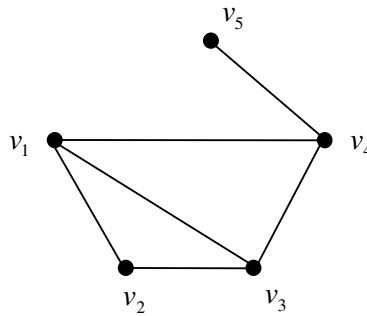


Figura 3.1: Os vértices  $v_1$ ,  $v_3$  e  $v_4$  são os mais centrais segundo a centralidade de grau.

Na Figura 3.1 temos:  $d_1 = d_3 = d_4 = 3$ ;  $d_2 = 2$  e  $d_5 = 1$ . Portanto, a medida de centralidade de grau nos indica que  $v_1$ ,  $v_3$  e  $v_4$  são os vértices mais centrais. Já usando a medida relativa da centralidade de grau podemos observar a proporção dos vértices adjacentes a cada vértice em relação à ordem do grafo, ou seja,  $c'_D(v_1) = c'_D(v_3) = c'_D(v_4) = \frac{3}{4}$ ;  $c'_D(v_2) = \frac{1}{2}$  e  $c'_D(v_5) = \frac{1}{4}$ . É claro que para um mesmo grafo, as medidas de centralidade absoluta e relativa produzirão sempre a mesma ordenação para os vértices em escala de importância, uma vez que diferem apenas em função da constante de normalização  $\frac{1}{n-1}$ . Sendo assim, a centralidade relativa será mais apropriada para se comparar a centralidade de um vértice em grafos com ordem diferentes.

### 3.2 Centralidade de Proximidade

As medidas de centralidade do vértice baseadas na proximidade foram desenvolvidas por BAVELAS [4], BEAUCHAMP [58], MOXLEY e MOXLEY [56] e SABIDUSSI [54]. A mais simples e natural destas medidas, proposta por Sabidussi, em 1966, chamada *centralidade de proximidade*, é baseada na soma das distâncias de um vértice em relação aos demais vértices do grafo.

**Definição 3.3.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade de proximidade* de  $v_k$  é dada pelo inverso da soma das distâncias de  $v_k$  a todos os demais vértices do grafo, ou seja,

$$c_C(v_k) = \frac{1}{\sum_{j=1}^n dist(v_j, v_k)}. \quad (3.3)$$

Note que  $v_k$  pode no mínimo estar a uma distância igual a 1 em relação a um dado vértice  $v_j$  num grafo conexo com  $n$  vértices e no máximo estar ligado a  $n-1$  outros vértices. Neste caso, o maior valor para a centralidade de proximidade de um vértice  $v_k$  é  $c_C(v_k) = \frac{1}{n-1}$ , pois o menor valor para  $\sum_{j=1}^n \text{dist}(v_j, v_k) = n-1$ . Também neste caso, podemos definir uma centralidade relativa de proximidade de um vértice:

**Definição 3.4.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade relativa de proximidade* de  $v_k$  é dada por:

$$c'_C(v_k) = \frac{n-1}{\sum_{j=1}^n \text{dist}(v_j, v_k)} = (n-1) \cdot c_C(v_k). \quad (3.4)$$

Também, neste caso,  $0 \leq c'_C(v_k) \leq 1$ , dado que  $\sum_{j=1}^n \text{dist}(v_j, v_k) \geq n-1$ . Novamente, o limite superior é atingido quando  $v_k$  é um vértice universal.

Tais medidas são de fato significativas apenas em grafos conexos, já que em grafos desconexos existe pelo menos um par de vértices  $v_j, v_k$  tal que  $\text{dist}(v_j, v_k) = \infty$ . Logo,

$$\sum_{j=1}^n \text{dist}(v_j, v_k) = \infty.$$

**Exemplo:**

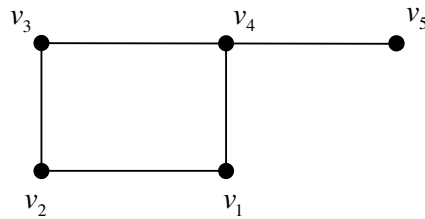


Figura 3.2: Grafo  $G$  onde  $v_4$  é o vértice mais central segundo  $c_C$ .

Para ilustrar a medida de centralidade de proximidade vamos considerar o grafo dado na Figura 3.2. As distâncias entre  $v_1$  e todos os demais vértices são dadas por:

$$dist(v_2, v_1) = 1; dist(v_3, v_1) = 2; dist(v_4, v_1) = 1; dist(v_5, v_1) = 2.$$

Pela equação (3.3), a centralidade de proximidade do vértice  $v_1$  é dada por:

$$c_c(v_1) = \frac{1}{\sum_{j=1}^5 dist(v_j, v_1)} = \frac{1}{1+2+1+2} = \frac{1}{6} = 0,166...$$

De modo análogo, podemos determinar esta medida para os demais vértices de  $G$  para obtermos:

$$c_c(v_2) = 0,142; c_c(v_3) = 0,166...; c_c(v_4) = 0,2 \text{ e } c_c(v_5) = 0,125.$$

Isto indica que o vértice  $v_4$  é o mais central de acordo com a medida  $c_c$ .

Para o grafo da Figura 3.2, a medida relativa de proximidade dos vértices dada pela equação (3.4) nos fornece:

$$c'_c(v_1) = c'_c(v_3) = 0,666...; c'_c(v_2) = 0,568; c'_c(v_4) = 0,8 \text{ e } c'_c(v_5) = 0,5.$$

**Observação:** Em um grafo com  $n$  vértices e  $m$  arestas, a medida de centralidade de proximidade pode ser calculada para todos os vértices com complexidade de  $O(m+n)$ , [14].

### 3.2.1 Centralidade de Eficiência

Em Pesquisa Operacional, alguns problemas de localização consistem em se determinar um local de modo que minimize o tempo máximo de viagem entre o mesmo e todas as demais localizações. Estes problemas possuem diversas aplicações práticas, como por exemplo, a instalação de um hospital, cujo objetivo é minimizar o tempo máximo de atendimento de uma ambulância a uma possível emergência. É neste sentido que HAGE e HARARY [29], em 1995, propuseram uma medida chamada **centralidade**

de eficiência, baseada no conceito de excentricidade de um vértice. Tal medida pode assim ser definida:

**Definição 3.5.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade de eficiência* de  $v_k$  é dada pelo inverso da excentricidade de  $v_k$ , isto é,

$$c_{eff}(v_k) = \frac{1}{e(v_k)}, \quad (3.5)$$

onde  $e(v_k) = \max\{dist(v_j, v_k) : v_j \in V\}$ .

Assim, esta medida indica que um vértice é mais eficiente quanto menor for a sua excentricidade.

**Exemplo:**

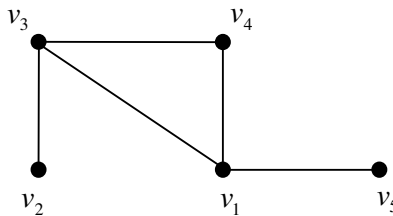


Figura 3.3: Grafo  $G$  onde  $v_1$ ,  $v_3$  e  $v_4$  são os mais centrais por  $c_{eff}$ .

Para o grafo  $G$  dado na Figura 3.3, temos que as distâncias entre  $v_1$  e todos os demais vértices são dadas por:

$$dist(v_2, v_1) = 2; \quad dist(v_3, v_1) = 1; \quad dist(v_4, v_1) = 1; \quad dist(v_5, v_1) = 1.$$

Logo, a excentricidade é dada pela distância de  $v_1$  a  $v_2$ , ou seja,  $e(v_1) = 2$ . Daí, segue que a centralidade de eficiência de  $v_1$  é:

$$c_{eff}(v_1) = \frac{1}{e(v_1)} = \frac{1}{2} = 0,5.$$

De maneira análoga, determinamos as centralidades de eficiência dos demais vértices,

$$c_{eff}(v_2) = c_{eff}(v_5) = 0,33; \quad c_{eff}(v_3) = c_{eff}(v_4) = 0,5.$$

Tal medida indica que os vértices  $v_1$ ,  $v_3$  e  $v_4$  são então os mais centrais.

Outras variações da centralidade baseadas na proximidade de um vértice merecem destaque. A primeira delas é a **centralidade centróide**, introduzida por BORGATTI em [9]. Segundo o autor, a idéia é medir a centralidade de qualquer vértice através da distância desse vértice ao centróide do grafo, que é o vértice, ou conjunto de vértices, para o qual a soma das distâncias aos demais vértices é a menor possível. A segunda, é a **centralidade radial** proposta por VALENTE e FOREMAN [31] e desenvolvida inicialmente para grafos orientados. Porém, uma versão desta medida é aplicada a grafos conexos não orientados, que é definida da seguinte maneira:

**Definição 3.6.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A

**centralidade radial** de  $v_k$  é dada por:

$$c_R(v_k) = \frac{\sum_{v_j \in V} (diam(G) + 1 - dist(v_k, v_j))}{n - 1}, \quad (3.6)$$

onde  $diam(G)$  denota o diâmetro do grafo.

Tal medida indica o quanto  $v_k$  está bem integrado na rede modelada pelo grafo  $G$ .

Por fim, STEPHENSON e ZELEN [40], em 1989, introduziram a **centralidade de informação** de um vértice que é definida como a média harmônica da informação contida em todos os possíveis caminhos tendo  $v_i$  como vértice terminal.

Matematicamente, a centralidade de informação é definida da seguinte maneira:



**Definição 3.7.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_i$  um vértice de  $G$ . A *centralidade de informação* de  $v_i$  é dada por:

$$c_I(v_i) = \frac{n}{\sum_{j=1}^n \frac{1}{I_{ij}}}, \quad (3.7)$$

onde  $I_{ij} = (c_{ii} + c_{jj} - 2c_{ij})^{-1}$  e  $C = [c_{ij}] = (L + J)^{-1}$ , sendo  $L$  a matriz Laplaciana e  $J$  a matriz de mesma ordem com todas as entradas iguais a 1.

**Exemplo:**

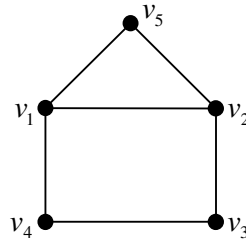


Figura 3.4: Grafo  $G$  onde  $v_1$  e  $v_2$  são os mais centrais segundo  $c_I$ .

Para ilustrar a medida de centralidade de informação vamos considerar o grafo dado na Figura 3.4. A centralidade de informação é calculada através da matriz  $C$ , obtida a partir das matrizes  $L$  e  $J$ :

$$L = \begin{bmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ -1 & 0 & -1 & 2 & 0 \\ -1 & -1 & 0 & 0 & 2 \end{bmatrix}, \quad J = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$C = \frac{1}{275} \begin{bmatrix} 76 & 1 & -29 & -4 & 11 \\ 1 & 76 & -4 & 29 & 11 \\ -29 & -4 & 116 & 16 & -44 \\ -4 & 29 & 16 & 116 & -44 \\ 11 & 11 & -44 & -44 & 121 \end{bmatrix}.$$

A informação contida em todos os caminhos tendo  $v_1$  como vértice terminal é:

$$I_{11} = \infty; I_{12} = 1,833...; I_{13} = 1,1; I_{14} = 1,375 \text{ e } I_{15} = 1,571.$$

Pela equação (3.7), a centralidade de informação do vértice  $v_1$  é dada por:

$$c_I(v_1) = \frac{5}{\sum_{j=1}^5 \frac{1}{I_{ij}}} = \frac{5}{2,817} = 1,774.$$

De maneira análoga, determinamos as centralidades de informação dos demais vértices,

$$c_I(v_2) = 1,774; \quad c_I(v_3) = c_I(v_4) = 1,41 \text{ e } c_I(v_5) = 1,375.$$

Tal medida indica que os vértices  $v_1$  e  $v_2$  são então os mais centrais.

### 3.3 Centralidade de Intermediação

Ao final da década de 70, MOXLEY e MOXLEY [56] observaram que as redes sociais do mundo real são sempre muito grandes e, geralmente, desconexas. Portanto, as medidas conhecidas, e até então apresentadas, não poderiam ser aplicadas. Motivado pelo artigo daqueles autores, FREEMAN [22], em 1977, propôs algumas medidas que, embora sejam também baseadas em interligações entre pares de vértices, **poderiam ser aplicadas a grafos desconexos.**

Para isso, Freeman primeiro introduziu o conceito de intermediação parcial de um vértice numa dada rede para então chegar a um valor que pudesse medir a centralidade deste vértice. Isto significa que este valor deveria ser capaz de expressar a influência que o indivíduo (representado pelo vértice) poderia exercer sobre os seus pares numa rede de comunicação do mundo real. Assim surgiu a seguinte definição:

**Definição 3.8.** Seja  $G$  um grafo (conexo ou não) com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . Considere um par de vértices  $v_i$  e  $v_j$  em  $G$ , tal que  $i \neq j$ ,  $i \neq k$  e  $j \neq k$ . A *intermediação parcial* de  $v_k$  com respeito a  $v_i$  e  $v_j$  é dada por:

$$b_{ij}(v_k) = \begin{cases} 0, & \text{se não existir caminho entre } v_i \text{ e } v_j; \\ \frac{g_{ij}(v_k)}{g_{ij}}, & \text{caso contrário,} \end{cases} \quad (3.8)$$

onde  $g_{ij}$  denota o número de geodésicas entre  $v_i$  e  $v_j$  e  $g_{ij}(v_k)$  denota o número de geodésicas entre  $v_i$  e  $v_j$  que passam por  $v_k$ .

Para se determinar a centralidade de intermediação de um vértice  $v_k$  num grafo  $G$  com  $n$  vértices, tudo o que precisamos fazer é somar todas as intermediações parciais de  $v_k$  em  $G$ . O valor encontrado será denotado por  $c_B(v_k)$  e definido como a centralidade de intermediação de  $v_k$  em  $G$ . Assim, temos:

$$c_B(v_k) = \sum_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} b_{ij}(v_k). \quad (3.9)$$

Para redes enormes, tanto o procedimento de identificação, quanto o de contagem de geodésicas, pode ser um trabalho tedioso. Isto acontece quando a ordem  $n$  de  $G$  for um número muito grande. Felizmente, existem algoritmos muito rápidos para executar estas tarefas, e o primeiro deles, baseado em métodos matriciais, foi desenvolvido por Harary e Flament [22]. NEWMAN [33], em 2005, observou que a centralidade de intermediação  $c_B(v_k)$  definida por (3.9) é uma das mais simples e, provavelmente, uma das mais úteis baseadas na inter-relação entre os vértices de  $G$  e, como já dissemos,  $c_B(v_k)$  indica o potencial que um ponto da rede, no caso  $v_k$ , tem para controlar o fluxo de informações entre os pares de vértices da rede.

FREEMAN, em seu artigo [22], determinou um limite superior para  $c_B(v_k)$  e mostrou que tal limite é alcançado somente quando o grafo  $G$  é uma estrela  $S_n$  e o vértice  $v_k$  é o único vértice com grau  $n-1$  em  $S_n$ . A seguir enunciaremos o Teorema de Freeman e apresentaremos uma prova mais detalhada para este resultado.

**Teorema 3.1.** Seja  $G$  um grafo conexo ou não com  $n \geq 3$  vértices. Seja  $v_k$  um vértice de  $G$  e  $c_B(v_k)$  a centralidade de intermediação de  $v_k$ . Então

$$c_B(v_k) \leq \frac{n^2 - 3n + 2}{2}, \quad (3.10)$$

e o limite é alcançado se e somente se  $G$  é o grafo estrela  $S_n$  e  $v_k$  é o vértice central da estrela.

**Prova:** Seja  $G$  um grafo com  $n \geq 3$  vértices. Primeiramente, vamos considerar  $G$  conexo. Neste caso, existe um caminho entre qualquer par de vértices, e assim,

$r = \binom{n}{2} = \frac{n(n-1)}{2}$  é o número total de pares não-ordenados de vértices  $v_i$  e  $v_j$  tal que

existe um caminho entre eles. Consideremos, agora, que  $G$  tenha um vértice de grau  $n-1$  e chamemos este vértice de  $v_k$ . Neste caso, a estrela  $S_n$  é um subgrafo gerador de  $G$  e qualquer par de vértices  $v_i$  e  $v_j$  são ligados por uma geodésica de comprimento 2 que passa por  $v_k$ . Só existirá uma geodésica para cada par de vértices de  $S_n$ . Portanto, da definição 3.8, segue-se que  $b_{ij}(v_k) = 1$ . Para calcularmos então a centralidade de intermediação do vértice  $v_k$  em  $S_n$ ,  $c_B(v_k)$ , basta utilizarmos a fórmula (3.9) que é

dada por  $c_B(v_k) = \sum_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} b_{ij}(v_k)$ . Tal somatório tem exatamente  $\binom{n}{2} - (n-1)$  parcelas.

Como cada parcela é igual a 1, teremos que:

$$c_B(v_k) = \frac{n(n-1)}{2} - (n-1) = \frac{n^2 - 3n + 2}{2}.$$

Provamos então que  $S_n$  é um grafo que atinge este valor para a centralidade de intermediação.

Para provarmos que  $S_n$  é o único grafo que tem um vértice central que satisfaz (3.10) na igualdade, vamos supor que  $G$  seja conexo e  $G \neq S_n$ . A estrela  $S_n$  é o único grafo em que o vértice de grau máximo igual a  $n-1$  é vértice terminal de toda aresta. Tome em  $G$  o vértice de grau máximo  $v_l$ . Como  $G \neq S_n$ , existirá em  $G$  uma aresta cujos vértices terminais  $v_i$  e  $v_j$  são diferentes de  $v_l$ . Logo é fácil ver que  $b_{ij}(v_l) = 0$ . Temos que  $b_{ij}(v_l)$  é uma das parcelas da centralidade de intermediação de  $v_l$  ligando  $v_i$  a  $v_j$ . Cada uma das outras parcelas poderá ser no máximo igual a 1. Logo a fórmula (3.10)

resultará um valor inferior a  $\frac{n^2 - 3n + 2}{2}$ .

Resta-nos considerar, agora o grafo  $G$  desconexo. Tal grafo terá no mínimo duas componentes conexas. Tomemos um vértice qualquer  $v_k$  em  $G$  e dois vértices  $v_i$  e  $v_j$  não adjacentes em  $G$ , para os quais  $i \neq j$ ,  $i \neq k$  e  $j \neq k$ . Se  $v_i$  e  $v_j$  estiverem em componentes conexas distintas,  $v_i$  e  $v_j$  são não alcançáveis, e pela definição (3.8) teremos  $b_{ij}(v_k) = 0$ . Assim, a fórmula dada por  $c_B(v_k)$  terá necessariamente alguma parcela nula e as outras serão no máximo igual a 1. Portanto, o valor máximo alcançado por  $c_B(v_k)$  terá que ser estritamente inferior a  $\frac{n^2 - 3n + 2}{2}$ , que como já vimos, é a maior soma possível alcançada por (3.10), quando todas as suas parcelas valem 1. Falta considerar  $v_i$  e  $v_j$  na mesma componente conexa. Neste caso, temos um subgrafo  $H$  de

$G$  conexo com  $t$  vértices,  $t < n$ . Então, no caso máximo teremos

$$\frac{t^2 - 3t + 2}{2} < \frac{n^2 - 3n + 2}{2}.$$

□

Sendo assim, uma medida relativa da centralidade de intermediação de um vértice  $v_k$  pode ser definida como:

**Definição 3.9.** Seja  $G$  um grafo (conexo ou não) com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade relativa de intermediação* de  $v_k$  é dada por:

$$c'_B(v_k) = \frac{2c_B(v_k)}{n^2 - 3n + 2}. \quad (3.11)$$

Pelo Teorema 3.1, também neste caso  $0 \leq c'_B(v_k) \leq 1$ .

**Observação:** Em um grafo com  $m$  arestas e  $n$  vértices, a medida de centralidade de intermediação pode ser calculada para todos os vértices com complexidade de  $O(mn)$ , [39].

**Exemplo:**

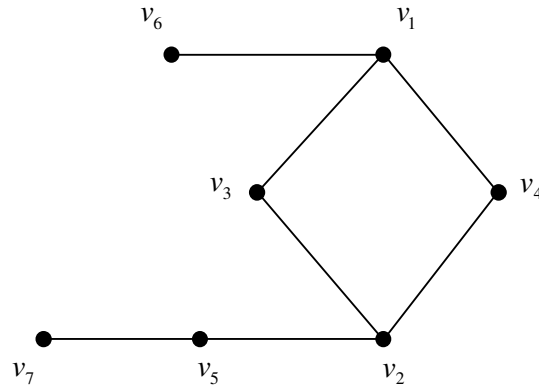


Figura 3.5: Grafo  $G$  onde  $v_2$  é o vértice mais central segundo  $c_B$ .

Na Figura 3.5, todas as geodésicas entre  $v_i$  e  $v_j$ ,  $v_i, v_j \in V$ ,  $i < j = 2, \dots, 7$ , bem como aquelas que passam por  $v_1$  são:

$$g_{23} = 1 \text{ e } g_{23}(v_1) = 0; \quad g_{24} = 1 \text{ e } g_{24}(v_1) = 0; \quad g_{25} = 1 \text{ e } g_{25}(v_1) = 0; \quad g_{26} = 2 \text{ e } g_{26}(v_1) = 2;$$

$$g_{27} = 1 \text{ e } g_{27}(v_1) = 0; \quad g_{34} = 2 \text{ e } g_{34}(v_1) = 1; \quad g_{35} = 1 \text{ e } g_{35}(v_1) = 0; \quad g_{36} = 1 \text{ e } g_{36}(v_1) = 1;$$

$$g_{37}=1 \text{ e } g_{37}(v_1)=0; g_{45}=1 \text{ e } g_{45}(v_1)=0; g_{46}=1 \text{ e } g_{46}(v_1)=1; g_{47}=1 \text{ e } g_{47}(v_1)=0;$$

$$g_{56}=2 \text{ e } g_{56}(v_1)=2; g_{57}=1 \text{ e } g_{57}(v_1)=0; g_{67}=2 \text{ e } g_{67}(v_1)=2.$$

Assim, a medida de centralidade de intermediação do vértice  $v_1$  é:

$$c_B(v_1) = \sum_{\substack{1 \leq i < j \leq 7 \\ i, j \neq 1}} b_{ij}(v_1) = \frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{2}{2} + \frac{0}{1} + \frac{1}{2} + \frac{0}{1} + \frac{1}{1} + \frac{0}{1} + \frac{0}{1} + \frac{1}{1} + \frac{0}{1} + \frac{2}{2} + \frac{0}{1} + \frac{2}{2} = \frac{11}{2} = 5,50.$$

Pelo mesmo raciocínio, teremos para os demais vértices:

$$c_B(v_2) = 8,50; c_B(v_3) = c_B(v_4) = 3; c_B(v_5) = 5 \text{ e } c_B(v_6) = c_B(v_7) = 0.$$

Portanto, a centralidade de intermediação indica o vértice  $v_2$  como sendo o mais central.

E usando a medida relativa teremos:

$$c'_B(v_1) = 0,366...; c'_B(v_2) = 0,566...; c'_B(v_3) = c'_B(v_4) = 0,2; c'_B(v_5) = 0,333... \text{ e}$$

$$c'_B(v_6) = c'_B(v_7) = 0.$$

### 3.3.1 Centralidade de Intermediação de Percursos Aleatórios

A centralidade de intermediação pode ser vista como uma medida da influência que um nó tem sobre a propagação do fluxo de informação (ou de qualquer conteúdo) através da rede. A Definição 3.9, proposta por FREEMAN [21], supõe que o conteúdo, considerado aqui como um fluxo unitário, percorre a rede somente ao longo das possíveis geodésicas. Entretanto, em grande parte das redes, o fluxo pode ser governado por outras regras que consideram caminhos geodésicos e não geodésicos.

Baseado nesta idéia mais realística de propagação da informação, BRANDES e FLEISCHER [14] introduziram uma medida de centralidade de intermediação baseada no fluxo de corrente em circuitos elétricos. Veja uma discussão mais detalhada em [14].

Apesar de seu reconhecimento, a centralidade baseada em fluxo de corrente não é freqüentemente utilizada por cientistas sociais, porque não há nenhuma razão especial para acreditar que tal fluxo tenha algo haver com processos em redes não elétricas. Sendo assim, NEWMAN [33] propôs uma derivação desta medida, chamada *centralidade de intermediação de percursos aleatórios*, fornecendo assim conceitos bem conhecidos e relevantes para redes sociais, assim como para outros tipos de redes. No entanto, o próprio Newman, em [33], mostra que as duas medidas são numericamente iguais. Tal medida considera um *percurso aleatório absorvente* que é um percurso aleatório que começa em  $v_s$  (chamado *fonte*), tem extremidade em  $v_t$  (chamado *alvo*) e continua vértice a vértice com  $v_j \neq v_t$ , escolhendo uma aresta  $e_{ij} \in E$  incidente a  $v_i$  com probabilidade

$$m_{ij} = \frac{a_{ij}}{d_j}, \quad (3.12)$$

para continuar o percurso até atingir o alvo  $v_t$ .

A matriz das probabilidades de transição entre os vértices é denotada por  $M$  e, em notação de matrizes, podemos escrever  $M = AD^{-1}$ , onde  $D$  é a matriz diagonal dos graus dos vértices. Como observado por Newman, o percurso aleatório é absorvente, ou seja,  $m_{ti} = 0$ , para todo  $v_i \in V$ . Sendo assim, a maneira mais conveniente para representar esta situação, segundo Newman, é justamente remover o vértice  $v_t$  do grafo, o que corresponde a removermos a  $t$ -ésima linha e a  $t$ -ésima coluna de  $M$ , não afetando as demais transições. Vamos denotar por  $M_t = A_t \cdot D_t^{-1}$ , a matriz com estes elementos removidos. Na Seção 2.2, vimos que a matriz  $(A^r)_{ij}$  fornece o número de percursos de comprimento  $r$  entre  $v_i$  e  $v_j$ . Entretanto, não estamos interessados no número de percursos aleatórios, mas na probabilidade que um percurso aleatório de



início em  $v_s$ , após  $r$  etapas, esteja em  $v_j$ . Isto é dado por  $[M_t^r]_{js}$ . Assim, a probabilidade que na etapa  $r+1$  estejamos em um vértice adjacente  $v_i$  é dada por  $d_j^{-1}[M_t^r]_{js}$ .

Note que todas as entradas de qualquer matriz  $M_t^r$  são valores entre 0 e 1 e, assim, o somatório sobre todos os possíveis percursos de comprimento 0 até  $\infty$ , é convergente [31]. Isto significa que:

$$\sum_{r=0}^{\infty} d_j^{-1}[M_t^r]_{js} = d_j^{-1}[(I_{n-1} - M_t)^{-1}]_{js}. \quad (3.13)$$

A equação (3.13) em notação de matriz nos dá:

$$p_{st}(v) = D_t^{-1}(I - M_t)^{-1}.s = (D_t - A_t)^{-1}.s, \quad (3.14)$$

onde  $s$  é um vetor fonte definido por

$$s_i = \begin{cases} 1, & \text{para } i = s; \\ 0, & \text{caso contrário.} \end{cases}$$

Logo, a probabilidade que um percurso aleatório passe ao longo da aresta entre  $v_i$  e  $v_j$  é dado pela diferença absoluta entre as probabilidades  $|p_{st}(v_i) - p_{st}(v_j)|$  e a probabilidade de passar através do vértice  $v_i$  é a metade da soma das probabilidades nas arestas incidentes a ele:

$$\tau_{st}(v_i) = \frac{1}{2} \sum_j a_{ij} |p_{st}(v_i) - p_{st}(v_j)|, \text{ para } v_i \neq v_s, v_t. \quad (3.15)$$

Quando  $v_i = v_s$  ou  $v_i = v_t$ ,  $\tau_{st}(v_i) = 1$ . Assim, a centralidade de intermediação do vértice  $v_i$  em percursos aleatórios entre  $v_s$  e  $v_t$  pode ser definida da seguinte maneira:

**Definição 3.10.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_i$  um vértice de  $G$ . A *centralidade de intermediação de  $v_i$  em percursos aleatórios* é dada por:

$$c_{CA}(v_i) = \frac{2 \sum_{s \leq t} \tau_{st}(v_i)}{n(n-1)}. \quad (3.16)$$

**Observação:** Conforme visto em [33], em um grafo com  $m$  arestas e  $n$  vértices, a centralidade de intermediação de percursos aleatórios pode ser calculada para todos os vértices com complexidade  $O((m+n)n^2)$ .

**Exemplo:** Tomemos novamente o grafo da Figura 3.5.

Para o grafo  $G$ , a partir de suas matrizes de adjacência e dos graus dos vértices, encontramos:

$$L = D - A = \begin{bmatrix} 3 & 0 & -1 & -1 & 0 & -1 & 0 \\ 0 & 3 & -1 & -1 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 2 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}; (D_3 - A_3)^{-1} = \begin{bmatrix} 3/4 & 1/4 & 1/2 & 1/4 & 3/4 & 1/4 \\ 1/4 & 3/4 & 1/2 & 3/4 & 1/4 & 3/4 \\ 1/2 & 1/2 & 1 & 1/2 & 1/2 & 1/2 \\ 1/4 & 3/4 & 1/2 & 7/4 & 1/4 & 7/4 \\ 3/4 & 1/4 & 1/2 & 1/4 & 7/4 & 1/4 \\ 1/4 & 3/4 & 1/2 & 7/4 & 1/4 & 11/4 \end{bmatrix}.$$

Considerando-se  $v_1$  como vértice fonte e  $v_3$  como vértice alvo e aplicando a equação (3.14) temos:

$$p_{13}(v_i) = (D_3 - A_3)^{-1} \cdot [1 \ 0 \ 0 \ 0 \ 0 \ 0]^t = [3/4 \ 1/4 \ 1/2 \ 1/4 \ 3/4 \ 1/4]^t.$$

Sendo assim, a probabilidade que um percurso aleatório entre  $v_1$  e  $v_3$  passe pelos vértices  $v_2, v_4$  e  $v_5$  é dada por:

$$p_{13}(v_2) = 1/4; \ p_{13}(v_4) = 1/2 \text{ e } p_{13}(v_5) = 1/4.$$

Portanto, a probabilidade que um percurso aleatório do vértice  $v_1$  ao vértice  $v_3$  passe através de  $v_2$  é:

$$\tau_{13}(v_2) = \frac{1}{2} \sum_{j=1}^n a_{2j} |p_{13}(v_2) - p_{13}(v_j)| = \frac{1}{2} [|p_{13}(v_2) - p_{13}(v_3)| + |p_{13}(v_2) - p_{13}(v_4)| + |p_{13}(v_2) - p_{13}(v_5)|] = \frac{1}{4}.$$

Analogamente, fazendo o cálculo para todos os pares  $(v_s, v_t)$  possíveis cujo percurso passe por  $v_2$  temos:

$$\tau_{12}(v_2) = 1; \tau_{14}(v_2) = \frac{1}{4}; \tau_{15}(v_2) = 1; \tau_{16}(v_2) = 0; \tau_{17}(v_2) = 1; \tau_{23}(v_2) = 1; \tau_{24}(v_2) = 1;$$

$$\tau_{25}(v_2) = 1; \tau_{26}(v_2) = 1; \tau_{27}(v_2) = 1; \tau_{34}(v_2) = \frac{1}{2}; \tau_{35}(v_2) = 1; \tau_{36}(v_2) = \frac{1}{4}; \tau_{37}(v_2) = 1;$$

$$\tau_{45}(v_2) = 1; \tau_{46}(v_2) = \frac{1}{4}; \tau_{47}(v_2) = 1; \tau_{56}(v_2) = 1; \tau_{57}(v_2) = 0 \text{ e } \tau_{67}(v_2) = 1. \text{ Portanto,}$$

$$\sum_{s < t} \tau_{st}(v_2) = \frac{31}{2}.$$

Logo, a medida de centralidade de intermediação do vértice  $v_2$  em percursos aleatórios é:

$$c_{CA}(v_2) = \frac{2 \sum_{s < t} \tau_{st}(v_2)}{n(n-1)} = \frac{31}{42} = 0,738.$$

Utilizando-se o mesmo raciocínio, podemos calcular tal medida para os demais vértices:

$$c_{CA}(v_1) = 0,619; c_{CA}(v_3) = c_{CA}(v_4) = 0,488; c_{CA}(v_5) = 0,523; c_{CA}(v_6) = c_{CA}(v_7) = 0,285,$$

e veremos que, de fato,  $v_2$  é aquele mais central.

Note que para o mesmo grafo da Figura 3.5, a relevância dos vértices obtida fazendo uso da centralidade de intermediação que utiliza apenas caminhos geodésicos é a mesma que a centralidade de intermediação de percursos aleatórios. Entretanto, no Capítulo 6 veremos que diferentes medidas podem nos conduzir a diferentes resultados, o que mostra a importância de se avaliar também o tipo de rede que estamos modelando, a fim de se utilizar uma medida mais apropriada.

### 3.4 Centralidade de Grafos

Além das medidas de centralidade dos vértices de um grafo até aqui apresentadas, intuitivamente parece claro que alguns grafos são mais dependentes dos vértices mais centrais que outros. O grafo estrela, por exemplo, é extremamente dependente de seu vértice mais central, e não importa a medida de centralidade considerada, tal vértice tem medida superior aos demais, e se suprimido, tornará o grafo estrela um grafo nulo, razão pela qual é dito ser o grafo mais central. É nesse sentido que FREEMAN [21], em 1978, propôs uma medida de centralidade de grafos, a ser calculada em relação aos valores de algumas das principais medidas de centralidade dos vértices. Tal medida nos permite comparar grafos de mesma ordem mas topologicamente diferentes. Antes disso, seja  $x$  uma das medidas de centralidade dos vértices de um grafo (grau, proximidade ou intermediação). As definições a seguir valem para qualquer grafo  $G$ , conexo ou não, exceto quando a medida  $x$  for a centralidade de proximidade, pois neste caso,  $G$  tem que ser conexo:

**Definição 3.11.** O *centro* de um grafo  $G = (V, E)$  é o subconjunto de vértices com valor máximo no que diz respeito à medida de centralidade  $x$  considerada, ou seja,

$$\varphi_x(G) = \{v_k \in V; c_x(v_k) \geq c_x(v_j), \forall v_j \in V\} \quad (3.17)$$

**Definição 3.12.** Seja  $G$  um grafo com  $n$  vértices,  $v_j$  um vértice de  $G$  e  $c_x(v_j)$  a centralidade do vértice  $v_j$  em  $G$  no que diz respeito à medida de centralidade  $x$ . Seja

$v_k \in \varphi_x(G)$  e considere  $\Phi_x(G) = \sum_{j=1}^n (c_x(v_k) - c_x(v_j))$ . A *centralidade do grafo*  $G$  é

dada por:

$$c_x(G) = \frac{\Phi_x(G)}{\max_{H \in \gamma} \Phi_x(H)}, \quad (3.18)$$

onde  $\gamma$  é a coleção de todos os grafos  $H$  de  $n$  vértices para o qual  $\Phi_x(H)$  está definido.

Observe que, quando  $x$  for a medida de centralidade de proximidade,  $G$  e  $H$  têm que ser conexos.

O próximo Teorema, cuja prova pode ser encontrada em [21], dá limites superiores para  $\Phi_x(G)$ , em função das centralidades de grau, intermediação e proximidade dos vértices de  $G$ . Além disso, mostra que dentre todos os grafos de mesma ordem, o grafo estrela é aquele que satisfaz os limites.

**Teorema 3.7.** Seja  $G$  um grafo com  $n \geq 3$  vértices,  $D$ ,  $B$  e  $C$  as medidas de centralidade de grau, intermediação e proximidade, respectivamente. Então,

1.  $\Phi_D(G) \leq n^2 - 3n + 2$ ;
2.  $\Phi_B(G) \leq n^3 - 4n^2 + 5n - 2$ ;
3.  $\Phi_C(G) \leq \frac{n^2 - 3n + 2}{2n - 3}$ ;

e os limites são alcançados se e somente se  $G$  é o grafo estrela  $S_n$ .

A medida  $c_x(G)$  estará sempre entre 0 e 1.  $c_x(G)$  será igual a 0 quando todos os vértices tiverem exatamente as mesmas centralidades, por exemplo se  $G = K_n$ , e será igual a 1 quando um único vértice “dominar” completamente os demais vértices, como é o caso de  $G = S_n$  (veja o Teorema 3.7). Isto significa que um grafo  $G$  é tanto mais central quanto mais próximo de 1 for sua medida de centralidade.

**Exemplo:**

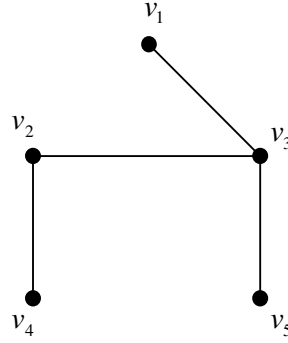


Figura 3.6: grafo  $H_1$

Para o grafo  $H_1$  dado na Figura 3.6, se a medida  $x$  for a centralidade de grau dos vértices considerada, temos  $d_1 = d_4 = d_5 = 1$ ,  $d_2 = 2$  e  $d_3 = 3$ . Dado que  $S_5$  é o grafo de 5 vértices, cujo máximo valor é dado por  $\Phi_D(S_5) = 12$ , utilizando-se a equação (3.18) temos que a centralidade do grafo  $H_1$  é dada por:

$$c_D(H_1) = \frac{\Phi_D(H_1)}{\Phi_D(S_5)} = \frac{\sum_{j=1}^5 (c_D(v_k) - c_D(v_j))}{n^2 - 3n + 2} = \frac{3 \cdot (3-1) + (3-2)}{12} = \frac{7}{12} = 0,583...$$

Para o mesmo grafo da Figura 3.6, considere agora  $x$  como medida de centralidade de proximidade dos vértices. Assim,  $c_c(v_1) = 0,125$ ;  $c_c(v_2) = 0,166...$ ;  $c_c(v_3) = 0,2$ ;  $c_c(v_4) = 0,111...$  e  $c_c(v_5) = 0,125$ . Novamente, o valor máximo para o denominador da equação (3.18) ocorre quando  $G = S_n$  e, neste caso, segue-se do Teorema 3.7 que

$$\max \Phi_C(G) = \max \Phi_C(S_5) = \frac{n^2 - 3n + 2}{2n - 3}. \text{ Portanto, } \Phi_C(S_5) = 1,714. \text{ Logo, de (3.18),}$$

temos:

$$c_C(H_1) = \frac{\sum_{j=1}^5 (c_C(v_k) - c_C(v_j))}{\frac{n^2 - 3n + 2}{2n - 3}} = \frac{2 \cdot (0,2 - 0,125) + (0,2 - 0,166...) + (0,2 - 0,111...)}{1,714} =$$

$$\frac{0,272...}{1,714} = 0,158.$$

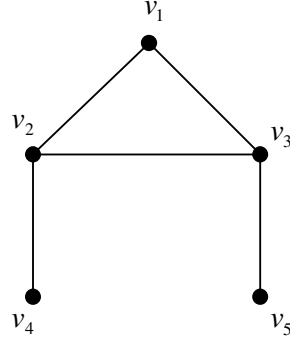


Figura 3.7: grafo  $H_2$

Vamos considerar agora o grafo da Figura 3.7, tendo  $x$  como medida de centralidade de grau dos vértices. Logo,  $d_1 = 2$ ,  $d_2 = d_3 = 3$  e  $d_4 = d_5 = 1$ . Neste caso,

$$c_D(H_2) = \frac{\sum_{j=1}^5 (c_D(v_k) - c_D(v_j))}{n^2 - 3n + 2} = \frac{2 \cdot (3 - 1) + (3 - 2)}{12} = \frac{5}{12} = 0,416...$$

Ainda, considerando-se o grafo dado na Figura 3.7, e tendo agora  $x$  como medida de centralidade de proximidade dos vértices,  $c_C(v_1) = 0,166...$ ;  $c_C(v_2) = c_C(v_3) = 0,2$  e  $c_C(v_4) = c_C(v_5) = 0,125$ . Neste caso,

$$c_C(H_2) = \frac{\sum_{j=1}^5 (c_C(v_k) - c_C(v_j))}{\frac{n^2 - 3n + 2}{2n - 3}} = \frac{2 \cdot (0,2 - 0,125) + (0,2 - 0,166...)}{1,714} = \frac{0,183...}{1,714} = 0,106.$$

Portanto, temos que  $c_D(H_1) > c_D(H_2)$  e  $c_C(H_1) > c_C(H_2)$ . Assim, podemos dizer que em relação às centralidades de grau e proximidade, o grafo  $H_1$  é mais central que o grafo  $H_2$ .

## Capítulo 4 – Medidas de Centralidade Espectrais

As medidas de centralidade espectrais buscam obter propriedades estruturais dos vértices de um grafo a partir das propriedades dos autovalores e autovetores das matrizes associadas a estes grafos, mediante conhecimentos de Álgebra Linear. Assim, serão apresentadas aqui duas medidas de centralidade que avaliam o grau de relevância dos vértices a partir da representação de um grafo por uma matriz. A *centralidade de autovetor* atribui alta relevância para um vértice, em função da sua relação com seus vizinhos, ou seja, mesmo se um vértice  $v_k$  está ligado somente a outro vértice  $v_i$  (tendo assim uma baixa centralidade de grau), os vizinhos de  $v_i$  podem ser importantes, e consequentemente, o vértice  $v_k$  também será importante, obtendo uma elevada centralidade de autovetor. Já a *centralidade via conectividade algébrica*, como será visto em 4.4, mede o grau de relevância de um vértice em relação a vulnerabilidade que ele oferece à rede caso tenha de ser dela retirado. Tais medidas de centralidade, agregadas àquelas já apresentadas ao longo do texto, são uma importante ferramenta para o estudo e análise de redes.

### 4.1 Centralidade de Autovetor

O Teorema de Perron-Frobenius, apresentado na Seção 2.2, garante que em grafos conexos, o raio espectral coincide com o índice do grafo, isto é,  $\rho(G) = ind(G)$ , pois grafos conexos têm matrizes de adjacência irredutíveis. Além disso, este Teorema garante que  $\rho(G) > 0$  e, como autovalor, tem multiplicidade algébrica igual a 1. Portanto, a multiplicidade geométrica é também igual a 1, isto é, o espaço gerado pelos autovetores associados ao índice é unidimensional. Pelo Teorema de Perron-Frobenius,



temos ainda que existe um autovetor associado ao índice com todas as coordenadas positivas. Isto nos permite introduzir mais uma medida de centralidade de vértice em um grafo. A medida proposta por BONACICH [6], em 1987, chamada *centralidade de autovetor*, é baseada no conceito de autovalores e autovetores da matriz de adjacência do grafo  $G$ . Podemos defini-la da seguinte maneira:

**Definição 4.1.** Seja  $G$  um grafo conexo com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade de autovetor* de  $v_k$  é dada por:

$$c_{eig}(v_k) = x_k, \quad (4.1)$$

onde  $x_k$  é a  $k$ -ésima coordenada do autovetor positivo unitário  $x$  associado ao índice do grafo, isto é,

$$x_k = \frac{1}{\rho} \sum_{j=1}^n a_{kj} x_j ; k = 1, \dots, n. \quad (4.2)$$

Como a multiplicidade do raio espectral é igual a 1, qualquer outro autovetor positivo associado ao índice será múltiplo escalar deste. Logo, sejam  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$  ambos autovetores positivos associados ao índice de  $G$  tais que  $y = \alpha \cdot x$ . Se  $x_i \leq x_j$  então  $y_i \leq y_j$ .

Muitos autores, como [27], utilizam o termo autocentralidade para indicar a centralidade de autovetor.

**Exemplo:**

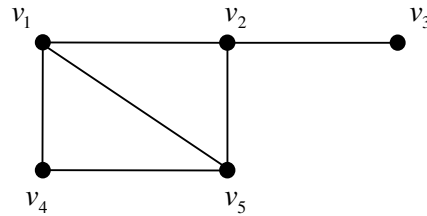


Figura 4.1: Grafo cujos vértices mais centrais segundo a autocentralidade são  $v_1$  e  $v_5$ .

Na Figura 4.1, o polinômio característico do grafo  $G$  e seu respectivo espectro são:

$$p_G(\lambda) = \lambda^5 - 6\lambda^3 - 4\lambda^2 + 3\lambda + 2$$

$$\text{spect } G = \begin{bmatrix} 2,641 & 0,723 & -0,589 & -1 & -1,775 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

O autovetor positivo e de norma 1 associado a  $\rho = 2,641$  é:

$$x = [0,537 \quad 0,474 \quad 0,179 \quad 0,406 \quad 0,537]^t.$$

Assim, as centralidades de autovetor dos respectivos vértices de  $G$  são:

$$c_{\text{eig}}(v_1) = c_{\text{eig}}(v_5) = 0,537; \quad c_{\text{eig}}(v_2) = 0,474; \quad c_{\text{eig}}(v_3) = 0,179; \quad c_{\text{eig}}(v_4) = 0,406.$$

Portanto, a medida indica os vértices  $v_1$  e  $v_5$  como mais centrais do grafo da Figura 4.1.

O significado da centralidade de autovetor como medida de propagação é porque o autovetor define como o vértice mais central aquele que está ligado a outros que por sua vez também estabelecem relações com vértices que estão em posições centrais, e assim por diante. A centralidade de autovetor de um vértice é uma combinação linear das centralidades dos vértices com ele conectados [35], e desta forma, pode ser justificado pelo Teorema a seguir, conhecido como *Método das Potências* [13]. Em linhas gerais, o método consiste em determinar o autovalor de maior valor absoluto de uma matriz e seu correspondente autovetor de maneira aproximada. A razão para a escolha deste método é que, podemos deduzir fatos interessantes sobre a propagação de informações, tecnologias ou doenças através da rede.

**Teorema 4.1. (Método das Potências)** Seja  $A$  uma matriz real de ordem  $n$  e sejam  $\lambda_1, \lambda_2, \dots, \lambda_n$  seus autovalores e  $x_1, x_2, \dots, x_n$  seus correspondentes autovetores. Suponha que os autovetores são linearmente independentes, e que:

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Seja a sequência  $y_k$  definida por:

$$y_{k+1} = Ay_k, \quad k = 0, 1, 2, \dots$$

onde  $y_0$  é um vetor arbitrário não nulo, que permite a expansão:

$$y_0 = \sum_{j=1}^n c_j x_j$$

com  $c_j$  escalares quaisquer e  $c_1 \neq 0$ . Então,

$$\lim_{k \rightarrow \infty} \frac{(y_{k+1})_r}{(y_k)_r} = \lambda_1,$$

onde o índice  $r$  indica a  $r$ -ésima coordenada dos vetores  $y_{k+1}$  e  $y_k$ . Além disso, quando  $k \rightarrow \infty$ ,  $y_k$  tende ao autovetor correspondente a  $\lambda_1$ .

**Prova:** Por hipótese,  $y_0 = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$ , com  $c_1 \neq 0$ .

Mas,  $Ax_i = \lambda_i x_i$  e  $y_1 = Ay_0$ . Então,

$$y_1 = c_1 Ax_1 + c_2 Ax_2 + \dots + c_n Ax_n \text{ e } y_1 = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \dots + c_n \lambda_n x_n. \text{ Assim,}$$

$$y_1 = \lambda_1 \left[ c_1 x_1 + c_2 \frac{\lambda_2}{\lambda_1} x_2 + \dots + c_n \frac{\lambda_n}{\lambda_1} x_n \right].$$

Por outro lado,  $y_2 = Ay_1 = A^2 y_0$ . Logo,

$$y_2 = \lambda_1 \left[ c_1 Ax_1 + c_2 \frac{\lambda_2}{\lambda_1} Ax_2 + \dots + c_n \frac{\lambda_n}{\lambda_1} Ax_n \right] \text{ que, por sua vez, é}$$

$$y_2 = \lambda_1 \left[ c_1 \lambda_1 x_1 + c_2 \frac{\lambda_2}{\lambda_1} \lambda_2 x_2 + \dots + c_n \frac{\lambda_n}{\lambda_1} \lambda_n x_n \right]. \text{Portanto,}$$

$$y_2 = \lambda_1^2 \left[ c_1 x_1 + c_2 \left( \frac{\lambda_2}{\lambda_1} \right)^2 x_2 + \dots + c_n \left( \frac{\lambda_n}{\lambda_1} \right)^2 x_n \right]. \text{Repetindo-se este procedimento,}$$

obtém-se, para todo  $k \geq 2$ ,  $y_k = A y_{k-1} = A^k y_0$  e

$$y_k = \lambda_1^k \left[ c_1 x_1 + c_2 \left( \frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + c_n \left( \frac{\lambda_n}{\lambda_1} \right)^k x_n \right].$$

Como por hipótese  $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ , tem-se  $\left| \frac{\lambda_i}{\lambda_1} \right| < 1$ , para  $i = 2, \dots, n$ . Portanto,

quando  $k \rightarrow \infty$ ,  $\left( \frac{\lambda_i}{\lambda_1} \right)^k \rightarrow 0$ .

Daí, segue que:

$$y_k = A^k y_0 \rightarrow \lambda_1^k c_1 x_1 \text{ quando } k \rightarrow \infty.$$

Como  $\lambda_1 \neq 0$  e  $c_1 \neq 0$ ,  $y_k$  se aproxima de um múltiplo não nulo de  $x_1$  (ou seja, de um autovetor correspondente a  $\lambda_1$ ). Além disso,  $\lambda_1$  pode ser assim obtido,

$$\lim_{k \rightarrow \infty} \frac{(y_{k+1})_r}{(y_k)_r} = \lim_{k \rightarrow \infty} \frac{(A^{k+1} y_0)_r}{(A^k y_0)_r} = \lim_{k \rightarrow \infty} \frac{(\lambda_1^{k+1} c_1 x_1)_r}{(\lambda_1^k c_1 x_1)_r} = \lambda_1. \quad \square$$

**Observação:** No limite, todas as coordenadas de  $\frac{(y_{k+1})_r}{(y_k)_r}$  tendem a  $\lambda_1$ . Entretanto,

na prática, uma das coordenadas converge mais rapidamente do que as outras. Assim, quando uma delas satisfizer a precisão desejada teremos o autovalor procurado. Além disso, a velocidade de convergência depende de  $\frac{\lambda_2}{\lambda_1}$ . Portanto, quanto maior for  $|\lambda_1|$ ,

quando comparado com  $|\lambda_2|$ , mais rápida será a convergência, [20].

Para obtermos  $\lambda_1$  com uma precisão  $\varepsilon$ , calculamos, em cada passo, aproximações para  $\lambda_1$  através das razões entre as coordenadas dos vetores  $y_{k+1}$  e  $y_k$ . O teste do erro relativo, dado por

$$\left| \lambda_1^{(k+1)} - \lambda_1^{(k)} \right|_r < \varepsilon,$$

é usado como critério de parada.

Como em todo processo iterativo, este método apresenta um resultado aproximado, que será tão próximo do resultado real, conforme o número de iterações realizadas.

**Exemplo:**

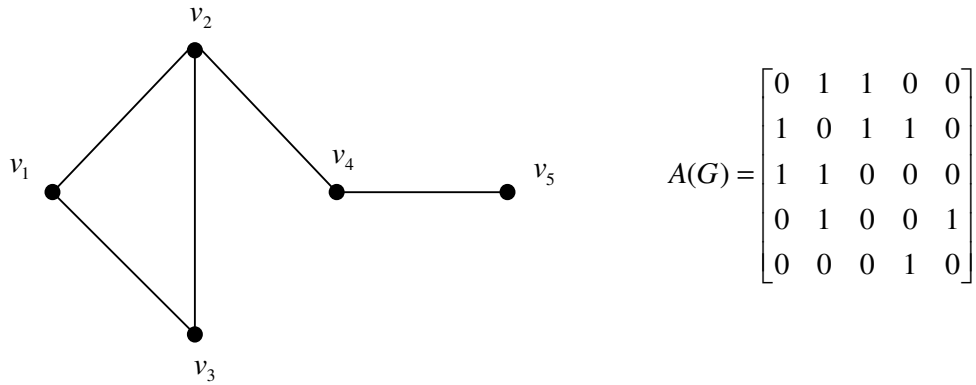


Figura 4.2: Grafo  $G$  e sua respectiva matriz de adjacência.

Utilizando o método das potências com precisão  $10^{-2}$  para os autovetores e autovalores do grafo da Figura 4.2, temos:

1ª iteração: Tomemos o vetor  $y_0 = (1 \ 1 \ 1 \ 1 \ 1)^t$ .

Assim,  $y_1 = Ay_0 = (2 \ 3 \ 2 \ 2 \ 1)^t$ . Veja que as coordenadas de  $y_1$  representam os graus dos vértices de  $G$ . Aqui temos  $(y_1)_2 = 3$  e  $(y_0)_2 = 1$ . Logo, uma primeira aproximação para  $\lambda_1$  é dada por:

$$\lambda_1^{(1)} = \frac{(y_1)_2}{(y_0)_2} = \frac{3}{1} = 3.$$

2ª iteração: Temos,  $y_2 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \\ 5 \\ 4 \\ 2 \end{pmatrix}$ , que representa o grau dos

vizinhos de segunda ordem. Neste caso,  $(y_2)_2 = 6$ ,  $(y_1)_2 = 3$  e, assim, na nova aproximação para  $\lambda_1$  temos:

$$\lambda_1^{(2)} = \frac{(y_2)_2}{(y_1)_2} = \frac{6}{3} = 2.$$

Calculando o erro relativo, obtemos:

$$\left| \lambda_1^{(2)} - \lambda_1^{(1)} \right|_2 = 1 > \varepsilon.$$

Portanto, devemos fazer pelo menos mais uma iteração.

3ª iteração: Seja então  $y_3 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \\ 5 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ 14 \\ 11 \\ 8 \\ 4 \end{pmatrix}$  que representa o grau dos

vizinhos de terceira ordem. Aqui temos  $(y_3)_2 = 14$ ,  $(y_2)_2 = 6$  e, assim, a aproximação para  $\lambda_1$  satisfaz:

$$\lambda_1^{(3)} = \frac{(y_3)_2}{(y_2)_2} = \frac{14}{6} = 2,333\dots$$

Calculando então o erro relativo, obtemos:

$$\left| \lambda_1^{(3)} - \lambda_1^{(2)} \right|_2 = 0,333\dots > \varepsilon.$$

Analogamente, seguimos com o processo fazendo novas iterações até que o critério de parada seja satisfeito. Na Tabela 4.1 vemos que o processo seguirá até a 14ª iteração.

$k$	$y_k$	$\lambda_1^{(k)}$
0	$(1 \ 1 \ 1 \ 1 \ 1)^t$	1
1	$(2 \ 3 \ 2 \ 2 \ 1)^t$	3
2	$(5 \ 6 \ 5 \ 4 \ 2)^t$	2
3	$(11 \ 14 \ 11 \ 8 \ 4)^t$	2,3333
4	$(25 \ 30 \ 25 \ 18 \ 8)^t$	2,14
5	$(55 \ 68 \ 55 \ 38 \ 18)^t$	2,26
6	$(123 \ 148 \ 123 \ 86 \ 38)^t$	2,17
7	$(271 \ 332 \ 271 \ 186 \ 86)^t$	2,24
8	$(603 \ 728 \ 603 \ 418 \ 186)^t$	2,19
9	$(1331 \ 1624 \ 1331 \ 914 \ 418)^t$	2,23
10	$(2955 \ 3576 \ 2955 \ 2042 \ 914)^t$	2,20
11	$(6531 \ 7952 \ 6531 \ 4490 \ 2042)^t$	2,22
12	$(14483 \ 17552 \ 14483 \ 9994 \ 4490)^t$	2,20
13	$(32035 \ 38960 \ 32035 \ 22024 \ 9994)^t$	2,219
14	$(70995 \ 86112 \ 70995 \ 48954 \ 22042)^t$	2,210

Tabela 4.1: Autovalores e autovetores correspondentes em suas respectivas iterações.

A Tabela 4.1 mostra, portanto, o passo dado na 14ª iteração como passo final do algoritmo que consiste na convergência para o autovetor dado por:

$$(70995 \ 86112 \ 70995 \ 48954 \ 22042)^t,$$

cujo índice  $\lambda_1 = 2,21$ .

Note que, embora o autovetor encontrado na 14<sup>a</sup> iteração não seja unitário, ele mantém a mesma ordenação dos vértices em escala de importância. O algoritmo reflete, desta forma, uma justificativa para o uso de autovalores e autovetores como medida de centralidade, tendo em vista que a cada iteração  $k$ , medimos a influência de cada vértice em relação a seus vizinhos de ordem  $k$ , dado pela soma de seus graus. Sendo assim, um vértice é mais central de acordo com a centralidade de autovetor, se ele está ligado a vértices que estão ligados a outros e que, por sua vez, estão ligados a muitos outros.

## 4.2 Coordenadas de um autovetor e suas relações com o grau

Como vimos na Seção 4.1, as coordenadas de um autovetor positivo  $x$  associado ao índice de um grafo  $G$ , não orientado, simples e conexo refletem a importância de cada vértice de acordo com a rotulação considerada no grafo. Alguns resultados relativos a tais coordenadas, bem como suas relações com outras medidas de centralidade, em especial a centralidade de grau, são apresentados a seguir. O primeiro deles, cuja prova se encontra em [27], fornece uma fórmula para comparar duas coordenadas de  $x$ ,  $x_i$  e  $x_j$ , associadas a vértices adjacentes em  $G$ .

**Teorema 4.2.** Seja  $G$  um grafo conexo com  $n$  vértices e  $v_i$  e  $v_j$  dois vértices adjacentes em  $G$ . Sejam  $x_i$  e  $x_j$  coordenadas do autovetor  $x$  associado ao índice  $\rho$  e  $\rho I - A_i$  a submatriz principal de  $\rho I - A$ , obtida retirando-se sua  $i$ -ésima linha e coluna. Então,



$$x_j = x_i \left[ \rho - \sum_{\substack{\{v_k, v_l\} \in E \\ k \neq j}} \frac{\det B_{k,i}}{\det(\rho I - A_i)} \right], \quad (4.3)$$

onde  $B_{k,i}$  é a submatriz de ordem  $n-1$ , obtida da matriz  $(\rho I - A)$ , deletando-se sua  $i$ -ésima linha e coluna e substituindo-se a  $k$ -ésima coluna pelo vetor

$$[a_{1i} \cdots a_{i-1,i} \ a_{i+1,i} \cdots a_{ni}]^t.$$

**Exemplo:**

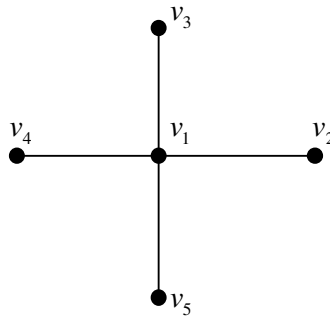


Figura 4.3: Grafo estrela  $S_5$

Considere o grafo  $S_5$  da Figura 4.3 com seus dois vértices adjacentes  $v_1$  e  $v_2$ . É fácil ver que para este grafo  $\rho = 2$  e  $\det(\rho I - A_1) = 2^4 = 16$ . Além disso, para todo  $k = 3, 4, 5$ ,  $\det B_{k,1} = 2^3 = 8$ . Aplicando-se (4.3), obtemos:

$$x_2 = x_1 \left[ 2 - 3 \cdot \frac{8}{16} \right] = 0,5x_1.$$

Analogamente, para  $j = 3, 4, 5$ ,  $x_j = 0,5x_1$ . Portanto, o vértice  $v_1$  (o centro da estrela) é o mais central dado que o valor de  $x_1$  é o dobro das demais coordenadas que tem a mesma centralidade de autovetor.

Como consequência do Teorema 4.2., podemos relacionar as coordenadas de um autovetor  $x$  com os vértices pendentes de um grafo  $G$  conexo. O Corolário 4.1 e uma prova similar à apresentada aqui podem também ser encontrados em [27].

**Corolário 4.1.** Seja  $G$  um grafo conexo com  $n \geq 2$  vértices,  $v_i$  e  $v_j$  dois vértices adjacentes em  $G$ . Sejam  $x_i$  e  $x_j$  coordenadas do autovetor  $x$  associado ao índice  $\rho$  e correspondente aos vértices  $v_i$  e  $v_j$ .

- i) Se  $n > 2$  e  $d_i = 1$  então  $x_j > x_i$ ;
- ii) Se  $n > 2$  e  $d_i > 1$  então  $x_j < \rho x_i$  e
- iii) Se  $n = 2$  então  $x_1 = x_2$ .

**Prova:**

i) Se  $n > 2$ , o somatório na expressão (4.3) é estendido para todos os vértices adjacentes a  $v_i$ , exceto o vértice  $v_j$ . Portanto, se  $d_i = 1$ , de (4.3), segue-se que  $x_j = \rho x_i$ , dado que o somatório em (4.3) é nulo. Como o grafo é conexo e  $n > 2$ , pelo menos um vértice tem grau maior que 1, e portanto, da Proposição 2.1, temos que  $1 = \delta < \bar{d} \leq \rho$ , então  $x_j = \rho x_i > x_i$ .

ii) Se  $n > 2$  e  $d_i > 1$ , no somatório da expressão (4.3), temos que  $\det B_{k,i} > 0$ , para todo vértice  $v_k$  adjacente a  $v_i$ . Novamente da expressão (4.3), segue-se que  $x_j < \rho x_i$ .

iii) No caso em que  $n = 2$  temos:  $\begin{cases} x_2 = \rho x_1 \\ x_1 = \rho x_2 \end{cases} \Rightarrow \rho = 1$  e então  $x_1 = x_2$ .

□

**Exemplo:**

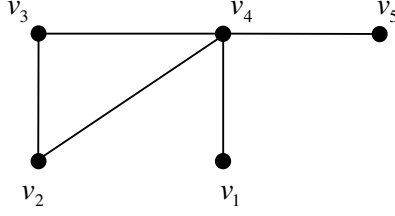


Figura 4.4: Grafo  $G$  com raio espectral  $\rho = 2,342$ .

A Figura 4.4 mostra um grafo  $G$  que possui raio espectral  $\rho = 2,342$  e o autovetor associado

$$x = [0,271 \quad 0,473 \quad 0,473 \quad 0,635 \quad 0,271]^t .$$

Os vértices pendentes  $v_1$  e  $v_5$  correspondem ao caso (i) do Corolário 4.1, verificando-se de fato que tais vértices possuem centralidades de autovetor menor que os demais vértices não pendentes,  $v_2$ ,  $v_3$  e  $v_4$ . Em relação ao caso (ii), temos que para cada vértice  $v_i$  ( $i = 2,3,4$ ) não pendente, vale que  $x_j < \rho x_i$ . De fato, para  $j = 1,3,4,5$ ,  $x_j < \rho x_2 = 1,109$ .

O Teorema 4.3, que virá a seguir, fornece novos limites (superior e inferior) para o raio espectral  $\rho$  em termos das coordenadas do autovetor  $x$  associado a  $\rho$ . Em se tratando de um resultado melhor que, por exemplo, as clássicas desigualdades,  $\delta \leq \rho \leq \Delta$ , apresentaremos aqui sua prova, que também pode ser encontrada em [27].

**Teorema 4.3.** Seja  $G = (V, E)$  um grafo conexo e  $A$  sua matriz de adjacência, com raio espectral  $\rho$  e autovetor correspondente  $x$ . Sejam  $v_k$  e  $v_w$  dois vértices em  $G$ ,  $d_k$  e  $d_w$  seus respectivos graus cujas coordenadas correspondentes em  $x$  são  $x_k$  e  $x_w$ . Sejam os conjuntos  $U = \{v_k \in V : x_k = \max_{v_i \in V} x_i\}$ ,  $W = \{v_w \in V : x_w = \min_{v_i \in V} x_i\}$  e os graus mínimo e

máximo dados por  $d_{\min}(U) = \min\{d_k : v_k \in U\}$  e  $d_{\max}(W) = \max\{d_w : v_w \in W\}$ . Então, temos que

$$d_{\max}(W) \leq \rho \leq d_{\min}(U), \quad (4.4)$$

com os limites superior e inferior como igualdades se e somente se  $G$  é um grafo regular.

**Prova:** Primeiro, faremos a prova do lado direito das desigualdades. Seja  $v_k \in U$  tal que

$d_k = d_{\min}(U)$ . Dado que  $Ax = \rho x$ , temos

$$\rho x_k = \sum_{\{v_i, v_k\} \in E} x_i, \quad (4.5)$$

onde o somatório se estende a todos os vértices  $v_i$  adjacentes a  $v_k$ . Sendo  $x_k = \max_{v_i \in V} x_i$ ,

da equação (4.5), obtemos

$$\rho = \frac{1}{x_k} \sum_{\{v_i, v_k\} \in E} x_i \leq \frac{1}{x_k} \sum_{\{v_i, v_k\} \in E} x_k = d_k = d_{\min}(U). \quad (4.6)$$

Se  $G$  é um grafo regular, todas as coordenadas de um autovetor  $x$  associado a  $\rho$  são iguais, então a desigualdade (4.6) é verdadeira na igualdade.

Por outro lado, se a igualdade acontece em (4.6), temos

$$\sum_{\{v_i, v_k\} \in E} x_i = \sum_{\{v_i, v_k\} \in E} x_k. \quad (4.7)$$

Sendo  $0 < x_i \leq x_k$ , quando (4.6) define uma igualdade, temos  $x_i = x_k$ , para todo vértice  $v_i$  adjacente a  $v_k$ . O mesmo procedimento pode ser aplicado para qualquer vértice  $v_w$  adjacente a  $v_k$ . Assim,  $x_j = x_w$ , para todo vértice  $v_j$  adjacente a  $v_w$ . Como  $G$  é conexo, podemos prosseguir sucessivamente desta maneira até concluirmos que  $x_i = x_k$ ,

para todo  $i = 1, \dots, n$ . Portanto todas as coordenadas do autovetor  $x$  são iguais, e isto só acontece quando  $G$  é grafo regular [27].

O mesmo argumento é aplicado do lado esquerdo da desigualdade, concluindo assim a prova.  $\square$

**Exemplo:**

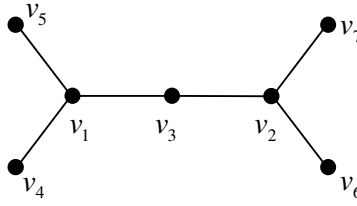


Figura 4.5: Grafo  $G$  com  $\rho = 2$ .

O grafo  $G$  da Figura 4.5 possui  $\delta = 1$ ,  $\Delta = 3$  e raio espectral  $\rho = 2$  que satisfaz os limites clássicos  $\delta \leq \rho \leq \Delta$ . Por outro lado, associado ao raio espectral  $\rho = 2$ , temos o autovetor  $x = [0,5 \ 0,5 \ 0,5 \ 0,25 \ 0,25 \ 0,25 \ 0,25]^t$ , de onde obtemos os conjuntos  $U = \{v_1, v_2, v_3\}$ ,  $W = \{v_4, v_5, v_6, v_7\}$  e os limites  $d_{\max}(W) = 1$  e  $d_{\min}(U) = 2$ , que pelo Teorema 4.3, satisfazem as desigualdades  $d_{\max}(W) \leq \rho \leq d_{\min}(U)$ . Observe que o Teorema 4.3 realmente melhora o resultado clássico  $\delta \leq \rho \leq \Delta$ , dado que  $d_{\min}(U)$  e  $d_{\max}(W)$  são graus mínimo e máximo apenas dos subconjuntos de vértices com respectivas coordenadas máxima e mínima no autovetor associado ao índice.

O próximo exemplo novamente mostra que diferentes medidas de centralidade podem não resultar num mesmo conjunto de vértices mais central. Ou seja, medidas distintas de centralidade podem definir diferentes subconjuntos de um ou mais vértices mais centrais. Para ilustrar isto, a Figura 4.6 mostra uma árvore onde o vértice mais

central, segundo a centralidade de grau, difere daquela obtida pela centralidade de autovetor.

**Exemplo:**

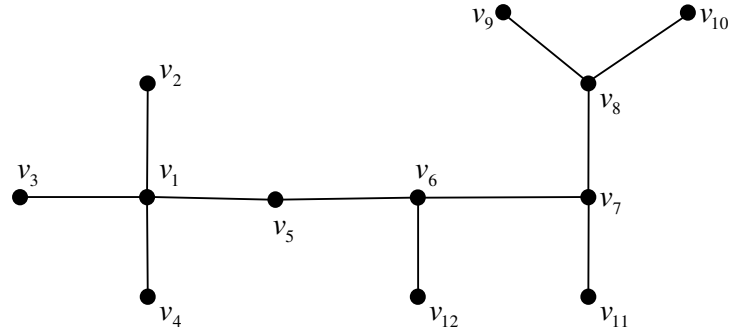


Figura 4.6: Árvore  $T$  com vértices mais centrais distintos para medidas de centralidade distintas.

Veja que a árvore  $T$  da Figura 4.6 tem como raio espectral  $\rho = 2,248$ , cujo o autovetor é

$$x = [0,42 \quad 0,186 \quad 0,186 \quad 0,186 \quad 0,383 \quad 0,442 \quad 0,414 \quad 0,305 \quad 0,135 \quad 0,135 \quad 0,184 \quad 0,197]^t.$$

Como  $x_6$  é a maior coordenada de  $x$ ,  $v_6$  é o vértice de maior centralidade de autovetor. No entanto,  $v_1$  é o vértice que possui maior centralidade de grau.

O Teorema 4.4, cuja prova pode ser encontrada em [27], atribui uma condição necessária para que o vértice com maior centralidade de autovetor possua, também, a maior centralidade de grau. Seja  $\varsigma$  o segundo maior grau de um vértice em um grafo. Sob a hipótese que  $\rho > \varsigma$ , obteremos grafos com a propriedade que desejamos.

**Teorema 4.4.** Seja  $G = (V, E)$  um grafo conexo e  $A$  sua matriz de adjacência cujo raio espectral  $\rho$  tem autovetor correspondente  $x$ . Seja  $v_k \in G$ ,  $d_k$  seu respectivo grau e

$x_k$  a  $k$ -ésima coordenada do autovetor  $x$  tal que  $x_k = \max_{v_i \in V} x_i$ . Seja  $\varsigma$  o segundo maior

grau de um vértice  $v_i \in G$ . Se  $\rho > \varsigma$  então  $d_k = \Delta$ .

**Exemplo:**

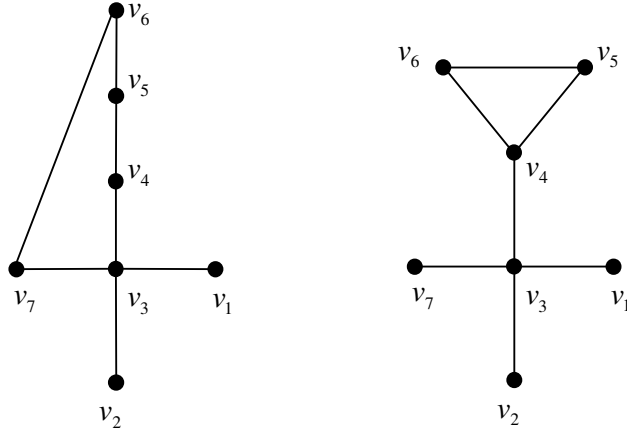


Figura 4.7: O grafo  $G_1$  mais à esquerda satisfaz às condições do Teorema 4.4 e  $G_2$ , o grafo mais à direita não satisfaz.

O grafo  $G_1$  da Figura 4.7 ilustra bem o Teorema 4.4. No grafo  $G_1$ , temos  $v_3$  como o vértice de grau máximo  $\Delta = 4$ ,  $v_4$ ,  $v_5$ ,  $v_6$  e  $v_7$  vértices com o segundo maior grau  $\varsigma = 2$ . O raio espectral de  $G_1$  é  $\rho(G_1) = 2,256$ , cujo autovetor  $x$  associado é  $x = [0,259 \ 0,259 \ 0,584 \ 0,400 \ 0,318 \ 0,318 \ 0,400]^t$ . Portanto, dado que  $x_3$  é a coordenada de maior valor em  $x$ , o vértice  $v_3$  é o que possui maior centralidade de autovetor. Veja que é também aquele que possui maior grau, pois  $\rho(G_1) > \varsigma$ . Por outro lado, o grafo  $G_2$  não satisfaz o Teorema 4.4. Nele,  $v_3$  é o vértice de grau máximo  $\Delta = 4$ , e  $v_4$  é aquele com o segundo maior grau  $\varsigma = 3$  e veja que o raio espectral é  $\rho(G_2) = 2,368$ , que é inferior a 3, ou seja,  $\rho(G_2) < \varsigma$ . O autovetor associado a  $\rho(G_2)$  é dado por  $x = [0,209 \ 0,209 \ 0,497 \ 0,547 \ 0,400 \ 0,400 \ 0,209]^t$ . Note que, neste

caso, o vértice de maior centralidade de autovetor é o  $v_4$ , enquanto que aquele com maior centralidade de grau é o  $v_3$ .

O Teorema 4.5, cuja prova pode ser encontrada em [27], relaciona as coordenadas do autovetor a vértices adjacentes em um grafo.

**Teorema 4.5.** Seja  $G=(V,E)$  um grafo conexo e  $A$  sua matriz de adjacência, cujo raio espectral é  $\rho$  e o autovetor correspondente é  $x$ . Sejam  $v_k$  e  $v_j \in V$ ,  $x_k$  a  $k$ -ésima coordenada de  $x$ , tal que  $x_k > x_j \geq x_i$ , para todo  $i=1,\dots,n$ ,  $i \neq k, j$ . Seja  $\varsigma$  o segundo maior grau de um vértice  $v_i \in V$ . Se  $\rho > \varsigma$  então  $v_k$  e  $v_j$  são vértices adjacentes em  $G$ .

**Exemplo:**

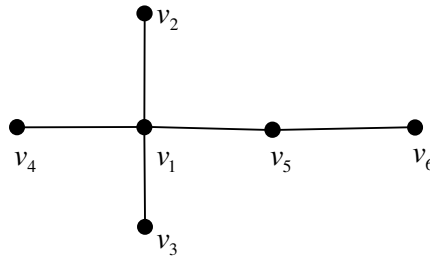


Figura 4.8: Os vértices  $v_1$  e  $v_5$  são adjacentes por satisfazerem as condições do Teorema 4.5.

O grafo  $G$  da Figura 4.8 possui raio espectral  $\rho = 2,074$  e autovetor associado  $x = [0,676 \ 0,326 \ 0,326 \ 0,326 \ 0,425 \ 0,204]^T$ . Note que  $x_1 > x_5 > x_i$ ,  $i = 2,3,4,6$  e também que  $\rho > \varsigma = 2$ . Do Teorema 4.5 podemos, antes mesmo de conhecer o grafo, afirmar que os vértices  $v_1$  e  $v_5$  são adjacentes. Isto é, ainda sob a condição  $\rho > \varsigma$ , o



vértice com maior centralidade de autovetor tem que estar diretamente ligado ao segundo mais central, em relação à mesma medida.

Para medir a centralidade de um vértice, em função dos valores obtidos para as coordenadas do autovetor associado ao índice da matriz de adjacência, vimos que, assim como na centralidade de proximidade, exige-se que o grafo seja conexo. A próxima subseção deste capítulo discute como tratar estes problemas em grafos desconexos.

### 4.3 Centralidade de autovetor em grafos desconexos

Da Proposição 4.1, cuja prova pode ser encontrada em [1], vemos que o espectro de um grafo desconexo pode facilmente ser encontrado através dos espectros de suas componentes conexas.

**Proposição 4.1** Se o grafo  $G$  é a união de dois grafos  $G_1$  e  $G_2$  então  $p_G(\lambda) = p_{G_1}(\lambda) \cdot p_{G_2}(\lambda)$ . Se  $G_1, G_2, \dots, G_r$  são componentes conexas de um grafo  $G$  então  $p_G(\lambda) = p_{G_1}(\lambda) \cdot p_{G_2}(\lambda) \cdot \dots \cdot p_{G_r}(\lambda)$ .

Como consequência da Proposição 4.1, temos que se  $G = G_1 \cup \dots \cup G_r$  então  $spect(G) = spect(G_1) \cup \dots \cup spect(G_r)$ .

Vejamos como tratar a centralidade de autovetor em grafos desconexos, utilizando-se a Proposição 4.1, através dos dois casos abaixo:

1) Seja  $H$  um grafo conexo com  $q$  vértices e  $G = H \cup \{v_1, \dots, v_p\}$  o grafo obtido de  $H$  acrescentando a ele  $p$  vértices isolados  $v_1, \dots, v_p$ . A matriz de adjacência de  $G$  é dada por

$$A(G) = \begin{bmatrix} A(H)_{q \times q} & 0_{q \times p} \\ 0_{p \times q} & 0_{p \times p} \end{bmatrix}.$$

Se  $\rho(H)$  é índice de  $H$  e  $u$  um autovetor de  $H$  associado a  $\rho(H)$ , então  $\rho(H)$  é o índice de  $G$  e  $\begin{bmatrix} u \\ 0 \end{bmatrix}$  com  $p$  coordenadas nulas, é um autovetor de  $G$  associado a  $\rho(H)$ , isto é,

$$A(G) \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} A(H) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} A(H)u \\ 0 \end{bmatrix} = \begin{bmatrix} \rho(H)u \\ 0 \end{bmatrix} = \rho(H) \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

O autovetor  $\begin{bmatrix} u \\ 0 \end{bmatrix}$  associado ao índice de  $G$  indica a importância dos vértices da componente conexa  $H$  e atribui valor zero a todos os demais vértices isolados. Assim, as multiplicidades algébrica e geométrica de  $\rho(H)$  pela Proposição 4.1, continua sendo igual a 1. Neste caso, os vértices mais centrais de  $G$  coincidirão com aquelas coordenadas de maior valor na única componente conexa de ordem maior que 1.

**Exemplo:**

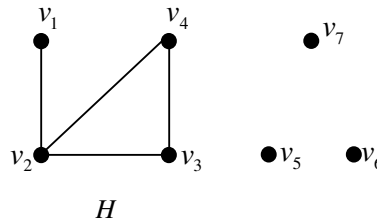


Figura 4.9: Grafo  $G$  desconexo com uma componentes conexa  $H$  e três vértices

isolados,  $v_5$ ,  $v_6$  e  $v_7$ .

O grafo  $G$  da Figura 4.9 possui raio espectral  $\rho(G) = \rho(H) = 2,17$ , e o autovetor associado é  $x = [0,281 \ 0,611 \ 0,522 \ 0,522 \ 0 \ 0 \ 0]^t$ . Observe que o autovetor  $x$  associado ao índice de  $G$  indica a importância dos vértices da componente conexa  $H$  e atribui valor zero a todos os demais vértices isolados.

2) Vamos considerar  $G$  composto por duas componentes conexas  $H_1$  e  $H_2$ , o primeiro com  $q$  vértices e índice  $\rho(H_1)$  e o segundo com  $p$  vértices e índice  $\rho(H_2)$ , tal que  $n = p + q$ . Seja  $\rho(H_1) > \rho(H_2)$ . Logo, o índice do grafo  $G = H_1 \cup H_2$  é  $\rho(H_1)$  com multiplicidade 1. O autovetor associado a  $\rho(H_1)$  é dado por  $\begin{bmatrix} u \\ 0 \end{bmatrix}$  com  $p$  coordenadas nulas, isto é,

$$A(G) \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} A(H_1) & 0 \\ 0 & A(H_2) \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} A(H_1)u \\ 0 \end{bmatrix} = \begin{bmatrix} \rho(H_1)u \\ 0 \end{bmatrix} = \rho(H_1) \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

Por esta razão, a centralidade dos vértices fornecida pelo autovetor  $\begin{bmatrix} u \\ 0 \end{bmatrix}$  só é capaz de indicar a importância dos vértices de  $H_1$ , pois todos os vértices da componente conexa  $H_2$  terão o mesmo valor, independentemente das ligações entre eles. Por exemplo, em  $H_2$  haverá um vértice central com coordenada máxima para  $\rho(H_2)$ . Este vértice tem alta influência em sua componente conexa  $H_2$  mas não podemos identificar sua importância na rede toda.

**Exemplo:**

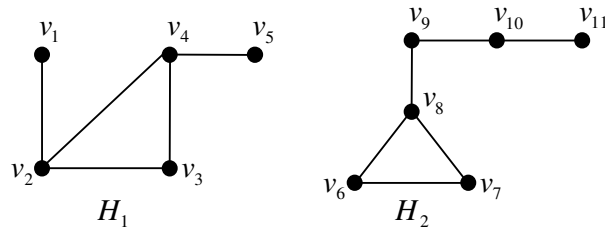


Figura 4.10: Grafo  $G$  desconexo com duas componentes conexas  $H_1$  e  $H_2$ .

O grafo  $G$  da Figura 4.10 ilustra bem o caso 2. Nele, temos duas componentes conexas  $H_1$  e  $H_2$ , cada qual com raio espectral igual a  $\rho(H_1) = 2,302$  e

$\rho(H_2) = 2,228$ . Como  $\rho(H_1) > \rho(H_2)$ , temos que o índice do grafo  $G = H_1 \cup H_2$  é  $\rho(H_1) = 2,302$  com multiplicidade 1. O autovetor associado é:

$$x = [0,245 \quad 0,565 \quad 0,490 \quad 0,565 \quad 0,245 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]^t.$$

Observe que o autovetor  $x$  só é capaz de indicar a importância dos vértices de  $G$  com base nos vértices da componente conexa  $H_1$ , já que os vértices referentes à componente conexa  $H_2$  possuem centralidade de autovetor nula.

#### 4.4 Centralidade de um vértice via conectividade algébrica

A conectividade algébrica de um grafo, apresentada na Seção 2.2, é um parâmetro espectral usado em análise de vulnerabilidade de redes. O objetivo da análise de vulnerabilidade é reduzir o risco em relação aos incidentes de segurança, detectando possíveis falhas que possam vir a comprometer a rede, [52]. Face a este conceito, podemos medir a importância de um vértice em relação a vulnerabilidade que ele oferece à rede caso tenha de ser dela retirado.

O Teorema 4.6, cuja prova pode ser encontrada em [53], fornece uma relação entre a conectividade algébrica de um grafo  $G$ , denotada por  $\alpha(G)$ , e a conectividade algébrica do subgrafo induzido de  $G$  obtido pela retirada de  $v_k$ , denotada por  $\alpha(G \setminus v_k)$ .

**Teorema 4.6.** Seja  $G = (V, E)$  um grafo conexo e  $v_k$  um vértice de  $G$ . Seja  $G \setminus v_k$  o subgrafo induzido de  $G$  obtido pela retirada de  $v_k$ . Então,  $\alpha(G) \leq \alpha(G \setminus v_k) + 1$ .

Se  $G \setminus v_k$  for desconexo, chamaremos  $v_k$  de *ponto de articulação*.

Observe que  $\alpha(G \setminus v_k)$  pode ser maior que, igual a, ou menor que  $\alpha(G)$ , ou seja, dependendo da escolha de  $v_k$  a ser retirado do grafo, a conectividade algébrica de  $G$  pode aumentar, se manter, ou reduzir. Isto serviu como ponto de partida para que KIRKLAND [30] introduzisse uma medida de centralidade de vértice baseada na conectividade algébrica de um grafo.

**Definição 4.2.** Seja  $G$  um grafo conexo e  $v_k$  um vértice de  $G$ . Seja  $G \setminus v_k$  o subgrafo induzido de  $G$  obtido pela retirada de  $v_k$ . A *centralidade do vértice  $v_k$  via conectividade algébrica* é definida por:

$$c_a(v_k) = \alpha(G) - \alpha(G \setminus v_k). \quad (4.8)$$

Note que da equação (4.8), a centralidade via conectividade algébrica pode, diferentemente das demais medidas de centralidade até aqui apresentadas, assumir valores negativos. No caso em que  $c_a(v_k) > 0$ , podemos tomar como interpretação que o vértice  $v_k$  e suas arestas incidentes servem para aumentar a conectividade algébrica de  $G$ . Se  $c_a(v_k) < 0$ , então  $v_k$  e suas arestas incidentes diminuem a conectividade algébrica de  $G$ , enquanto que se  $c_a(v_k) = 0$ , então  $v_k$  e suas arestas incidentes não têm nenhum efeito na conectividade algébrica de  $G$ . Assim, temos que  $c_a(v_k)$  pode ser usada para fornecer uma medida da contribuição de um vértice (e suas arestas incidentes) para  $\alpha(G)$ .

**Exemplo:** Considere o grafo  $G$  da Figura 4.11

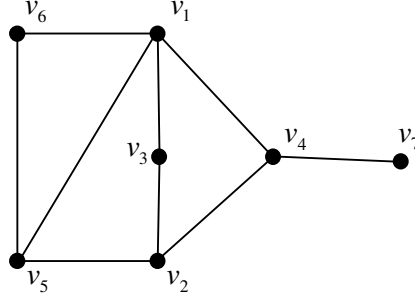


Figura 4.11: Grafo  $G$  com ponto de articulação  $v_4$ .

As matrizes  $D(G)$ , diagonal dos vértices e  $A(G)$  de adjacência, responsáveis por determinar a matriz laplaciana de  $G$ ,  $L(G)$ , são apresentadas abaixo:

$$D(G) = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad A(G) = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad L(G) = \begin{bmatrix} 4 & 0 & -1 & -1 & -1 & -1 & 0 \\ 0 & 3 & -1 & -1 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 3 & 0 & 0 & -1 \\ -1 & -1 & 0 & 0 & 3 & -1 & 0 \\ -1 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

Para o grafo  $G$  considerado, o polinômio característico da matriz Laplaciana de  $G$  é  $p_{L(G)}(\mu) = \mu^7 - 18\mu^6 + 127\mu^5 - 446\mu^4 + 811\mu^3 - 708\mu^2 + 224\mu$ , cujas raízes formam o espectro  $\zeta(G) = (5,72, 4, 3,345, 2,681, 1,547, 0,705, 0)$ . A conectividade algébrica de  $G$  é então  $\alpha(G) = 0,705$ . Por exemplo, se retirarmos o vértice  $v_1$  de  $G$ , obtemos o subgrafo induzido  $G \setminus v_1$  cujo espectro é  $\zeta(G \setminus v_1) = (4,302, 2,618, 2, 0,697, 0,381, 0)$ . A conectividade algébrica de  $G \setminus v_1$  é dada por  $\alpha(G \setminus v_1) = 0,381$ . Assim,  $c_a(v_1) = \alpha(G) - \alpha(G \setminus v_1) = 0,705 - 0,381 = 0,324$  é a centralidade de  $v_1$  via conectividade algébrica. Analogamente, podemos encontrar a centralidade dos demais vértices,  $c_a(v_2) = 0,221$ ,  $c_a(v_3) = -0,015$ ,  $c_a(v_4) = 0,705$ ,  $c_a(v_5) = 0,049$ ,  $c_a(v_6) = -0,057$  e  $c_a(v_7) = -0,732$ .

Como podemos observar, o ponto de articulação representado pelo vértice  $v_4$  é indispensável para manter o grafo conexo, e considerado o mais central, segundo a medida  $c_a$ . Portanto, a centralidade via conectividade algébrica pode ser considerada como uma medida capaz de exibir uma sequência de vértices que devem exigir maior vigilância por parte dos administradores de uma dada rede para que o grafo que a modela não se rompa.

A Tabela 4.2 mostra um comparativo entre a centralidade via conectividade algébrica e algumas das principais medidas de centralidade já abordadas no texto:

<i>Ranking</i>	Grau	Intermediação	Proximidade	Autovetor	Conectividade Algébrica
1	$v_1(4)$	$v_1(5,5)$	$v_1(0,125)$	$v_1(0,52)$	$v_4(0,705)$
2	$v_4(3)$	$v_4(5,3)$	$v_4(0,11)$	$v_5(0,44)$	$v_1(0,324)$
3	$v_2(3)$	$v_2(2,5)$	$v_2(0,11)$	$v_2(0,4)$	$v_2(0,221)$
4	$v_5(3)$	$v_5(1,3)$	$v_5(0,1)$	$v_4(0,37)$	$v_5(0,049)$
5	$v_3(2)$	$v_3(0,3)$	$v_3(0,09)$	$v_6(0,33)$	$v_3(-0,015)$
6	$v_6(2)$	$v_6(0)$	$v_6(0,09)$	$v_3(0,32)$	$v_6(-0,057)$
7	$v_7(1)$	$v_7(0)$	$v_7(0,07)$	$v_7(0,13)$	$v_7(-0,732)$

Tabela 4.2: *Ranking* dos vértices para o grafo  $G$  da Figura 4.11. As centralidades estão entre parênteses.

Observe que apenas a centralidade via conectividade algébrica atribui maior relevância à  $v_4$ , pois sua retirada desconecta o grafo, fato que não é contemplado pelas demais medidas de centralidade.

## Capítulo 5 – Centralidade em Grafos Valorados

Há certas redes que quando modeladas por grafos, faz-se necessário avaliar não somente as possíveis ligações entre os pares de vértices como também a intensidade de tais ligações. Assim, precisamos atribuir valores às arestas dos grafos e estes serão considerados grafos valorados. Neste caso, os grafos são valorados por arestas. Há outros casos em que a valoração se dá sobre os vértices, mas estes não serão aqui estudados.

Um grafo é *valorado* se existe uma função  $w: V \times V \rightarrow \mathbb{R}$  que relaciona  $E(G)$  a valores numéricos, referindo-se cada número como o *peso* da aresta, denotado por  $\omega$ . Estes valores podem ser custos, distâncias, tempo gasto no percurso, confiabilidade da transmissão, etc. As próximas seções tratam dos fundamentos das medidas de centralidade de um vértice em grafos valorados.

### 5.1 Medidas de centralidade em grafos valorados

As medidas de centralidade de grau e centralidade de proximidade de um vértice podem ser estendidas para grafos valorados. Assim, a **centralidade de grau** de  $v_k$  é dada pela soma dos pesos das arestas incidentes a  $v_k$ , e a **centralidade de proximidade** de  $v_k$  é dada pelo inverso da soma dos pesos das arestas referentes a geodésica que liga pares de vértices. Outras duas medidas de grande relevância, tais como *centralidade de autovetor* e *centralidade de intermediação de fluxo*, serão mais detalhadamente apresentadas.



## 5.2 Centralidade de autovetor em grafos valorados

A centralidade de autovetor em grafos valorados considera uma matriz onde as entradas correspondem aos valores de cada aresta, denominada matriz dos pesos. Podemos defini-la da seguinte maneira:

**Definição 5.1.** Seja  $G$  um grafo conexo valorado com  $n$  vértices. A *matriz dos pesos*  $P(G)$  de  $G$  é uma matriz de ordem  $n$  cujas entradas são:

$$p_{ij} = \begin{cases} \omega_{ij}, & \text{se } \{v_i, v_j\} \in E; \\ 0, & \text{caso contrário.} \end{cases} \quad (5.1)$$

onde  $\omega_{ij}$  é dado pela função  $w: V \times V \rightarrow \mathbb{R}$ .

**Observação:** Consideraremos apenas grafos valorados com  $p_{ij} \geq 0$ , para todo  $i, j$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq n$ . A matriz  $P = P(G)$  será, portanto, simétrica e não-negativa.

Se  $G$  é um grafo conexo, a matriz dos pesos é irredutível, isto significa que as hipóteses do Teorema de Perron-Frobenius [18] são satisfeitas e podemos então definir a centralidade de autovetor para estes tipos de grafos.

**Definição 5.2.** Seja  $G$  um grafo conexo valorado com  $n$  vértices e seja  $v_k$  um vértice de  $G$ . A *centralidade de autovetor* de  $v_k$  é dada por:

$$c_{\text{eig}}(v_k) = x_k, \quad (5.2)$$

onde  $x_k$  é a  $k$ -ésima coordenada do autovetor positivo unitário  $x$  associado ao índice do grafo, isto é

$$x_k = \frac{1}{\rho} \sum_{j=1}^n \omega_{ij} x_j ; k = 1, \dots, n. \quad (5.3)$$

**Exemplo:**

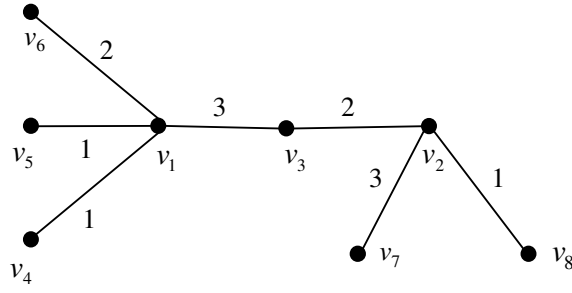


Figura 5.1: Grafo valorado  $G$  cujo vértice  $v_3$  é o mais central, segundo a centralidade de autovetor.

Consideremos o grafo  $G$  da Figura 5.1. Sua respectiva matriz dos pesos é dada abaixo:

$$P(G) = \begin{bmatrix} 0 & 0 & 3 & 1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 3 & 1 \\ 3 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

O polinômio característico de  $G$  e seu respectivo espectro são:

$$p_G(\lambda) = \lambda^8 - 29\lambda^6 + 174\lambda^4 \quad \text{e}$$

$$\text{spect } G = \begin{bmatrix} 4,529 & 2,911 & 0 & 0 & 0 & 0 & -2,911 & -4,529 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

O autovetor positivo e de norma 1 associado a  $\rho = 4,529$  é:

$$x = [0,520 \quad 0,478 \quad 0,555 \quad 0,114 \quad 0,114 \quad 0,114 \quad 0,229 \quad 0,317]^t.$$

Assim, as centralidades de autovetor dos respectivos vértices de  $G$  são:

$$c_{eig}(v_1) = 0,520; c_{eig}(v_2) = 0,478; c_{eig}(v_3) = 0,555;$$

$$c_{eig}(v_4) = c_{eig}(v_5) = c_{eig}(v_6) = 0,114; c_{eig}(v_7) = 0,229; c_{eig}(v_8) = 0,317.$$

Portanto, a medida indica o vértice  $v_3$  como mais central do grafo da Figura 5.1, segundo a centralidade de autovetor.

Observe que, se todas as arestas do grafo da Figura 5.1 tivessem peso 1, teríamos a matriz de adjacência associada ao grafo. Neste caso,  $\rho = 2,148$  e o autovetor positivo e de norma 1 correspondente é:

$$x = [0,601 \quad 0,371 \quad 0,452 \quad 0,279 \quad 0,279 \quad 0,279 \quad 0,172 \quad 0,172]^t,$$

o que indica o vértice  $v_1$  como mais central do grafo.

Desta forma, podemos concluir que em certas situações, pode não ser o bastante medir apenas a importância de um vértice em relação a seus vizinhos, mas também avaliar a intensidade destas relações, que pode melhor ser representada pela matriz dos pesos referente ao grafo valorado.

### 5.3 Centralidade de Intermediação de Fluxo

Um *fluxo* em um grafo  $G$  é o movimento de algum recurso quantificável ao longo das arestas de  $G$  e indicado por uma valoração que representa o quanto está sendo transferido. As arestas da rede são consideradas como os canais de uma comunicação ligando pares de vértices; o valor da conexão entre  $v_i$  e  $v_j$  determina a capacidade  $w_{ij}$  do canal que os liga, ou a quantidade máxima de informação que pode passar entre eles. A informação flui ao longo destes canais. Se  $f_{ij}$  é a quantidade de informação que passa em um canal que liga  $v_i$  diretamente a  $v_j$ , então  $f_{ij} \leq w_{ij}$ , isto é, a quantidade de informação que flui ao longo da aresta de um canal que conecta diretamente dois

vértices não pode exceder a capacidade desse canal. O que é relevante não é apenas o fluxo direto entre vértices conectados, mas o fluxo total entre pares de vértices ao longo de todos os possíveis caminhos que os conectam. Se um vértice  $v_s$  é escolhido como fonte ou transmissor da informação, e um outro vértice  $v_t$ , como um alvo ou receptor da informação, esta informação que sai de  $v_s$  pode alcançar  $v_t$  ao longo de uma aresta que liga  $v_s$  diretamente a  $v_t$  ou ao longo de qualquer caminho indireto que começa em  $v_s$ , passa por um ou vários vértices intermediários e termina em  $v_t$ . Assim, o fluxo entre dois vértices é um fenômeno global. Ele depende, não apenas da capacidade do canal ligando dois vértices diretamente, mas das capacidades de todos os canais e todos os possíveis caminhos - direto e indireto - que os conecta.

Seja  $G = (V, E, f)$  um grafo com fluxo  $f$  e seja  $S, T \subset V$  uma partição de  $V$  tal que  $v_s \in S$  e  $v_t \in T$ . Um *corte*  $v_s - v_t$ , denotado por  $C(S, T)$ , é um subconjunto de arestas de  $G$  tal que todo caminho de  $v_s$  para  $v_t$  contém uma aresta em  $C(S, T)$ . A *capacidade do corte*, denotado por  $cap(C(S, T))$ , é o somatório dos valores das arestas que pertencem ao corte. O *fluxo máximo* é igual ao valor mínimo da capacidade do corte dentre todos os cortes possíveis do grafo (Ford e Fulkerson, 1956). A quantidade de fluxo máximo entre dois vértices  $v_i$  e  $v_j$  é denotado por  $\theta_{ij}$ . Seja  $v_k \in V$ , dizemos que  $v_k$  pertence ao corte que corresponde ao fluxo máximo de  $v_i$  a  $v_j$ , se o mesmo depende de  $v_k$  para ocorrer. Denotamos por  $\theta_{ij}(v_k)$  a quantidade de fluxo máximo de  $v_i$  a  $v_j$  que passa por  $v_k$ .

A medida proposta por FREEMAN [24], em 1991, estende o conceito de centralidade de intermediação de um vértice para grafos valorados, baseada na idéia do

fluxo máximo, de Ford e Fulkerson (1956). Esta medida é chamada de *centralidade de intermediação de fluxo* e podemos defini-la da seguinte maneira:

**Definição 5.3.** Seja  $G = (V, E, f)$  um grafo com  $n$  vértices e fluxo  $f$ . Seja  $v_k$  um vértice de  $G$ . A *centralidade de intermediação de fluxo* através de  $v_k$  é dada por:

$$c_F(v_k) = \sum_{\substack{i < j \\ i, j \neq k}}^n \theta_{ij}(v_k), \quad (5.4)$$

onde  $\theta_{ij}(v_k)$  denota a quantidade de fluxo máximo de  $v_i$  a  $v_j$  que passa por  $v_k$ .

Se nós dividirmos a quantidade de fluxo que passa através de  $v_k$  pelo fluxo máximo total entre todos os pares de vértices, onde  $v_k$  não é fonte nem alvo, nós podemos determinar a proporção do fluxo que depende de  $v_k$ .

**Definição 5.4.** Seja  $G = (V, E, f)$  um grafo com  $n$  vértices e fluxo  $f$ . Seja  $v_k$  um vértice de  $G$ . A *centralidade de intermediação relativa de fluxo* através de  $v_k$  é dada por:

$$c'_F(v_k) = \frac{\sum_{\substack{i < j \\ i, j \neq k}}^n \theta_{ij}(v_k)}{\sum_{\substack{i < j \\ i, j \neq k}}^n \theta_{ij}}, \quad (5.5)$$

onde  $\theta_{ij}$  denota a quantidade de fluxo máximo de  $v_i$  a  $v_j$  e  $\theta_{ij}(v_k)$  denota a quantidade de fluxo máximo de  $v_i$  a  $v_j$  que passa por  $v_k$ .

**Observação:** O fluxo máximo de uma dada fonte a todos os alvos alcançáveis pode ser calculado com complexidade de  $O(m^2)$ , e portanto, a medida de centralidade de intermediação de fluxo pode ser calculada para todos os vértices com complexidade de  $O(m^2n)$ , [33].

**Exemplo:**

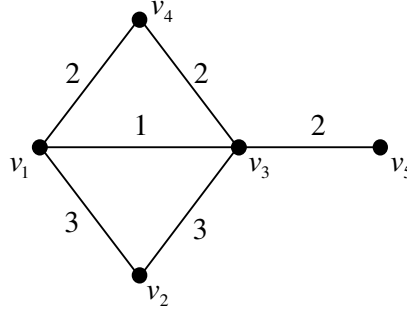


Figura 5.2: Grafo  $G$  com  $v_3$  sendo o vértice mais central, segundo a centralidade de intermediação de fluxo.

Na Figura 5.2, todos os fluxos máximos entre  $v_i$  e  $v_j$ ,  $v_i, v_j \in V$ ,  $i, j \neq 2$ , são dados por  $\theta_{13} = 6$ ,  $\theta_{14} = 4$ ,  $\theta_{15} = 2$ ,  $\theta_{34} = 4$ ,  $\theta_{35} = 2$ ,  $\theta_{45} = 2$  e as quantidades que passam necessariamente por  $v_2$  de modo que o fluxo seja máximo são  $\theta_{13}(v_2) = 3$ ,  $\theta_{14}(v_2) = 1$ ,  $\theta_{34}(v_2) = 1$ . As demais quantidades são nulas, isto é,  $\theta_{15}(v_2) = \theta_{35}(v_2) = \theta_{45}(v_2) = 0$ .

Assim, a medida de centralidade de intermediação de fluxo através de  $v_2$  é:

$$c_F(v_2) = \sum_{\substack{1 \leq i < j \leq 5 \\ i, j \neq 2}} \theta_{ij}(v_2) = 3 + 1 + 0 + 1 + 0 + 0 = 5.$$

Pelo mesmo raciocínio, teremos para os demais vértices:

$$c_F(v_1) = 7; c_F(v_3) = 13; c_F(v_4) = 6 \text{ e } c_F(v_5) = 0.$$

Portanto, a centralidade de intermediação de fluxo indica o vértice  $v_3$  como sendo o mais central.

Por outro lado, a medida relativa da intermediação de fluxo através dos vértices dada pela equação (5.5) para o grafo da Figura 5.2 nos fornece:

$$c'_F(v_1) = 0,35; c'_F(v_2) = 0,25; c'_F(v_3) = 0,65; c'_F(v_4) = 0,25 \text{ e } c'_F(v_5) = 0.$$

Note que a centralidade de intermediação de fluxo difere da clássica centralidade de intermediação de um vértice por considerar caminhos geodésicos e não-geodésicos, e também por sua aplicabilidade em grafos valorados, [24].

## Capítulo 6 – Aplicações e Resultados

Este capítulo se destina a análise e comparação das medidas de centralidade até aqui apresentadas. Inicialmente, vamos introduzir o conceito de isomorfismo em grafos e vértices automorficamente equivalentes, que também pode ser encontrado em [28]. Além disso, mostraremos que vértices universais, quando eles existem em determinados grafos não valorados, são mais centrais que os demais vértices em relação a qualquer uma das cinco medidas de centralidade estudadas (centralidade de grau, centralidade de proximidade, centralidade de eficiência, centralidade de intermediação e centralidade de autovetor).

### 6.1 Vértices automorficamente equivalentes e vértices universais

**Definição 6.1.** Dois grafos  $G_1 = (V_1, E_1)$  e  $G_2 = (V_2, E_2)$  são ditos *isomorfos* se existir uma função bijetiva  $\phi: V_1 \rightarrow V_2$  que preserve as adjacências entre os vértices, isto é,  $\{v_i, v_j\} \in E_1$  se e somente se  $\{\phi(v_i), \phi(v_j)\} \in E_2$ . A bijeção  $\phi$  é um isomorfismo entre  $G_1$  e  $G_2$  e usamos a notação  $G_1 \approx G_2$  para dizer que  $G_1$  e  $G_2$  são isomorfos.

**Exemplo:**

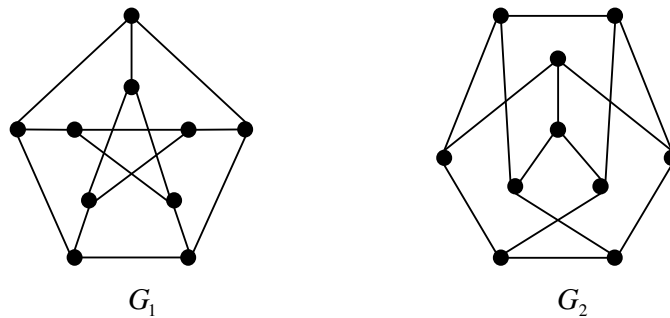


Figura 6.1: Grafos isomorfos.



Um automorfismo sobre um grafo  $G$  é um isomorfismo  $\phi: V \rightarrow V$ , ou seja, é uma função bijetiva entre os vértices de um mesmo conjunto.

**Definição 6.2.** Dois vértices  $v_i$  e  $v_j \in V$  são chamados *automorficamente equivalentes* em  $G$  se  $v_j = \phi(v_i)$  para algum automorfismo  $\phi$  de  $G$ .

**Exemplo:**

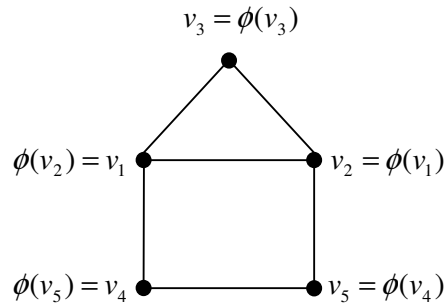


Figura 6.2: Grafo  $G$  cujos vértices  $v_1$  e  $v_2$  são automorficamente equivalentes.

Também são automorficamente equivalentes os vértices  $v_4$  e  $v_5$ .

Da Definição 6.2, podemos observar que nos grafos completos  $K_n$  e ciclos  $C_n$ , que são tipos de grafos regulares, todos os vértices são automorficamente equivalentes entre si e, portanto, para qualquer medida de centralidade  $x$ , é válida  $c_x(v_i) = c_x(v_j)$ , para todo  $v_i, v_j \in V$ , [28]. No entanto, nem todo grafo  $k$ -regular possui todos os vértices com um mesmo valor para medida de centralidade  $x$ , independentemente da escolha de  $x$ .

**Exemplo:**

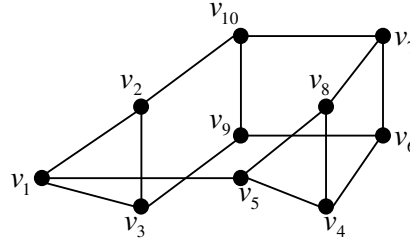


Figura 6.3: Grafo  $G$  3-regular com diferentes centralidades de intermediação.

Na Figura 6.3, todos os vértices do grafo  $G$  têm grau 3 mas  $c_B(v_1) = c_B(v_5) = 6$ ,  $c_B(v_2) = c_B(v_3) = c_B(v_4) = c_B(v_8) = 2,83...$  e  $c_B(v_6) = c_B(v_7) = c_B(v_9) = c_B(v_{10}) = 4,66...$

O Teorema 6.1, cuja prova parcial pode ser encontrada em [28], garante que em todo grafo que possui um ou mais vértices universais, isto é, vértices adjacentes a todos os demais, estes vértices são sempre mais centrais que quaisquer outros em relação a algumas das medidas de centralidade apresentadas ao longo do texto. Além disso, para cada uma dessas medidas, todos eles possuem a mesma centralidade.

**Teorema 6.1.** Seja  $G$  um grafo conexo com  $n$  vértices,  $n > 2$  e  $v_i \in V$  tal que  $d_i = n - 1$ . Então, para todo  $v_j$  tal que  $d_j < n - 1$ ,

1.  $c_D(v_i) > c_D(v_j)$ ;
2.  $c_C(v_i) > c_C(v_j)$ ;
3.  $c_{eff}(v_i) > c_{eff}(v_j)$ ;
4.  $c_B(v_i) > c_B(v_j)$ ;
5.  $c_{eig}(v_i) > c_{eig}(v_j)$ .

**Prova:**

1. Trivial.

2. Se  $d_i = n-1$  então  $\sum_{k=1}^n \text{dist}(v_k, v_i) = 1 \cdot (n-1) = n-1$ . Assim, como a centralidade de

proximidade dada por (3.3) é inversamente proporcional a soma das distâncias,

$c_C(v_i) = \frac{1}{n-1}$  é o maior valor que  $c_C(v_i)$  pode alcançar. Portanto,  $c_C(v_i) > c_C(v_j)$ , uma

vez que sendo  $d_j < n-1$ , existe pelo menos um vértice  $v_k$  tal que  $\text{dist}(v_k, v_j) = 2$  e,

assim,  $\sum_{k=1}^n \text{dist}(v_k, v_j) > \sum_{k=1}^n \text{dist}(v_k, v_i)$ .

3. Se  $d_i = n-1$  então a eficiência de  $v_i$  é  $e(v_i) = \max\{\text{dist}(v_i, v_k) : v_k \in V\} = 1$ . Assim, a

centralidade de eficiência dada por (3.5) é inversamente proporcional à eficiência de  $v_i$

e, então,  $c_{\text{eff}}(v_i) = \frac{1}{e(v_i)} = 1$  é o maior valor que  $c_{\text{eff}}(v_i)$  pode alcançar. Portanto,

$c_{\text{eff}}(v_i) > c_{\text{eff}}(v_j)$ , uma vez que sendo  $d_j < n-1$ , existe pelo menos um vértice  $v_k$  tal

que  $\text{dist}(v_j, v_k) = 2$ . Logo  $c_{\text{eff}}(v_j) = \frac{1}{2}$ .

4. Seja  $v_i \in V$  tal que  $d_i = n-1$ . Primeiro vamos provar que, para todo  $v_j \in V$ ,  $v_i \neq v_j$ ,

é verdadeira a desigualdade  $c_B(v_i) \geq c_B(v_j)$ . Para todo  $v_l, v_k \in V$ , se  $\{v_l, v_k\} \in E$  então

$\text{dist}(v_l, v_k) = 1$ ; se  $\{v_l, v_k\} \notin E$  então  $\text{dist}(v_l, v_k) = 2$ , sendo  $v_l$  e  $v_k$  ambos adjacentes a

$v_i$ . Seja  $r$  o número de geodésicas ligando  $v_l$  e  $v_k$ . Considere o vértice  $v_j$  tal que

$v_j \neq v_i, v_j \neq v_l$  e  $v_j \neq v_k$ . Se  $\{v_l, v_k\} \notin E$  então,

$$g_{lk}(v_j) = \begin{cases} 1, & \text{se } \{v_j, v_l\}, \{v_j, v_k\} \in E; \\ 0, & \text{caso contrário.} \end{cases}$$

Assim,  $\frac{g_{lk}(v_j)}{g_{lk}} = \frac{1}{r} \leq \frac{g_{lk}(v_i)}{g_{lk}}$  e a igualdade vale somente se  $\{v_j, v_l\}, \{v_j, v_k\} \in E$ . Vamos

computar a intermediação do vértice  $v_j$ :

$$c_B(v_j) = \sum_{l < k} \frac{g_{lk}(v_j)}{g_{lk}} = \sum_{\{v_l, v_k\} \in E} \frac{g_{lk}(v_j)}{g_{lk}} + \sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_j)}{g_{lk}} = \sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_j)}{g_{lk}} \geq 0.$$

Se  $\sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_j)}{g_{lk}} = 0$ , então  $c_B(v_j) = 0$  e  $c_B(v_i) \geq c_B(v_j)$ . Se  $\sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_j)}{g_{lk}} > 0$ , como

todas as parcelas  $\frac{g_{lk}(v_j)}{g_{lk}}$  positivas pertencem também ao  $\sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_i)}{g_{lk}}$ , temos

$$c_B(v_i) = \sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_i)}{g_{lk}} \geq \sum_{\{v_l, v_k\} \notin E} \frac{g_{lk}(v_j)}{g_{lk}} = c_B(v_j).$$

Logo,  $c_B(v_i) \geq c_B(v_j)$ .

Agora, suponha por absurdo que  $c_B(v_i) = c_B(v_j)$  com  $d_j < n-1$ . Então, existe pelo menos um vértice  $v_k$  tal que  $\{v_j, v_k\} \notin E$ . Para todo  $v_l \in V$ , se  $\{v_l, v_k\} \in E$ , então  $\text{dist}(v_k, v_l) = 1$ ; se  $\{v_l, v_k\} \notin E$  então  $v_k$  e  $v_l$  não contribuem para  $c_B(v_j)$ . Assim,  $c_B(v_i) > c_B(v_j)$ .

5. Seja  $x$  autovetor positivo de  $G$  com relação ao índice  $\rho$ . Sejam  $v_i$  e  $v_j \in V$  tal que

$d_i = n-1$  e  $d_i > d_j$ . Sabemos que  $Ax = \rho x$ . Logo,

$$\rho x_i = \sum_{\{v_k, v_i\} \in E} x_k \text{ e } \rho x_j = \sum_{\{v_k, v_j\} \in E} x_k.$$

Suponha por absurdo que:

$$x_i \leq x_j. \quad (6.1)$$

Daí  $\rho x_i \leq \rho x_j$  pois  $\rho > 1$  e, portanto,  $\sum_{\{v_k, v_i\} \in E} x_k \leq \sum_{\{v_k, v_j\} \in E} x_k$ . Isto significa que:

$$\sum_{\{v_k, v_i\} \in E} x_k - \sum_{\substack{\{v_k, v_j\} \in E \\ v_k \neq v_i}} x_k \leq x_i,$$

$$x_j + \sum_{\substack{\{v_k, v_i\} \in E \\ \{v_k, v_j\} \in E \\ v_k \neq v_j}} x_k \leq x_i. \quad (6.2)$$

De (6.1) e (6.2) temos que  $x_i + \sum_{\substack{\{v_k, v_i\} \in E \\ \{v_k, v_j\} \in E \\ v_k \neq v_j}} x_k \leq x_j + \sum_{\substack{\{v_k, v_i\} \in E \\ \{v_k, v_j\} \in E \\ v_k \neq v_j}} x_k \leq x_i$ . Isto implica em  $\sum_{\substack{\{v_k, v_i\} \in E \\ \{v_k, v_j\} \in E \\ v_k \neq v_j}} x_k \leq 0$ ,

Absurdo! Logo  $x_i > x_j$  e, portanto,  $c_{eig}(v_i) > c_{eig}(v_j)$ , para todo  $v_j \in V$  tal que  $d_j < n-1$ . □

**Exemplo:**

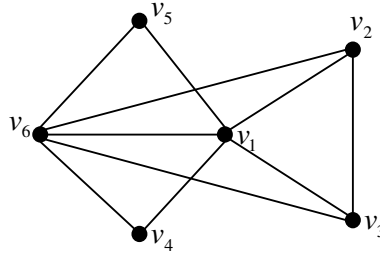


Figura 6.4: Grafo  $G$  com dois vértices universais,  $v_1$  e  $v_6$ .

De fato, para o grafo  $G$  da Figura 6.4 temos:

1.  $c_D(v_1) = c_D(v_6) = 5 > c_D(v_2) = c_D(v_3) = 3 > c_D(v_4) = c_D(v_5) = 2$ .
2.  $c_C(v_1) = c_C(v_6) = 0,2 > c_C(v_2) = c_C(v_3) = 0,142 > c_C(v_4) = c_C(v_5) = 0,125$ .
3.  $c_{eff}(v_1) = c_{eff}(v_6) = 1 > c_{eff}(v_2) = c_{eff}(v_3) = c_{eff}(v_4) = c_{eff}(v_5) = 0,5$ .
4.  $c_B(v_1) = c_B(v_6) = 2,5 > c_B(v_2) = c_B(v_3) = c_B(v_4) = c_B(v_5) = 0$ .
5.  $c_{eig}(v_1) = c_{eig}(v_6) = 0,515 > c_{eig}(v_2) = c_{eig}(v_3) = 0,392 > c_{eig}(v_4) = c_{eig}(v_5) = 0,284$ .

## 6.2. Aplicações

Esta seção ilustra como as medidas de centralidade podem ser usadas em grafos que modelam redes reais. Dois exemplos serão discutidos. O primeiro deles corresponde a um grafo que modela a rede de co-autorias do grupo de pesquisa em Teoria Espectral de Grafos. Tal grupo é formado por estudantes e pesquisadores das instituições CEFET, ENCE, IME, UERJ, UFF, UFMT, UFRJ e UNIRIO. Nós calculamos as medidas de centralidade dos atores desta rede comparando as centralidades, a correlação entre as medidas e fornecemos um *ranking* para os vértices, como sugerem vários artigos. O segundo exemplo é um grafo que modela um grupo de 40 homossexuais masculinos que têm o vírus da AIDS. Eles têm parceiros sexuais com alguns atores da rede e, desta maneira, será inserida uma aresta entre atores que têm contato sexual. Será investigada a relevância dos vértices com base na medida de centralidade de autovetor em comparação com outras medidas já estudadas em [40], utilizando a mesma rede.

### 6.2.1. Rede de Co-autorias

A primeira aplicação das medidas de centralidade considera uma rede de co-autorias entre os pesquisadores do grupo de pesquisa em Teoria Espectral de Grafos do Rio de Janeiro, TEG-RIO, e representada pelo grafo  $G$  da Figura 6.5.

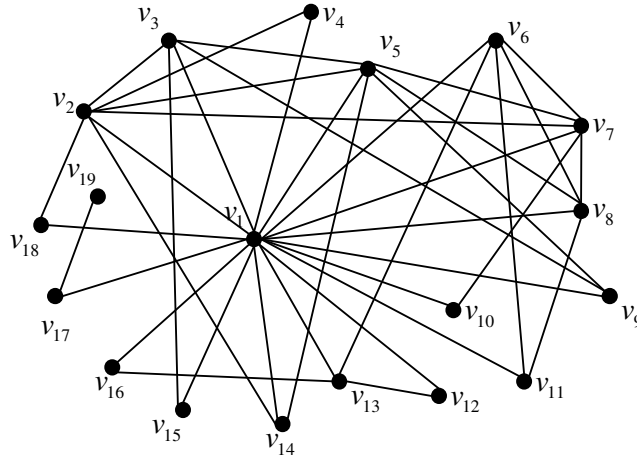


Figura 6.5: Grafo  $G$

O grupo TEG-RIO conta com 19 pesquisadores oriundos de diferentes instituições de ensino superior e sob a coordenação da professora Dr<sup>a</sup> Nair Abreu buscam ampliar a capacidade de produção do conhecimento científico e a transmissão desse conhecimento a futuros membros do grupo na área de Teoria Espectral de Grafos. Os vértices representam os 19 pesquisadores assim definidos:

$v_1$  : Nair Maria Maia de Abreu (UFRJ)

$v_2$  : Renata Raposo Del-Vecchio (UFF)

$v_3$  : Cybele Tavares Maia Vinagre (UFF)

$v_4$  : Ana Isabel de Azevedo Spinola Dias(UFF)

$v_5$  : Maria Aguiéiras Alvarez de Freitas (UFRJ)

$v_6$  : Claudia Marcela Justel (IME)

$v_7$  : Leonardo Silva de Lima (CEFET)

$v_8$  : Carla Silva Oliveira (ENCE)

$v_9$  : Andréa Soares Bonifácio (UNIRIO)

$v_{10}$ : Aroldo José de Oliveira (UFMT)

$v_{11}$ : Joelma Ananias de Oliveira (UFMT)

$v_{12}$ : Luciana Lee (UFRJ)

$v_{13}$ : Lilian Markenzon (UFRJ)

$v_{14}$ : Laura Patuzzi (UFRJ)

$v_{15}$ : Beatriz Malajovich (UNIRIO)

$v_{16}$ : Claudio Santos de Souza (UERJ)

$v_{17}$ : Montauban Moreira de Oliveira Junior (UNIRIO)

$v_{18}$ : Leandro Quintanilha de Freitas (UFRJ)

$v_{19}$ : Jean Guillaume Eon (UFRJ)

e as arestas a relação de co-autoria em trabalhos científicos entre eles.

Na Tabela 6.1, nós mostramos um retrato da rede em questão com as respectivas centralidades medidas com o auxílio dos softwares Netdraw 2.086 e CentiBin. A primeira coluna estabelece um *ranking* para os vértices:

<i>Ranking</i>	Grau	Intermediação	Proximidade	Percursos aleatórios	Autovetor	Eficiência
1	$v_1$ (17)	$v_1$ (110,83)	$v_1$ (0,052)	$v_1$ (0,709)	$v_1$ (0,535)	$v_1$ (0,5)
2	$v_5$ (7)	$v_{17}$ (17)	$v_5$ (0,033)	$v_2$ (0,202)	$v_5$ (0,333)	$v_{17}$ (0,5)
3	$v_2$ (7)	$v_2$ (5,5)	$v_2$ (0,033)	$v_5$ (0,179)	$v_2$ (0,309)	$v_2$ (0,33)
4	$v_7$ (6)	$v_5$ (4,16)	$v_7$ (0,032)	$v_7$ (0,161)	$v_7$ (0,297)	$v_7$ (0,33)
5	$v_8$ (5)	$v_7$ (3,16)	$v_8$ (0,031)	$v_6$ (0,132)	$v_8$ (0,258)	$v_8$ (0,33)



6	$v_3(5)$	$v_3(1,8)$	$v_3(0,031)$	$v_3(0,127)$	$v_3(0,245)$	$v_3(0,33)$
7	$v_6(5)$	$v_6(1,8)$	$v_6(0,031)$	$v_{13}(0,125)$	$v_6(0,234)$	$v_6(0,33)$
8	$v_{13}(4)$	$v_{13}(1,5)$	$v_{13}(0,03)$	$v_8(0,112)$	$v_{14}(0,194)$	$v_{13}(0,33)$
9	$v_9(3)$	$v_8(1,16)$	$v_9(0,029)$	$v_{17}(0,111)$	$v_9(0,183)$	$v_9(0,33)$
10	$v_{11}(3)$	$v_{11}(0)$	$v_{11}(0,029)$	$v_9(0,05)$	$v_{11}(0,169)$	$v_{11}(0,33)$
11	$v_{14}(3)$	$v_{14}(0)$	$v_{14}(0,029)$	$v_{11}(0,049)$	$v_{13}(0,164)$	$v_{14}(0,33)$
12	$v_{17}(2)$	$v_9(0)$	$v_{17}(0,029)$	$v_{14}(0,047)$	$v_{18}(0,139)$	$v_5(0,33)$
13	$v_{10}(2)$	$v_{10}(0)$	$v_{10}(0,028)$	$v_{12}(0,036)$	$v_{10}(0,139)$	$v_{10}(0,33)$
14	$v_4(2)$	$v_4(0)$	$v_4(0,028)$	$v_{16}(0,036)$	$v_4(0,137)$	$v_4(0,33)$
15	$v_{18}(2)$	$v_{18}(0)$	$v_{18}(0,028)$	$v_{15}(0,031)$	$v_{15}(0,128)$	$v_{18}(0,33)$
16	$v_{15}(2)$	$v_{15}(0)$	$v_{15}(0,028)$	$v_{18}(0,028)$	$v_{12}(0,115)$	$v_{15}(0,33)$
17	$v_{16}(2)$	$v_{16}(0)$	$v_{16}(0,028)$	$v_4(0,028)$	$v_{16}(0,115)$	$v_{16}(0,33)$
18	$v_{12}(2)$	$v_{12}(0)$	$v_{12}(0,028)$	$v_{10}(0,026)$	$v_{17}(0,09)$	$v_{12}(0,33)$
19	$v_{19}(1)$	$v_{19}(0)$	$v_{19}(0,019)$	$v_{19}(0)$	$v_{19}(0,014)$	$v_{19}(0,33)$

Tabela 6.1: Os pesquisadores estão representados pelos vértices  $v_i$ , listados em ordem

decrecente de acordo com cada medida de centralidade. As centralidades dos vértices

estão entre parênteses.

Como mostra a Tabela 6.1, todas as medidas indicam que o vértice  $v_1$  é o mais central. O que já era natural de se esperar, sendo a professora Nair Abreu, criadora do grupo TEG-RIO, responsável pela orientação da grande maioria dos trabalhos, e quem reuniu os pesquisadores para trabalharem em grupo. Os pesquisadores que correspondem aos vértices  $v_2$  e  $v_5$  possuem o segundo mais alto grau em relação aos demais vértices da rede. Isto significa que tais pesquisadores são mais ativos em

pesquisas relativas à Teoria Espectral de Grafos e possuem o segundo maior número de co-autorias. Também estão estruturalmente bem posicionados na rede, segundo a centralidade de proximidade, possuindo mais chances que os demais vértices de aumentar o número de suas publicações. Isto se justifica por estes pesquisadores estarem mais próximos dos outros na rede. Outros resultados merecem destaque. Podemos observar, por exemplo, que o vértice  $v_{17}$  possui uma centralidade de grau menor que a centralidade dos vértices  $v_2$  e  $v_5$  e, no entanto, em relação à centralidade de intermediação, é considerado o segundo mais relevante. A justificativa disso parece estar no fato que o vértice  $v_{17}$  é o único responsável pela troca de informações entre o vértice  $v_{19}$  e os demais vértices do grafo. A medida de centralidade de eficiência é a mais limitada em relação a capacidade de distinguir diferenças entre as centralidades dos vértices, pois 90% possuem a mesma centralidade por esta medida.

Ainda para avaliar as relações entre as medidas de centralidade dos vértices no grafo da Figura 6.5, foi usada análise de regressão linear. O objetivo desta análise é explicar o quão próximos ou não estão os valores das centralidades obtidos neste exemplo. Para isso, plotam-se os dados num diagrama de dispersão. Se o diagrama indicar uma relação que de modo geral é linear, ajusta-se os dados por uma reta que forneça a melhor aproximação. A estimação dos parâmetros desta reta é determinada pelo Método dos Mínimos Quadrados. As fórmulas utilizadas para isso são:

$$y = a + bx, \quad (6.2)$$

$$b = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} \text{ e} \quad (6.3)$$

$$a = \bar{y} - b\bar{x}, \quad (6.4)$$

onde  $x_i$  e  $y_i$  são os valores correspondentes às medidas de centralidades sob análise e,  $\bar{x}$  e  $\bar{y}$ , as respectivas médias aritméticas de  $x_i$  e  $y_i$ .

Também calculamos o coeficiente de correlação simples, que mede o grau de associação linear entre cada par de medidas de centralidade, cuja fórmula é dada por:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (6.5)$$

A Figura 6.6 mostra dois diagramas de dispersão, o primeiro relaciona os valores da centralidade de proximidade  $\times$  centralidade de grau, e o segundo, valores da centralidade de autovetor  $\times$  centralidade de grau.

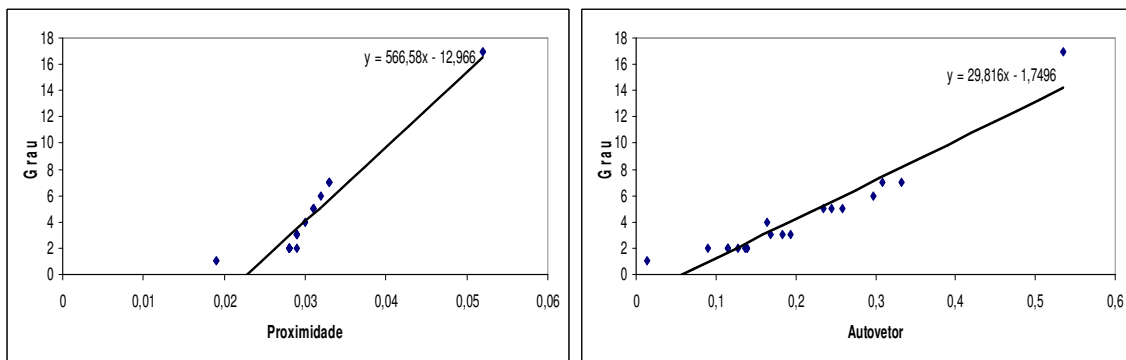


Figura 6.6: Gráficos de dispersão entre as medidas de centralidade de proximidade e grau (à esquerda) e centralidade de autovetor e grau (à direita), referentes ao grafo da

Figura 6.5.

Como mostra a Figura 6.6, podemos verificar que a centralidade de grau está altamente correlacionada com a centralidade de proximidade ( $r = 0,9504$ ) e com a centralidade de autovetor ( $r = 0,9527$ ). Assim, neste caso, vértices com grau ou proximidade mais elevada tendem também a ter mais elevada centralidade de autovetor. Por outro lado, a centralidade de autovetor mostrou-se estar moderadamente correlacionada com a centralidade de intermediação ( $r = 0,7049$ ). Desta forma, é possível que a contagem dos caminhos mais curtos não seja a medida mais apropriada

para avaliar a centralidade dos vértices neste exemplo. A Tabela 6.2 mostra as correlações entre as demais medidas de centralidade.

	Grau	Intermediação	Proximidade	Percursos aleatórios	Autovetor	Eficiência
Grau	1	0,8646	0,9504	0,9471	0,9527	0,5179
Intermediação		1	0,8787	0,9414	0,7049	0,7831
Proximidade			1	0,9526	0,9194	0,5944
Percursos aleatórios				1	0,8785	0,6673
Autovetor					1	0,3450
Eficiência						1

Tabela 6.2: Coeficientes de correlação entre as seis medidas de centralidade usadas para avaliar a rede de co-autorias do grupo TEG-RIO.

De um modo geral, podemos concluir que as centralidades de grau e proximidade podem ser consideradas similares à análise da rede neste contexto, sendo assim, tanto as ligações diretas quanto a proximidade entre os pesquisadores os tornam mais integrados em determinada pesquisa, o que pode facilitar a criação de trabalhos na área em comum.

### 6.2.2. Rede de AIDS

A rede composta por 40 homossexuais que possuem contatos sexuais e foram diagnosticados com o vírus da AIDS foi previamente estudada por STEPHENSON e ZELEN [40]. Nós usaremos esta mesma rede para analisar a relevância dos vértices com base na centralidade de autovetor e comparar com as medidas já estudadas em

[40], dado que a centralidade de autovetor não foi analisada no referente artigo. O grafo  $H$  da Figura 6.7 ilustra esta rede.

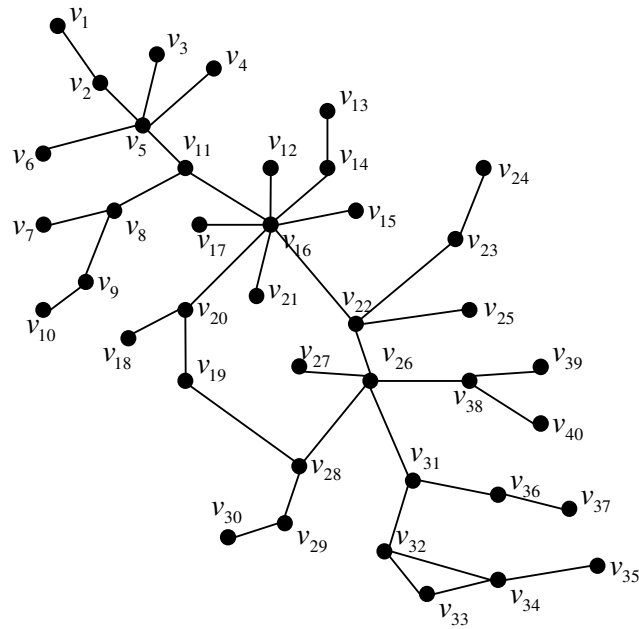


Figura 6.7: Grafo  $H$ . Rede que exhibe a relação sexual entre 40 homossexuais.

A Tabela 6.3 resume os resultados para as cinco medidas de centralidade dos vértices do grafo  $H$  e fornece um *ranking* para os mesmos de acordo com cada medida.

Ranking	Grau	Intermediação	Proximidade	Percursos aleatórios	Autovetor
1	$v_{16}$ (8)	$v_{16}$ (476)	$v_{16}$ (0,0090)	$v_{16}$ (0,6722)	$v_{16}$ (0,5866)
2	$v_{26}$ (5)	$v_{26}$ (380)	$v_{22}$ (0,0088)	$v_{26}$ (0,5472)	$v_{22}$ (0,3394)
3	$v_5$ (5)	$v_{22}$ (365)	$v_{26}$ (0,0082)	$v_{22}$ (0,4289)	$v_{11}$ (0,2726)
4	$v_{22}$ (4)	$v_{11}$ (314)	$v_{11}$ (0,0077)	$v_{11}$ (0,4237)	$v_{20}$ (0,2540)
5	$v_{31}$ (3)	$v_{31}$ (206)	$v_{20}$ (0,0071)	$v_{31}$ (0,2780)	$v_{26}$ (0,2414)

6	$v_{28} (3)$	$v_5 (179)$	$v_{14} (0,0068)$	$v_{28} (0,2739)$	$v_{14} (0,2097)$
7	$v_8 (3)$	$v_8 (110)$	$v_{19} (0,0068)$	$v_{20} (0,2503)$	$v_{17} (0,1881)$
8	$v_{32} (3)$	$v_{32} (108)$	$v_{31} (0,0068)$	$v_5 (0,2415)$	$v_{21} (0,1881)$
9	$v_{34} (3)$	$v_{28} (100)$	$v_{17} (0,0067)$	$v_{19} (0,2073)$	$v_{15} (0,1881)$
10	$v_{20} (3)$	$v_{20} (85)$	$v_{28} (0,0067)$	$v_8 (0,1484)$	$v_{12} (0,1881)$
11	$v_{38} (3)$	$v_{38} (75)$	$v_{21} (0,0067)$	$v_{32} (0,1466)$	$v_5 (0,1517)$
12	$v_{11} (3)$	$v_{19} (45)$	$v_{15} (0,0067)$	$v_{38} (0,1012)$	$v_{28} (0,1324)$
13	$v_{14} (2)$	$v_{14} (38)$	$v_{23} (0,0067)$	$v_{34} (0,0679)$	$v_{19} (0,1239)$
14	$v_9 (2)$	$v_9 (38)$	$v_{12} (0,0067)$	$v_9 (0,0512)$	$v_{23} (0,1213)$
15	$v_2 (2)$	$v_2 (38)$	$v_{25} (0,0066)$	$v_2 (0,0512)$	$v_8 (0,1118)$
16	$v_{36} (2)$	$v_{36} (38)$	$v_{38} (0,0064)$	$v_{36} (0,0512)$	$v_{25} (0,1089)$
17	$v_{19} (2)$	$v_{34} (38)$	$v_5 (0,0063)$	$v_{14} (0,0512)$	$v_{31} (0,1058)$
18	$v_{23} (2)$	$v_{23} (38)$	$v_{27} (0,00628)$	$v_{23} (0,0512)$	$v_{38} (0,0975)$
19	$v_{29} (2)$	$v_{29} (38)$	$v_8 (0,00621)$	$v_{29} (0,0512)$	$v_{18} (0,0815)$
20	$v_{33} (2)$	$v_{33} (0)$	$v_{18} (0,0056)$	$v_{33} (0,0332)$	$v_{27} (0,0774)$
21	$v_{17} (1)$	$v_{17} (0)$	$v_{32} (0,0055)$	$v_{17} (0)$	$v_{13} (0,0672)$
22	$v_{39} (1)$	$v_{39} (0)$	$v_{36} (0,00546)$	$v_{39} (0)$	$v_2 (0,0542)$
23	$v_{13} (1)$	$v_{13} (0)$	$v_{13} (0,0054)$	$v_{13} (0)$	$v_{32} (0,0505)$
24	$v_6 (1)$	$v_6 (0)$	$v_{29} (0,0054)$	$v_6 (0)$	$v_6 (0,0486)$
25	$v_{21} (1)$	$v_{21} (0)$	$v_{24} (0,0053)$	$v_{21} (0)$	$v_3 (0,0486)$
26	$v_{24} (1)$	$v_{24} (0)$	$v_{39} (0,0051)$	$v_{24} (0)$	$v_4 (0,0486)$

27	$v_{10} (1)$	$v_{10} (0)$	$v_2 (0,0051)$	$v_{10} (0)$	$v_{29} (0,0473)$
28	$v_{18} (1)$	$v_{18} (0)$	$v_6 (0,0051)$	$v_{18} (0)$	$v_9 (0,0399)$
29	$v_{40} (1)$	$v_{40} (0)$	$v_{40} (0,0051)$	$v_{40} (0)$	$v_{24} (0,0389)$
30	$v_3 (1)$	$v_3 (0)$	$v_3 (0,0051)$	$v_3 (0)$	$v_{36} (0,0378)$
31	$v_4 (1)$	$v_4 (0)$	$v_4 (0,0051)$	$v_4 (0)$	$v_7 (0,0358)$
32	$v_{15} (1)$	$v_{15} (0)$	$v_9 (0,005)$	$v_{15} (0)$	$v_{39} (0,0312)$
33	$v_{30} (1)$	$v_{30} (0)$	$v_7 (0,005)$	$v_{30} (0)$	$v_{40} (0,0312)$
34	$v_{12} (1)$	$v_{12} (0)$	$v_{34} (0,00467)$	$v_{12} (0)$	$v_{34} (0,0269)$
35	$v_7 (1)$	$v_7 (0)$	$v_{33} (0,00465)$	$v_7 (0)$	$v_{33} (0,0248)$
36	$v_1 (1)$	$v_1 (0)$	$v_{37} (0,0045)$	$v_1 (0)$	$v_1 (0,0174)$
37	$v_{27} (1)$	$v_{27} (0)$	$v_{30} (0,0044)$	$v_{27} (0)$	$v_{30} (0,0151)$
38	$v_{25} (1)$	$v_{25} (0)$	$v_1 (0,0043)$	$v_{25} (0)$	$v_{10} (0,0128)$
39	$v_{35} (1)$	$v_{35} (0)$	$v_{10} (0,0042)$	$v_{35} (0)$	$v_{37} (0,0121)$
40	$v_{37} (1)$	$v_{37} (0)$	$v_{35} (0,0039)$	$v_{37} (0)$	$v_{35} (0,0086)$

Tabela 6.3: Os homens estão indicados com os números de 1 a 40 e listados em ordem decrescente de acordo com cada medida de centralidade. As centralidades estão entre parênteses.

Como podemos observar, todas as cinco medidas mostram que o ator representado pelo vértice  $v_{16}$  é o mais central da rede. Os vértices  $v_{17}$ ,  $v_{12}$ ,  $v_{21}$  e  $v_{15}$ , apesar de terem apenas um contato sexual, pois estão ligados somente ao vértice  $v_{16}$ , possuem centralidade de autovetor relativamente alta, o que os fazem figurar entre as 10 primeiras posições do *ranking*. Isto se explica pelo fato de que um dado ator pode ter

apenas um contato sexual, mas se seu correspondente vizinho for bastante influente, o primeiro também será. Por exemplo, este vértice terá alta possibilidade de transmitir o vírus HIV a outros elementos da rede. Em relação às 20 últimas posições, as centralidades de grau, intermediação e percursos aleatórios não são capazes de distinguir a relevância dos vértices, diferentemente das centralidades de autovetor e proximidade. Outro fator interessante são as semelhantes posições entre os vértices, ao se avaliar a centralidade de proximidade e centralidade de autovetor.

Desta forma, podemos concluir que em um contexto de transmissão de doenças, por exemplo a AIDS, investigada também em [10], [42] e [60], uma pessoa central na rede deve estar relacionada à sua contribuição para a propagação da doença. Neste caso, a centralidade de autovetor é uma boa medida do poder de propagação. A propagação ocorrerá vértice a vértice e dependerá da centralidade do vértice contaminado. Os vizinhos do vértice contaminado serão contaminados e contaminarão seus próprios vizinhos, e assim por diante. Se um vértice tem o poder de propagação elevado (isto é, uma alta centralidade de autovetor), contaminará outros vértices mais rapidamente do que outros vértices com baixo poder de propagação.

### **6.2.3. Outras Aplicações**

De modo geral, podemos dizer que há situações em que a centralidade de um ator depende apenas do número ligações diretas entre o mesmo e os demais atores da rede. Como é o caso, por exemplo, da capacidade de um ator trocar uma determinada informação diretamente com outros. Neste caso, um ator na rede é mais central se possui o maior número de contatos. No setor de transportes, as medidas de centralidade também têm sido usadas com frequência, como observado em [59] e [61]. Em [59], por exemplo, foi realizado um estudo visando identificar soluções aos problemas na



área de influência de um corredor ferroviário e no entorno da estação ferroviária do ramal Saracuruna (de Barão de Mauá até Gramacho), na Região Metropolitana do Rio de Janeiro. O estudo investigou a relação entre algumas das medidas de centralidade abordadas nesta dissertação. Em relação aos resultados, o autor destaca duas medidas de grande relevância neste contexto. A centralidade de proximidade, onde um bairro é tão mais central quanto menor o caminho que se precise percorrer para, a partir dele, alcançar os outros bairros da rede, e a centralidade de intermediação, que é o potencial dos bairros que servem de intermediários, atuando como "ponte", facilitando o fluxo de viagens na rede. Desta forma, o conhecimento das medidas de centralidade contribuiu na identificação de áreas centrais, indicando locais favoráveis à integração modal, possibilitando uma melhoria na qualidade do serviço de transporte. Já em [30], foi realizado um estudo de centralidade em uma rede de alimentos em um ecossistema. Os vértices do grafo correspondem às várias espécies presentes dentro de um ecossistema e, para cada par de vértices predador/presa, há uma aresta entre o vértice predador e o vértice presa. Neste caso, a medida utilizada foi a centralidade via conectividade algébrica, buscando medir a mudança no ecossistema devido à remoção de uma espécie, que corresponde a relevância de um vértice para a rede caso tenha de ser omitido da mesma. Apesar da falta de simetria no relacionamento entre o predador e a presa, o autor considera as redes de alimento como grafos não orientados, isto mostra a necessidade de se buscar outras aplicações desta medida e que, por ser recente, é algo a ser estudado com mais detalhes. Recentemente, em [62], medidas de centralidade foram aplicadas no estudo de confiabilidade em redes. Foi constatado, através de simulações, que no caso da ampliação de uma rede, a inserção de uma aresta entre o vértice menos central e o mais distante dele levou ao melhor resultado em relação à melhoria na confiabilidade do novo grafo gerado em comparação com o original. Além disso, foi

apresentado um exemplo de aplicação prática considerando a análise da ampliação de uma rede de entrega de mercadorias, modelada por um grafo, de uma suposta empresa distribuidora.

Diante destas análises, nós tentamos mostrar através destas aplicações que a interpretação do conceito “centralidade” é muito dependente do contexto, e cada medida de centralidade ajuda a analisar uma situação particular representada pela rede. É importante notar que nenhuma destas medidas é superior a outras. Todas são apropriadas para se responder algumas, mas não todas as perguntas referentes à análise da rede.

## Capítulo 7 – Considerações Finais

Nesta dissertação foram reunidos parâmetros de grafos utilizados para medir a centralidade de um vértice bem como os resultados teóricos a eles relacionados e as possíveis aplicações em problemas de natureza real. Com isso, pudemos analisar o comportamento de tais medidas em diversas situações.

Uma das principais constatações é que não há uma medida universal, ou seja, aquela que, sob qualquer hipótese e em qualquer grafo, poderia ser melhor que as demais. Assim, cada medida pode ser no máximo considerada como mais apropriada a determinados tipos de problemas. A centralidade de grau é aplicável, por exemplo, em redes que representam uma situação estática e nós queremos avaliar a importância dos vértices em função de seus contatos diretos. A centralidade de grau é uma medida local, porque o valor da centralidade de um vértice é determinado somente pelo número de seus vértices adjacentes. Em problemas que tratam da difusão de algum conteúdo, o tempo que leva uma determinada informação, por exemplo, para se espalhar de um vértice a todos os demais, depende da proximidade entre eles. Assim, medidas de centralidade baseadas em proximidade são mais interessantes neste contexto. Se a importância de um vértice está no potencial para controle do fluxo através da rede, medidas de centralidade baseadas em intermediação são mais apropriadas. As posições são vistas como estruturalmente centrais, na medida em que estão entre outras e pode, conseqüentemente, facilitar, impedir ou permitir a transmissão do fluxo, que pode ser uma mensagem, um recurso, transporte, etc. Para processos de propagação na rede, como por exemplo, a propagação de uma doença, a centralidade de autovetor é mais indicada dado que mede a importância de um vértice levando em consideração a importância de seus vizinhos. Se a importância de um vértice está atrelada à

vulnerabilidade da rede, uma medida baseada em conectividade algébrica é mais apropriada e, por fim, redes que apresentam a força ou intensidade da relação entre os vértices sugerem medidas de centralidade que são aplicáveis em grafos valorados.

Como parte das sugestões para pesquisas futuras propomos analisar o comportamento das medidas apresentadas considerando grafos orientados. Há uma série de medidas aplicáveis a estas estruturas, inclusive muitas delas variações das medidas de centralidade apresentadas nesta dissertação, por exemplo, o *Page Rank*, que avalia a importância de um site, ou página na internet, considerando a quantidade de *links* que cada página possui. Trata-se de uma variação da centralidade de autovetor para grafos orientados, usado pelo motor de busca Google para fornecer uma ordenação das páginas relevantes de acordo com determinado assunto pesquisado.

Por fim, foi apresentada em [62] uma relação entre as medidas de centralidade e a confiabilidade em redes. Constatou-se, através de simulações, que as medidas de centralidade são uma importante ferramenta na maximização da confiabilidade da rede, ao indicar os principais vértices que deveriam ser conectados para aumentar a confiabilidade da mesma. Além disso, um estudo recente relacionando a centralidade de um vértice com a conectividade algébrica de um grafo foi apresentado na Seção 4.4, e, como ainda não se tem investigadas comparações com as demais medidas de centralidade, acreditamos ser uma boa linha de pesquisa. Estas são as principais sugestões para futuras investigações.

## Referências Bibliográficas

- [1] ABREU, N.M.M., DEL-VECCHIO, R.R., VINAGRE, C.T.M., et al, 2007, “Introdução à Teoria espectral de Grafos com aplicações”, Notas em Matemática aplicada, CNMAC/SBMAC, UFSC.
- [2] ABREU, N.M.M, 2005, “Teoria espectral dos grafos: um híbrido entre a álgebra linear e a matemática discreta e combinatória com origens na química quântica”, Uma publicação da sociedade brasileira de matemática aplicada e computacional, v. 6, n 1, pp. 1-10.
- [3] ANDRADE, L. N., 2004, **Introdução à Computação Algébrica com o Maple**. Sociedade Brasileira de Matemática.
- [4] BAVELAS, A., 1950, “Communication patterns in task oriented groups”, *Journal of the Acoustical Society of America*, v. 22, pp. 725-30.
- [5] BIGGS, N., 1993, **Algebraic Graph Theory**. 2 ed. New York, Cambridge.
- [6] BONACICH, P., 1987, “Power and Centrality: A Family of Measures”, *The American Journal of Sociology*, v. 92, n 5, pp. 1170-1182.
- [7] BONACICH, P., 2007, “Some unique properties of eigenvector centrality”, *Social Networks*, v. 29, pp. 555-564.

- [8] BOAVENTURA NETTO, P. O., 2001, **Grafos: Teoria, Modelos, Algoritmos**. 2 ed, São Paulo, Editora Blucher.
- [9] BORGATTI, S. P., EVERETT, M. G., 2006, “A Graph-Theoretic perspective on centrality”, *Social Networks*, v.28, pp. 466-484.
- [10] BORGATTI, S. P., 1995, “Centrality and AIDS”, *Connections*, 18(1), pp. 112-114.
- [11] BORGATTI, S. P., 2005, “Centrality and network flow”, *Social Networks*, 27(1), pp. 55-71.
- [12] BORGATTI, S. P., 2002, **NetDraw: Graph Visualization Software**. Harvard: Analytic Technologies.
- [13] BORTOLI, A.L., CARDOSO, C., FACHIN, M.G.P., CUNHA, R.D., 2003, “Introdução ao Cálculo Numérico”, 2 edição. Disponível em <http://www.chasqueweb.ufrgs.br/~carolina.manica/cap6.pdf>, acesso em 17/03/2010.
- [14] BRANDES, U., FLEISCHER, D., 2005, “Centrality Measures Based on Current Flow”, V. Diekert and B. Durand (Eds.): STACS 2005, LNCS 3404, pp. 533–544.
- [15] DEL-VECCHIO, R.R., GALVÃO, D.J.C., LIMA, L.S., et al., 2009, “Medida de centralidade da Teoria dos Grafos aplicada a Fundos de Ações no Brasil”. In: Pesquisa Operacional na Gestão do conhecimento, pp. 533-545.

- [16] DIESTEL, R., 2005, **Graph Theory**. Eletronic version of third Edition, New York, Springer-Verlag.
- [17] KOSCHUTZKI, D., 2006, **CentiBiN: Centralities in Biological Networks Software**. Leibniz Institute of Plant Genetics and Crop Plant Research, Germany.
- [18] DONADELLI, J., 2006, “Uma Introdução à teoria espectral de Grafos”, Monografia preliminar do Departamento de Informática da Universidade Federal do Paraná, UFRPR. Disponível em <http://www.inf.ufpr.br/jair/MANUSCRIPTS/>, acesso em 04/04/2009.
- [19] EVERETT, M. G., SINCLAIR, P., 2004, “Some centrality results new and old”, *Journal of Mathematical Sociology*, v. 28, pp. 215-227.
- [20] FRANCO, N.M.B., 2006. **Cálculo Numérico**. Editora Pretice Hall.
- [21] FREEMAN, L. C., 1978/79, “Centrality in Social Networks: Conceptual Clarification”, *Social Networks*, v. 1, pp. 215-239.
- [22] FREEMAN, L. C., 1977, “A Set of Measures of Centrality Based on Betweenness”, *Sociometry*, v. 40, n. 1, pp. 35-41.
- [23] FREEMAN, L. C., ROEDER, D., MULHOLLAND, R. R., 1979/80, “Centrality in Social Networks: II. Experimental Results”, *Social Networks*, v. 2, pp. 119-141.

- [24] FREEMAN, L. C., BORGATTI, S. P., WHITE, D. R., 1991, “Centrality in valued graphs: A measure of betweenness based on network flow”. *Social Networks*, v. 13, pp. 141-154.
- [25] FREEMAN, L. C., 1980, “The Gatekeeper, Pair-Dependency and Structural Centrality”, *Quality and Quantity*, v. 14, pp. 585-592.
- [26] GRASSI, R., STEFANI, S., TORRIERO, A., 2009, “Centrality in Organizational Networks”, *International Journal of Intelligent Systems*, v. 25, pp. 253-265.
- [27] GRASSI, R., STEFANI, S., 2007, “Some New Results on the Eigenvector Centrality”, *Journal of Mathematical Sociology*, v. 31, pp. 237–248.
- [28] GRASSI, R., STEFANI, S., 2006, “Central vertices in networks: a unified approach”, Rapporto di ricerca n 105, Dipartimento di metodi Quantitativi per la Scienze Economiche Ed Aziendali, Università di Milano Bicocca.
- [29] HAGE, P., HARARY, F., 1995, “Eccentricity and centrality in networks”, *Social Networks*, v. 17, pp. 57-63.
- [30] KIRKLAND, S., 2009, “Algebraic connectivity for vertex-deleted subgraphs, and a notion of vertex centrality”, *Discrete Mathematics*.



- [31] KOSCHUTZKI, D., LEHMANN, K.A., PEETERS, L., et al, 2005, “Centrality Indices”, *Network Analysis*, pp. 16-61.
- [32] NAIMZADA, A.K., et al., 2009, “Networks, Topology and Dynamics”, Lecture Notes in Economic and Mathematical Systems, Springer-Verlag, Berlin Heidelberg.
- [33] NEWMAN, M. E. J., 2005, “A measure of Betweenness centrality based on random walks”, *Social Networks*, v. 27, pp. 39-54.
- [34] POOLE, D., 2004, **Algebra Linear**. Editora Thomson Pioneira.
- [35] RUHNAU, B., 2000, “Eigenvector-centrality—a node-centrality?”, *Social Networks*, v. 22, pp. 357–365.
- [36] SLATER, P.J., 1975, “Maximin facility location”, **Journal of National Bureau of Standards**, 79B:107-115.
- [37] WASSERMAN, S., FAUST, K., 1994, **Social Network Analysis: Methods and Applications**. Cambridge: Cambridge University Press.
- [38] WILF, H. S., 2001, “Searching the web with eigenvectors”, Disponível em <http://www.math.upenn.edu/~wilf/>, acesso em 30/04/2010.
- [39] BRANDES, U., 2001, “A faster algorithm for betweenness centrality”, *Journal of Mathematical Sociology*, v.25, pp. 163-177.

- [40] STEPHENSON, K.A., ZELEN, M., 1989, "Rethinking centrality: methods and examples", *Social Networks*, v.11, pp. 1-37.
- [41] BORGATTI, S. P., EVERETT, M. G., 1999, "The Centrality of Groups and Classes", *Journal of Mathematical Sociology*, v.23, pp. 181-201.
- [42] BELL, D. C., ATKISON, J. S., CARLSON, J. W., 1999, "Centrality for disease transmission networks", *Social Networks*, v.21, pp. 1-21.
- [43] SINCLAIR, P., 2005, "Betweenness Centralization for Bipartite Graphs", *Journal of Mathematical Sociology*, v.29, pp. 25-31.
- [44] FRIEDKIN, N. E., 1991, "Theoretical Foundations for Centrality Measures", *The American Journal of Sociology*, v.96, pp 1478-1504.
- [45] BORGATTI, S. P., EVERETT, M. G., 1997, "Network analisys of 2-mode data", *Social Networks*, v.19, pp. 243-269.
- [46] NEWMAN, M. E. J., 2001, "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality", *Physical Review E*, v.64, 016132.
- [47] BOLDRINI, J. L., 1980, **Álgebra Linear**. 3 ed, São Paulo, Editora Harbra.

- [48] BERNÓ, B., 2007, “Networks with Closeness Incentives”, In: Bielefeld University, Institute of Mathematical Economics/working Papers.
- [49] GRASSI, R., SCAPELLATO, R., STEFANI, S., TORRIERO, A., 2008, “Betweenness centrality: extremal values and structural properties”. Lecture notes in economic and mathematical systems: network, topology and dynamics: theory and applications to economic and social systems. Springer, Heidelberg.
- [50] MARTELETO, R. M., TOMAÉL, M. I., 2006, “Redes Sociais: posições dos atores no fluxo da informação”, revista eletrônica de biblioteconomia e ciência da informação, v.10, pp. 75-91.
- [51] HUSSAIN, D. M. A., 2007, “Destabilization of Terrorist Networks through Argument Driven Hypothesis Model”, *Journal of software*, v. 2, n. 6, pp. 22-29.
- [52] NETO, B. M. R., 2005, *Parâmetros de grafos em vulnerabilidade de redes*. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- [53] FIEDLER, M., 1973, “Algebraic connectivity of graphs”, *Czechoslovak Mathematical Journal*, v. 23, pp. 298-305.
- [54] SABIDUSSI, G., 1966, “The centrality index of a graph”, *Psychometrika*, v. 31, pp. 581-603.

- [55] MACKENZIE, K. D., 1966, “Structural centrality in communications networks”, *Psychometrika*, v. 31, pp. 17-25.
- [56] MOXLEY, R. L., MOXLEY, N. F., 1974, “Determining point-centrality in uncontrived social networks”, *Sociometry*, v. 37, pp. 120-133.
- [57] BEAUCHAMP, M. A., 1965, “An improved index of centrality”, *Behavioral Science*, v. 10, pp. 161-163.
- [58] SHAW, M. E., 1964, **Communication networks**, In L. Berkowitz (ed). Disponível em <http://books.google.com.br/books>, acesso em 18/07/2009.
- [59] GONÇALVES, J. A. M., PORTUGAL, L. S., BOAVENTURA NETTO, P. O., 2005, “As potencialidades de indicadores de centralidade no estudo de um corredor ferroviário”, XIX ANPET – Congresso de pesquisa e ensino em transportes, Recife.
- [60] ROTHENBERG, R. B., et al, 1995, “Choosing a centrality measure: Epidemiologic correlates in the Colorado Springs study of social networks”, *Social Networks*, v. 17, pp. 273-297.
- [61] SCHEUER, J., PORTA, S., 2006, “Centrality and Connectivity in Public Transport Networks and their Significance fir Transport Sustainability in Cities”, presented at World Planning Schools Congress, Mexico.

[62] SILVA, T. S. A., 2010, *Um estudo de Medidas de Centralidade e Confiabilidade em Redes*, Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

[63] HORN, R. A., JOHNSON, C. R., 1985, **Matrix Analysis**, Cambridge Univ. Press.

# Índice Remissivo

## A

Arco, 7  
Arestas, 1  
Árvore, 9  
Ator, 1  
Autocentralidade, 41  
Automorfismo, 73  
Autovalor, 3,10,40  
Autovetor, 3,41,42

## C

Cadeia, 8  
Caminho, 8  
Capacidade do corte, 68  
Centralidade, 1  
- de grau, 2,17,18,64  
- de proximidade, 2,17,20,64  
- de eficiência, 3,17,22,23  
- centróide, 24  
- radial, 24  
- de informação, 24,25  
- de intermediação, 2,17,26,27  
- de intermediação de fluxo,  
3,64,69  
- de percursos aleatórios,  
3,17,31,33,34  
- de grafos, 36,37  
- de autovetor, 3,40,41,65  
- via conectividade algébrica,  
4,40,60,61  
Centralidade relativa, 19,20,30,69  
Centro, 36  
Ciclo, 8  
Circuito, 8  
Círculo espectral, 12  
Complexidade, 22,30,34,69  
Componentes conexas, 8,57  
Comprimento de caminho, 8  
Conjunto independente, 9  
Conectividade Algébrica, 15  
Corte, 68

## D

Diâmetro, 8,24  
Distância, 8

## E

Espectro, 10  
Excentricidade, 8

## F

Fluxo, 67  
Fluxo máximo, 3,68

## G

Grafo, 1,6  
- bipartido, 9  
- completo, 7  
- conexo, 8  
- desconexo, 8,57  
- estrela, 9  
- não orientado, 7,10  
- orientado, 7  
- regular, 7,52  
- simples, 7,10  
- valorado, 5,64  
Grau, 7  
- máximo, 7,55  
- médio, 7  
- mínimo, 7  
Geodésicas, 3,8

## I

Índice do grafo, 10  
Intermediação parcial, 26  
Isomorfismo, 72

## L

Laço, 6

## M

Matriz  
- de adjacência, 9  
- diagonal, 14  
- dos pesos, 65  
- irredutível, 12  
- laplaciana, 15  
- simétrica, 3,10  
Medidas de centralidade, 2,17  
Método das potências, 42  
Método dos mínimos quadrados, 82

## Multiplicidade

- algébrica, **10**
- geométrica, **12**

## O

Ordem, **14,19**

## P

Polinômio característico, **10**

Polinômio característico laplaciano, **15**

Percurso, **8**

- aleatório, **32**
- aleatório absorvente, **32**

Perron-Frobenius, **12**

Peso, **64**

Ponto de articulação, **60**

## R

Raio espectral, **12,51,57**

Raleigh-Ritz, **13**

Ranking, **80**

Rede, **1**

- de AIDS, **84**
- de co-autorias, **78**
- social, **1,17**

Regressão linear, **82**

## S

Subgrafo, **7**

- induzido, **8,60**
- gerador, **8**

Submatriz principal, **48**

## T

Trilha, **8**

## V

Vértice, **1**

- alvo, **32,68**
- automorficamente

equivalentes, **73**

- fonte, **32,68**
- isolado, **19**
- pendente, **51**
- terminal, **29**
- universal, **7,74**

Vizinhança, **6**

Vulnerabilidade, **60**