CEFET/RJ
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO

CENTRALIDADE DE ARESTAS: APLICAÇÕES NO PROBLEMA DA ÁRVORE MÍNIMA DE STEINER E RELAÇÕES COM A CENTRALIDADE DE VÉRTICES.

Alayna Marcia Garcia Conceição

Tese de Mestrado apresentado ao Programa de Pós-graduação em Engenharia de Produção e Sistemas, do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Produção e Sistemas.

Orientador(a): Celso Marques da Silva Junior

Coorientador(a): Ana Luísa Carvalho Furtado

Rio de Janeiro

Agosto de 2025

Centralidade de arestas: aplicações no problema da árvore mínima de Steiner e relações com a centralidade de vértices.

Tese de Mestrado em Engenharia de Produção e Sistemas, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ.

Alayna Marcia Garcia Conceição

Aprovado por:

_____

Celso Marques da Silva Junior, D.Sc. (orientador)

_____

Ana Luísa Carvalho Furtado, D.Sc. (Coorientadora)

_____

Carla Silva Oliveira, D.Sc.

_____

Claudia Marcela Justel, D.Sc.

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

C744 Conceição, Alayna Marcia Garcia
Centralidade de arestas: aplicações no problema da árvore mínima de Steiner e relações com a centralidade de vértices / Alayna Marcia Garcia Conceição. — 2025.
94f. : il. color. , enc.

Dissertação (Mestrado) Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, 2025.
Bibliografia : f. 93-94
Orientador: Celso Marques da Silva Junior
Coorientadora: Ana Luísa Carvalho Furtado

1. Teoria dos grafos. 2. Algoritmos. 3. Análise de redes (Planejamento). 4. Árvores mínimas de Steiner. I. Silva Junior, Celso Marques da. (Orient.). II. Furtado, Ana Luísa Carvalho. (Coorient.). III. Título.

CDD 511.5

Elaborada pela bibliotecária Tania Mello – CRB/7 nº 5507/04

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder forças, sabedoria e perseverança durante toda essa jornada.

À minha mãe, pelo amor incondicional, apoio constante e por sempre acreditar em mim, mesmo nos momentos mais difíceis.

Este trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Sou grata pelo suporte que possibilitou a continuidade dos meus estudos e a realização desta pesquisa.

RESUMO

Centralidade de arestas: aplicações no problema da árvore mínima de Steiner e relações com a centralidade de vértices

No estudo da Teoria dos Grafos, a análise de centralidade tem sido uma ferramenta fundamental para compreender a importância de vértices nas estruturas de rede. Em 2001, Girvan e Newman [2001] propuseram o estudo da centralidade de arestas, através do conceito de centralidade de intermediação de arestas. Neste trabalho estudamos duas potencialidades da centralidade de arestas, considerando tanto a centralidade de intermediação de arestas como também a centralidade de proximidade de arestas, posteriormente definida. A primeira trata-se de uma investigação sobre a frequência em que as arestas mais centrais estão associadas aos vértices mais centrais em um grafo de acordo com suas respectivas medidas de centralidade. Inicialmente, utilizamos a medida de intermediação como ponto de partida para análise. Como resultado, identificamos particularidades em grafos do tipo vassoura dupla equilibrada. A partir desse resultado, estudamos a associação entre as arestas mais centrais e os vértices mais centrais em grafos vassoura dupla equilibrada de acordo com suas respectivas medidas de centralidade de intermediação e de proximidade. A segunda potencialidade é uma proposta de duas novas variações da heurística KMB, um dos métodos mais utilizados para resolver o Problema da Árvore de Steiner Mínima, utilizando o conceito de centralidade de proximidade de arestas. Essas variações foram testadas computacionalmente em grafos do repositório SteinLib, com o intuito de comparar seus desempenhos entre si e em relação à heurística KMB original.

Palavras-chave:

Centralidade de Arestas; Heurística KMB; Problema da Árvore de Steiner Mínima; Vassoura dupla equilibrada

ABSTRACT

Edge Centrality: Applications to the Minimum Steiner Tree Problem and Relationships with Vertex Centrality.

In the study of Graph Theory, centrality analysis has been a fundamental tool for understanding the importance of vertices within network structures. In 2001, Girvan and Newman [2001] proposed the study of edge centrality through the concept of edge betweenness centrality. In this work, we explore two potential applications of edge centrality, considering both edge betweenness centrality and edge closeness centrality, which is defined later. The first application investigates the frequency with which the most central edges are associated with the most central vertices in a graph according to their respective centrality measures. Initially, we use the betweenness measure as a starting point for the analysis. As a result, we identify particular characteristics in balanced double broom graphs. Based on this finding, we study the association between the most central edges and the most central vertices in balanced double broom graphs according to their respective betweenness and closeness centrality measures. The second application proposes two new variations of the KMB heuristic, one of the most widely used methods to solve the Minimum Steiner Tree Problem, utilizing the concept of edge closeness centrality. These variations were computationally tested on graphs from the SteinLib repository to compare their performances with each other and against the original KMB heuristic.

Keywords:

Edge Centrality; KMB Heuristic; Minimum Steiner Tree Problem; Balanced Double Broom

SUMÁRIO

1 INTRODUÇÃO 10

2 CONCEITOS PRELIMINARES 15

2.1 Teoria dos Grafos: 15

2.2 Algoritmos para obtenção de caminhos mínimos e algoritmos para árvores geradoras mínimas 18

3 MEDIDAS DE CENTRALIDADE: 26

3.1 Centralidade de vértices: 26

3.2 Centralidade de Arestas: 28

3.3 Relação entre Centralidade de intermediação de vértices e centralidade de intermediação de arestas: 31

4 RELAÇÃO ENTRE CENTRALIDADES DE VÉRTICES E ARESTAS EM VASSOURAS DUPLAS EQUILIBRADAS 40

4.1 Vassoura Dupla: 40

4.2 Centralidade de intermediação de vértices: 41

4.2.1. Centralidade de intermediação de vértices quando o número de vértices na espinha é ímpar: 43

4.2.2. Centralidade de intermediação de vértices quando o número de vértices na espinha é par: 47

4.3 Centralidade de intermediação de arestas: 50

4.3.1. Centralidade de intermediação de arestas quando o número de vértices da espinha é ímpar: 52

4.3.2. Centralidade de intermediação de arestas quando o número de vértices na espinha é par: 53

4.4 Comparando centralidade de intermediação de vértices e arestas: 54

4.4.1. Número ímpar de vértices na espinha: 55

4.4.2. Número par de vértices na espinha: 56

4.5 Centralidade de proximidade de vértices: 57

4.5.1. Centralidade de proximidade de vértices quando número de vértices na espinha é ímpar: 60

4.5.2. Centralidade de proximidade de vértices quando o número de vértices na espinha é par: 61
4.6 Centralidade de proximidade de arestas: 63
4.6.1. Centralidade de proximidade de arestas quando o número de vértices na espinha é ímpar: 66
4.6.2. Centralidade de proximidade de arestas quando o número de vértices na espinha é par: 67
4.7 Comparando centralidade de proximidade de vértices e arestas: 68
4.7.1. Número ímpar de vértices na espinha: 69
4.7.2. Número par de vértices na espinha: 69
5 O PROBLEMA DA ÁRVORE DE STEINER MÍNIMA E A HEURÍSTICA KMB..... 71
5.1 Problema da Árvore de Steiner mínima em Grafos 71
5.2 Heurística KMB 73
5.3 Variações da heurística KMB: 75
5.4 Comparação dos resultados das variações da heurística KMB: 79
6 CONSIDERAÇÕES FINAIS 90

LISTA DE ILUSTRAÇÕES

FIGURAS

Figura 2.1: Grafo completo Kn 16

Figura 2.2: Exemplo de grafo simples. 17

Figura 2.3: Árvore T 17

Figura 2.4: Exemplo de um Subgrafo gerador de G 18

Figura 2.5: Exemplo para aplicação do algoritmo de Dijsktra. Fonte: a auto ... 19

Figura 2.6: Exemplo para aplicação do algoritmo de Bellman- Ford.Fonte: a autora. 22

Figura 2.7: Exemplo para aplicação do algoritmo de Prim. Fonte: a autora. 24

Figura 3.1: Exemplo para cálculo das medidas de centralidades apresentadas 28

Figura 3.2: Exemplo para cálculo da medida de centralidade de aresta 30

Figura 3.3 Exemplos de grafos para cada configuração analisada 35

Figura 3.4: Exemplo de grafo considerado fora das categorias analisadas 35

Figura 3.5:Exemplos de casos restritos em grafos simples com 6 vértices. 38

Figura 3.6:Exemplo de discordância entre centralidade de aresta de intermediação e de vértice de intermediação 39

Figura 4.1: Vassoura dupla equilibrada 40

Figura 4.2: funções de centralidades de vértices analisadas e as raízes obtidas para γ 45

Figura 4.3: Mudança na centralidade do grafo vassoura equilibrada com $n = 11$. 47

Figura 4.4: funções de centralidades de vértices analisadas e as raízes obtidas para γp 49

Figura 4.5:Mudança na centralidade do grafo vassoura equilibrada com $n = 12$. 50

Figura 4.6: Ilustração das arestas mais centrais quando n é ímpar e $n \geq 7$. ... 53

Figura 4.7: Ilustração da aresta mais central quando n é par e $n \geq 8$..... 54

Figura 4.8: Representação dos vértices e arestas mais centrais para e ímpar 55

Figura 4.9:Representação dos vértices e arestas mais centrais para e par. 57

Figura 4.10: Ilustração dos vértices e arestas mais centrais para n ímpar e $n \geq 7$. 69

Figura 4.11: Ilustração dos vértices e das arestas mais centrais para $n \geq 6$ e n par. 70

Figura 5.1:Exemplo de Árvore de Steiner mínima S de G. Adaptado de Goldbarg e Goldbarg [2012]..... 72

Figura 5.2:Exemplo de aplicação da Heurística KMB 75

Figura 5.3:Exemplo de aplicação da heurística KMB_EB_1 78

TABELAS

Tabela 1: Iterações do algoritmo de Dijsktra 20

Tabela 2: Iterações do algoritmo de Bellman-Ford..... 22

Tabela 3: Iterações do algoritmo de Prim..... 24

Tabela 4: Valores de centralidade obtidos 28

Tabela 5: Valores obtidos para centralidade de intermediação e proximidade de aresta 30

Tabela 6: Resultado da análise de associações entre centralidade de intermediação de vértices e arestas nos grafos simples com $6 \leq n \leq 8$. 36

Tabela 7: Resultado da análise de associações entre centralidade de intermediação de vértices e arestas nas árvores com $6 \leq n \leq 16$ 36

Tabela 8: Exemplos de valores obtidos para função γ(p)..... 46

Tabela 9: Exemplos de valores obtidos para função γ 49

Tabela 10: Teste da monotonicidade de KMB_EB_1 para diferentes valores de α 82

Tabela 11: Teste da monotonicidade de KMB_EB_2 para diferentes valores de β. 82

Tabela 12: heurística KMB_EB_1 para diferentes valores de α no conjunto C. 83

Tabela 13:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto B. 84

Tabela 14:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto C 84

Tabela 15:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto B 86

Tabela 16:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto C 87

1 INTRODUÇÃO

O tema desta pesquisa insere-se na Teoria dos Grafos. A Teoria dos Grafos é uma área fundamental da matemática que examina a relação entre objetos de um conjunto. Um grafo pode ser definido como uma estrutura abstrata que representa os elementos de um dado conjunto, denominados de vértices, e suas relações de interdependências, denominadas de arestas. Essa área de estudo começou formalmente com o trabalho de Leonhard Euler em 1736, ao resolver o problema das sete pontes de Königsberg, um marco considerado o primeiro teorema da Teoria dos Grafos [Biggs, 1993]. Desde então, a Teoria dos Grafos evoluiu de forma notável, influenciando campos diversos, como a ciência da computação, a biologia e as ciências sociais.

No século XVII, Fermat propôs o problema de encontrar um ponto adicional, em um plano euclidiano, tal que a soma das distâncias deste ponto a três pontos dados fosse mínima. Este problema, posteriormente, foi expandido para grafos, sendo conhecido como o problema da Árvore de Steiner mínima. Este tema serviu de ponto de partida para esta dissertação, que surgiu do interesse desenvolvido durante o meu Trabalho de Conclusão de Curso (TCC), no qual abordei o caso euclidiano do problema.

Durante a pesquisa para o presente trabalho, explorei algoritmos para a obtenção de Árvores de Steiner, destacando-se a heurística desenvolvida por Kou, Markowsky e Berman em 1981, conhecida como heurística KMB [Kou et al., 1981], devido à sua eficiência na construção de árvores mínimas. Contudo, o aprofundamento no estudo revelou variações da heurística KMB baseadas na centralidade de intermediação de arestas, um tema que despertou um interesse significativo.

Dessa forma, esta dissertação apresenta duas partes principais. Uma parte concentra-se na análise da heurística KMB e suas variações, enquanto uma segunda parte explora o conceito de centralidade de arestas, com destaque para as relações entre centralidade de arestas e de vértices.

O problema da árvore de Steiner mínima é um problema clássico dentro do escopo da Teoria dos Grafos que visa encontrar uma árvore de custo mínimo que conecta no grafo um subconjunto específico de vértices previamente determinados, denominados de terminais. O problema da árvore de Steiner mínima é amplamente conhecido na literatura para otimização de rede em várias áreas, como roteamento e estimativa de comprimento de fios, redes WDM óticas, design de VLSI, roteamento de redes, sistemas de comunicação óptica e sem fio, redes de transporte e distribuição, além da reconstrução de árvores filogenéticas [Chen e Hsieh 2022]. Este problema é classificado como NP-difícil e, portanto, encontrar uma solução exata em tempo polinomial é considerado impraticável para grandes instâncias devido à complexidade computacional [Boudjelida e Lemouari 2021]. Por essa razão, geralmente são utilizados métodos de aproximação para obter a árvore de Steiner mínima.

Um dos métodos mais populares para encontrar uma aproximação de uma árvore de Steiner mínima é a Heurística KMB, a qual constrói a árvore de Steiner utilizando caminhos de comprimento mínimo entre os terminais e uma árvore geradora mínima de um grafo auxiliar cujo vértices são todos terminais [Boudjelida e Lemouari 2021]. Embora a heurística KMB possa produzir aproximações eficientes, ela não estabelece uma ordem de prioridade quando existem vários caminhos mínimos possíveis entre pares de vértices terminais. Consequentemente, o custo da árvore de Steiner pode variar dependendo dos caminhos escolhidos [Fujita *et* al. 2016], [Boudjelida e Lemouari 2021].

Fujita *et* al. [2016a, 2016b] introduziram variações da heurística KMB, denominadas neste trabalho por KMB_EB_1 e KMB_EB_2, que utilizam a centralidade de intermediação de arestas para ajustar os custos das arestas e, assim, potencialmente reduzir o custo total da árvore de Steiner obtida. A centralidade de intermediação de uma aresta mede a importância de uma aresta com base na frequência com que ela é usada nos caminhos mínimos entre pares de vértices. Posteriormente, Boudjelida e Lemouari [2021], investigaram a viabilidade e os benefícios das abordagens propostas por Fujita *et* al.[2016a, 2016b].

Tanto em Fujita *et* al. [2016a] como em Fujita *et* al. [2016b], com o intuito de comparar o desempenho das heurísticas KMB_EB_1 e KMB_EB_2 com a heurística KMB, os autores utilizaram a biblioteca *Steinlib*, um repositório de instâncias de teste de diferentes tamanhos para o Problema da Árvore de Steiner Mínima em Grafos [SteinLib 2024]. Para cada conjunto de grafos, a biblioteca apresenta valores ótimos que heurísticas e algoritmos devem encontrar ao obter uma árvore de Steiner mínima. Esta biblioteca foi utilizada no desafio "*11th DIMACS Implementation Challenge in Collaboration with ICERM: Steiner Tree Problems*," organizado pelo DIMACS (*Center for Discrete Mathematics and Theoretical Computer Science*), um centro dedicado à matemática discreta e ciência da computação aberto à comunidade. Um dos objetivos do desafio foi incentivar a implementação e avaliação experimental de heurísticas de aproximação para a resolução do problema de árvore de Steiner mínima [DIMACS 2024].

As variações da heurística KMB mostraram-se particularmente intrigantes por utilizarem centralidade de arestas, despertando o interesse na relação entre este conceito com a centralidade de vértices. O conceito de centralidades em um grafo foi inicialmente proposto por Bavelas em 1948. A partir da ideia de que os vértices do grafo são atores e as arestas são as relações entre eles, a centralidade de vértices indica os atores mais importantes, ou centrais, nas relações com outros atores. Dessa forma, um ator mais central está estrategicamente localizado de forma a transmitir ou até interromper o fluxo de informação entre os membros do grupo. Muitas medidas de centralidades foram propostas ao longo do tempo para investigar as interações entre os atores de um grafo. Em 2001, Girvan e Newman [2001] propõe analisar as arestas mais centrais de um grafo, propondo a medida de intermediação de aresta.

Com intuito de explorar o papel da medida de centralidade de arestas escolhida nas heurísticas KMB_EB_1 e KMB_EB_2, propomos duas novas heurísticas, KMB_EC_1 e KMB_EC_2, que utilizam a centralidade de proximidade de arestas no lugar da centralidade de intermediação de arestas. Esta medida foi escolhida por avaliar a importância de uma aresta em relação aos caminhos mínimos, verificando o quão crucial ela é para minimizar a

distância total entre os vértices que conecta e todos os outros vértices no grafo [Bröhl e Lehnertz 2022].

Nesse contexto, este trabalho busca colaborar com a literatura analisando a variação da heurística KMB na resolução do problema da árvore de Steiner mínima, a partir do estudo da influência de um parâmetro de controle utilizado nas heurísticas KMB_EB_1 e KMB_EB_2. Além disto, a eficácia das heurísticas KMB_EB_1, KMB_EB_2, KMB_EC_1 e KMB_EC_2 são comparadas com a da heurística original KMB, complementando a investigação já iniciada em Fujita *et* al. [2016a, 2016b]. São destacados também aspectos restritivos nestas heurísticas, em especial quanto ao parâmetro de controle, que foram omitidos nos artigos em que são apresentados originalmente.

Para avaliar os pontos discutidos, foram realizados testes computacionais utilizando os conjuntos de grafos B e C do repositório *SteinLib*. Esses testes analisaram o desempenho das heurísticas mencionadas e geraram conclusões que são detalhadas na Seção 5.4 deste texto. Vale destacar que esses resultados foram submetidos e aceitos para publicação nos anais do Simpósio Brasileiro de Pesquisa Operacional (SBPO) em 2024.

Adicionalmente, durante o desenvolvimento desta pesquisa, surgiu um questionamento pertinente em relação à centralidade de arestas: "As arestas mais centrais estão associadas aos vértices mais centrais em um grafo?" Para investigar essa questão, foram gerados dois conjuntos de grafos: um composto por grafos simples, conexos e com até 8 vértices, e outro formado por árvores com até 16 vértices. A partir desses conjuntos, analisou-se a frequência com que os vértices incidentes às arestas mais centrais — conforme a medida de centralidade de intermediação de arestas — também figuram entre os vértices mais centrais segundo a respectiva medida de centralidade de vértices. Os resultados dessa análise, apresentados detalhadamente na Seção 3.3, incluem a descrição dos grafos utilizados, os critérios adotados para identificar as arestas e vértices mais centrais e a comparação entre as centralidades. Limitações técnicas e computacionais encontradas durante os testes também são abordadas.

A partir dessa análise inicial, definiu-se o foco em uma classe específica de grafos, as vassouras duplas equilibradas, com o propósito de examinar a relação entre centralidade de vértices e centralidade de arestas. Nessa etapa, são apresentados exemplos em que os vértices mais centrais estão associados às arestas mais centrais, bem como casos em que essa correspondência não se verifica.

Este trabalho está organizado da seguinte maneira: no Capítulo 2 são apresentados os conceitos preliminares necessários para a compreensão dos temas tratados nesta dissertação. Na Seção 2.1, discutimos fundamentos da Teoria dos Grafos, enquanto na Seção 2.2, abordamos algoritmos clássicos para obtenção de caminhos e árvores mínimas, que serão utilizados nas heurísticas discutidas posteriormente no Capítulo 5. No Capítulo 3, exploramos as medidas de centralidade, que são ferramentas essenciais para a análise da importância relativa de elementos em um grafo. A centralidade de vértices é apresentada na Seção 3.1 e a centralidade de arestas na Seção 3.2. Na Seção 3.3, discutimos a relação entre ambas as centralidades em um conjunto de grafos conexos e árvores, utilizando a medida de centralidade de intermediação. O Capítulo 4 é dedicado ao estudo das relações entre centralidades de vértices e arestas especificamente em grafos vassoura dupla equilibrada. Este capítulo apresenta os métodos de análise utilizados e discute os padrões observados nesta família de grafos. No Capítulo 5, introduzimos o problema da árvore de Steiner mínima, e a heurística KMB, uma das abordagens para sua solução. Na Seção 5.1, são descritos os fundamentos do problema da árvore de Steiner mínima. Na Seção 5.2, detalha-se a heurística KMB e, na Seção 5.3, apresentam-se variações dessa heurística que incorporam centralidades de arestas. Por fim, na Seção 5.4, realiza-se uma comparação entre os resultados obtidos pelas diferentes variações. No Capítulo 6 apresentamos as considerações finais deste trabalho.

2 CONCEITOS PRELIMINARES

Neste capítulo, serão apresentados os conceitos preliminares essenciais para a compreensão dos temas abordados ao longo desta dissertação. Na Seção 2.1, discutiremos conceitos fundamentais da Teoria dos Grafos. Na Seção 2.2, abordaremos alguns dos algoritmos clássicos para a obtenção de caminhos e árvores mínimas que serão utilizados nas heurísticas discutidas no Capítulo 5.

2.1 Teoria dos Grafos:

As principais definições e conceitos sobre grafos apresentados a seguir foram extraídos do livro de Szwarcfiter [2018].

Um *grafo* $G = (V, E)$ é um conjunto finito não vazio V e um conjunto E de pares não ordenados de elementos distintos de V. Utilizaremos a notação $n = |V|$ e $m = |E|$. Os elementos de V são os *vértices* e os de E são as *arestas* de G, respectivamente. Cada aresta $e \in E$ será denotada pelo par de vértices $e = (v, w)$ que a forma. Nesse caso, os vértices v, w são os extremos (ou extremidades) da aresta e, sendo denominados vértices *adjacentes ou vizinhos*. A aresta e é dita *incidente* a ambos v, w. Duas arestas que possuem um extremo comum são chamadas de *adjacentes*. A *vizinhança* de um vértice $v \in V$, denotada por $N(v)$, é o conjunto de todos os vértices que são adjacentes a v. O *grau do vértice* $v \in \mathrm{V}$, $d(v)$, indica o número de vértices adjacentes a v possui. Um vértice de grau 1 é dito *vértice pendente ou folha*. Um grafo é dito *não direcionado* quando suas arestas (v, w) não são direcionadas. Assim, se (v, w) é aresta de $G = (V, E)$ então tanto v é adjacente a w como w é adjacente a v.

Um *grafo* é uma estrutura abstrata que formaliza relações de interdependência entre elementos de um conjunto. Esta estrutura pode ser visualizada através de uma representação geométrica, na qual seus vértices correspondem a pontos distintos do plano em posições arbitrárias, e cada aresta (v, w) é associada uma linha arbitrária unindo os pontos correspondentes a v e w .

Nesta dissertação, consideraremos apenas *grafos simples*, ou seja, grafos que não possuem laços (arestas que conectam um vértice a si mesmo),

não apresentam arestas múltiplas (mais de uma aresta entre o mesmo par de vértices) e não são orientados (grafos não direcionados). Esses grafos servirão como base para todas as análises e discussões que se seguirão.

Dados dois grafos $G_1 = (V_1, E_1)$ e $G_2 = (V_2, E_2)$, com $|V_1| = |V_2| = n$, existe uma função unívoca $f : V_1 \rightarrow V_2$, tal que $(v, w) \in E_1$ se e somente se $\big(f(v), f(w)\big) \in E_2$, para todo $v, w \in V_1$? Em caso positivo, G_1 e G_2 são ditos *isomorfos* entre si.

Certos grafos têm características que os diferenciam dos outros. Denomina-se *grafo completo*, K_n , com n vértices aquele em que quaisquer dois vértices distintos são adjacentes. A Figura 2.1 apresenta o grafo K_4.

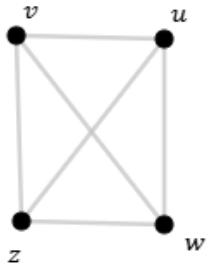
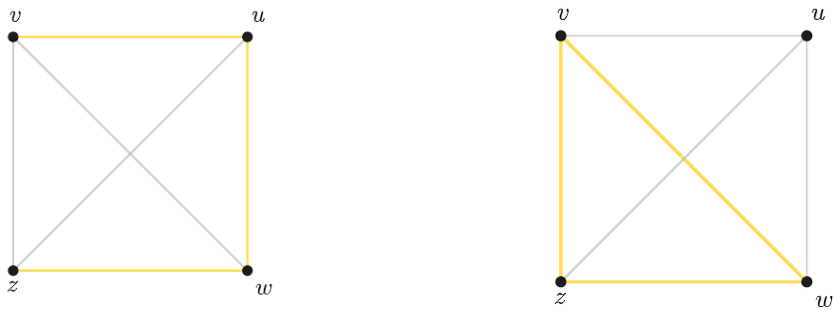


Figura 2.1: Grafo completo K_n

Uma sequência de vértices distintos $v_1, v_2, \dots, v_k$ tal que $\big(v_j, v_{j+1}\big) \in E$, $1 \leq j < k$ é denominada de *caminho* de v_1 até v_k. Diz-se, então, que v_1 alcança v_k. Um caminho de k vértices é formado por $k - 1$ arestas. O valor $k - 1$ é o comprimento do caminho. Denomina-se ciclo o grafo obtido a partir do caminho $v_1, v_2, \dots, v_k, k \geq 3,$ com a inclusão da aresta (v_1, v_k). Um grafo que não possui ciclos é denominado de acíclico. Além disso, um grafo é dito *conexo* quando existe um caminho entre cada par de vértices de G. Caso contrário G é *desconexo*. Na Figura 2.2, ilustramos um exemplo (em amarelo) de um caminho (Figura 2.2 a)) e de um ciclo (Figura 2.2 b)) em um grafo G, respectivamente como v, u, w, z e v, z, w, v.



a) Caminho em G b) Ciclo em G

Figura 2.2: Exemplo de grafo simples.

Denomina-se distância $d(v,w)$ entre dois vértices v,w de um grafo o comprimento do menor caminho entre v e w, quando este existir. Este caminho é chamado de *caminho mínimo*. Diversos algoritmos foram desenvolvidos e estabelecidos na literatura para a obtenção de caminhos mínimos em grafos, tais como Dijkstra [1959] e Bellman- Ford [1958], como veremos na Seção 2.2. No contexto destes algoritmos, usa-se que se não houver caminho entre v e w, então $d(v,w) = \infty$.

Um *grafo ponderado* $G = (V,E,c)$ (nas arestas) é um grafo juntamente com uma função c, que associa um número real não negativo, $c(e)$, a cada aresta $e \in E$, denotado custo da aresta e.

Denomina-se *árvore* um grafo $T = (V,E)$ que seja acíclico e conexo. Se um vértice v da árvore T possuir grau igual a um, então v é uma *folha*. Em uma árvore T existe um, e somente um, caminho entre qualquer par de vértices. Na Figura 2.3 há uma exemplificação de uma árvore T com seis vértices e cinco arestas.

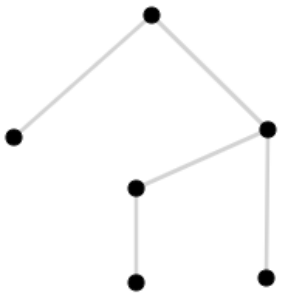
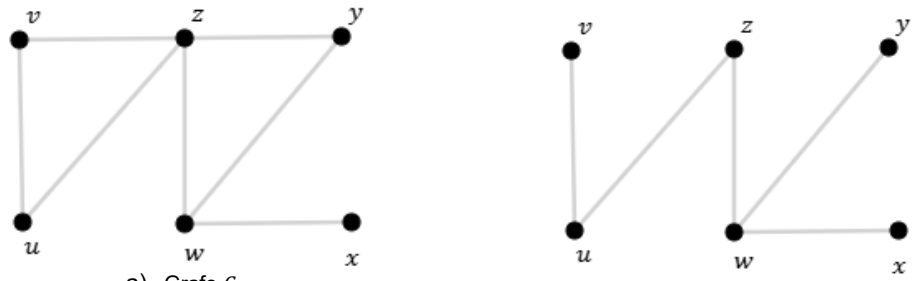


Figura 2.3: Árvore T

As árvores são fundamentais na Teoria dos Grafos, com aplicações úteis na construção de estruturas de dados e na resolução de uma ampla gama de problemas práticos. São algoritmos consolidados na literatura para a obtenção de árvores geradoras mínimas Kruskal [1956] e Prim [1957].

Nomeia-se por *subgrafo* H de $G\,(V,\,E)$ um grafo no qual $V(H) \subseteq V(G)$ e $E(H) \subseteq E\,(G)$. No caso em que $V(H) = V(G)$, este subgrafo é denominado *subgrafo gerador,* como ilustramos na Figura 2.4. Quando o subgrafo gerador é uma árvore, ele recebe o nome de *árvore geradora*. Sendo o custo de uma árvore geradora de um grafo ponderado G determinado pela soma das ponderações de suas arestas, uma *árvore geradora mínima de* G é aquela que possui o menor custo entre todas as árvores geradoras possíveis em G .



a) Grafo G b) Subgrafo gerador de G

Figura 2.4: Exemplo de um Subgrafo gerador de G

2.2 Algoritmos para obtenção de caminhos mínimos e algoritmos para árvores geradoras mínimas

Para obtenção de caminhos mínimos entre vértices de um grafo, dois algoritmos são amplamente reconhecidos na literatura, são eles: Dijkstra [1959] e Bellman- Ford [1958]. Além dos caminhos mínimos, outro problema central na Teoria dos Grafos é a determinação de uma árvore geradora mínima. Um dos algoritmos para obtenção de árvores geradoras mínimas mais reconhecido na literatura, e que será utilizado neste trabalho, é o algoritmo de Prim [1957].

2.2.1 Dijkstra [1959]:

O algoritmo de Dijkstra é um algoritmo para encontrar o caminho mais curto de um vértice de origem para todos os outros vértices em um grafo ponderado. O algoritmo assume que todos os custos das arestas são não negativos. O pseudocódigo do algoritmo é apresentado a seguir.

Algoritmo de Dijsktra

Entrada: Grafo $G = (V, E, c)$, custos $c(u, v) \geq 0$ para cada $aresta\ (u, v) \in E,\ u, v \in V$ e um vértice de origem $F \in V$.

Inicialização:
$N' = \{F\}$ // Conjunto de vértices processados, começando pelo vértice origem F.
$d(F) = 0$ // Defina a distância da origem para ela mesma como 0.

Para cada $v \in V - N'$:

$d(v) = c(F, v)$ // Inicialize as distâncias dos outros vértices a partir de F.

Atualização das distâncias:

Enquanto $N' \neq V$:

Escolher $u \in V - N'$que minimize $d(u)$ // Escolha o vértice não processado com a menor distância.

$N' = N' \cup \{u\}$ // Adicione u ao conjunto de vértices processados.

Para cada $v \in V - N'$:

$d(v) \coloneqq \min\{d(v), d(u) + c(u, v)\}$// Atualize a distância de v se uma rota passando por u for mais curta.

Saída: Uma lista com os menores caminhos de F até todos os outros vértices.

Fonte: Adaptado de Szwarcfiter [2018].

Na Figura 2.5 apresentamos um exemplo de aplicação do algoritmo de Dijsktra. Neste exemplo, o vértice origem, do qual se deseja obter os caminhos mínimos é o vértice v. Cada processo do algoritmo é indicado pela iteração i. Em cada iteração os valores das distâncias são alterados apenas se for menor do que o valor anterior, conforme demonstrado na Tabela 1.

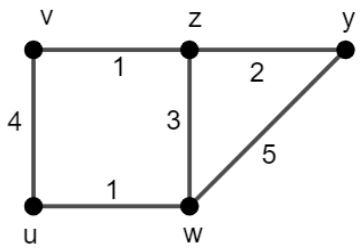


Figura 2.5: Exemplo para aplicação do algoritmo de Dijsktra. Fonte: a auto

Tabela 1: Iterações do algoritmo de Dijsktra

Iteração	Vértice Escolhido	Distâncias Atualizadas	Conjunto de Distâncias (d)	Observações
$i = 0$	$F = v\,(0)$	$d(v,z) = 1$, $d(v,u) = 4$	$d(v) = 0, d(z) = 1$, $d(u) = 4, d(y) = \infty$, $d(w) = \infty$	Inicialmente, todos os vértices têm distância infinita, exceto $v(0)$.
$i = 1$	$z\,(1)$	$d(z,y) = 3$, $d(z,w) = 4$	$d(v) = 0, d(z) = 1$, $d(u) = 4, d(y) = 3$, $d(w) = 4$	Vértice z é processado, atualizando y e w.
$i = 2$	$y\,(3)$	Nenhuma	$d(v) = 0, d(z) = 1$, $d(u) = 4, d(y) = 3$, $d(w) = 4$	Nova distância para w não é menor, então não há atualizações.
$i = 3$	$u\,(4)$	Nenhuma	$d(v) = 0, d(z) = 1$, $d(u) = 4, d(y) = 3$, $d(w) = 4$	Nenhuma atualização ocorre, pois, as distâncias calculadas já são ótimas.
$i = 4$	$w\,(4)$	Nenhuma	$d(v) = 0, d(z) = 1$, $d(u) = 4, d(y) = 3$, $d(w) = 4$	Todos os vértices foram processados, encerrando o algoritmo.

Inicialmente, a distância para a origem $F = v$ é 0, e para todos os outros vértices é definida a partir de v, sendo indicada como infinita para os vértices que não são vizinhos a v. Na primeira iteração $(i = 0)$, o vértice v (distância 0) é escolhido e as distâncias de seus vizinhos z e u são atualizadas com 1 e 4, respectivamente. Na segunda iteração $(i = 1)$, o vértice z (distância 1) é o escolhido e as distâncias de seus vizinhos y e w são atualizadas com 3 e 4, respectivamente. Na terceira iteração $(i = 2)$, escolhe-se o vértice y (distância 3), mas nenhuma distância é atualizada pois a nova distância calculada para w não é menor que a já registrada. Na quarta iteração $(i = 3)$, escolhe-se o vértice u (distância 4), mas novamente nenhuma distância é atualizada. Na quinta iteração $(i = 4)$, escolhe-se o vértice w (distância 4), e o algoritmo termina, pois todos os vértices foram processados. O resultado apresenta a distância mínima da origem v para todos os outros vértices: $d(v,z) = 1,\ d(v,y) = 3,\ d(v,u) = 4,$ $d(v,w) = 4$.

2.2.2 Bellman – Ford [1958]:

O algoritmo de Bellman-Ford é um algoritmo para encontrar o caminho mais curto de um vértice de origem para todos os outros vértices em um grafo ponderado, que pode conter arestas com custos negativos. O algoritmo é capaz de detectar ciclos de custo negativo no grafo. Ciclo de custo negativo refere-se a um ciclo no grafo em que a soma dos custos das arestas que o compõem é negativa. Esses ciclos são problemáticos ao tentar determinar caminhos mínimos, pois permitem reduzir indefinidamente o custo de um caminho ao passar repetidamente pelo ciclo. O algoritmo de Bellman-Ford é capaz de verificar e indicar que tais ciclos existem. Caso um ciclo de custo negativo seja detectado, o algoritmo interrompe a execução e informa sua existência, já que isso inviabiliza a determinação de caminhos mínimos válidos para todos os vértices. O pseudocódigo do algoritmo é apresentado a seguir.

Algoritmo Bellman – Ford

Entrada: $G = (V, E, c)$, onde G é um grafo direcionado e ponderado, que pode conter custos negativos, e F um vértice origem.

Passo 1: Inicialização

$dist = [\,]$ // defina uma lista para armazenar as distâncias

$dist[F] = 0$ // defina a distância da origem F para si mesma como 0

Para cada vértice v em V:

$dist[v] = \infty$ // A distância inicial para todos os vértices é infinita.

Passo 2: Relaxamento

Para i de 1 até $|V|$: // Repetir $|V|$vezes

Para cada aresta (u, v) em E:

Se $dist[u] + c(u, v) < dist[v]$: // Se um caminho mais curto for encontrado

$dist[v] = dist[u] + c(u, v)$ // Atualize a distância

Passo 3: Verificação de Ciclos Negativos

Para cada aresta (u, v) em E:

Se $dist[u] + c(u, v) < dist[v]$:

Parar: O grafo contém um ciclo de custo negativo.

Parar: O grafo não contém ciclo de custo negativo.

Saída: Uma lista com os menores caminhos de F até todos os outros vértices.

Fonte: Adaptado de Szwarcfiter [2018].

O algoritmo de Bellman-Ford executa o passo de relaxamento para todas as arestas repetidamente $|V|$ vezes, onde V é o número de vértices. Esta

abordagem garante que todas as distâncias mais curtas sejam calculadas corretamente, mesmo que existam arestas com custos negativos.

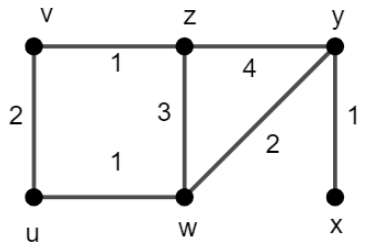


Figura 2.6: Exemplo para aplicação do algoritmo de Bellman- Ford.Fonte: a autora.

Na Figura 2.6 apresentamos um exemplo de aplicação do algoritmo de Bellman-Ford. Neste exemplo, o vértice origem, do qual se deseja obter os caminhos mínimos, é o vértice v. Cada processo do algoritmo é indicado pela iteração i (Tabela 2). Em cada iteração, os valores das distâncias são alterados apenas se for menor do que o valor anterior.

Tabela 2: Iterações do algoritmo de Bellman-Ford

Iteração	**Aresta Relaxada**	**Distâncias Atualizadas**	**Conjunto de Distâncias (dist)**	**Observações**
$i = 0$	-	-	$dist[v] = 0, dist[z] = \infty,$ $dist[u] = \infty, dist[y] = \infty,$ $dist[w] = \infty, dist[x] = \infty$	Inicialização: a distância para o vértice origem v é 0; todas as outras são ∞.
$i = 1$	$(v, z),$ (v, u)	$dist[z] = 1,$ $dist[u] = 2$	$dist[v] = 0, dist[z] = 1,$ $dist[u] = 2, dist[y] = \infty,$ $dist[w] = \infty, dist[x] = \infty$	As distâncias de z e u são atualizadas a partir de v.
$i = 2$	$(z, y),$ (u, w)	$dist[y] = 5,$ $dist[w] = 3$	$dist[v] = 0, dist[z] = 1,$ $dist[u] = 2, dist[y] = 5,$ $dist[w] = 3, dist[x] = \infty$	A distância de y é atualizada a partir de z e a de w, a partir de u.
$i = 3$	(y, x)	$dist[x] = 6$	$dist[v] = 0, dist[z] = 1,$ $dist[u] = 2, dist[y] = 5,$ $dist[w] = 3, dist[x] = 6$	A distância de x é atualizada a partir de y.
$i = 4$	-	-	$dist[v] = 0, dist[z] = 1,$ $dist[u] = 2, dist[y] = 5,$ $dist[w] = 3, dist[x] = 6$	Todas as arestas foram relaxadas. As distâncias mínimas foram calculadas.

Em $i = 0$, a distância para a origem v é 0, e para todos os outros vértices é infinita. Na primeira iteração, as distâncias de z e u são atualizadas a partir de v, para 1 e 2, respectivamente. Na segunda iteração, a distância de y é atualizada a partir de z e a distância de w é atualizada a partir de u, para 5 e 3, respectivamente. Na terceira iteração, as distâncias são revisadas e a distância de x é atualizada a partir de y para 6. O resultado apresenta a distância mínima da origem A para todos os outros vértices: $d(v,z) = 1,\ d(v,y) = 5,\ d(v,u) = 2,$ $d(v,w) = 3,\ d(v,x) = 6.$

2.2.3 Prim [1957]:

O algoritmo de Prim é um algoritmo para encontrar a árvore geradora mínima de um grafo não direcionado e ponderado. A árvore geradora mínima é um subgrafo que conecta todos os vértices do grafo original com o menor custo total possível. O pseudocódigo do algoritmo é apresentado a seguir.

Algoritmo de Prim

Entrada: $G = (V, E, c)$, onde G é um grafo não direcionado com custos não negativos.

Passo 1: Inicialização

Escolha qualquer vértice $i \in V$.

$AGM = \{ i \}$ // Inicializa o conjunto AGM com o vértice inicial i.

$N = V \setminus \{i\}$ // Conjunto de vértices não processados.

$T_{min} = \emptyset$ // Inicializa o conjunto das arestas da árvore geradora mínima.

Passo 2: Construção da AGM

Enquanto $AGM \neq V$ faça:

Encontrar a aresta $(j,k) \in E\ tal\ que\ j \in AGM, k \in N$, e $c(j,k)$ é o mínimo.

$AGM = AGM \cup \{k\}$ // Adicionar o vértice k ao conjunto AGM.

$N = N \setminus \{k\}$ // Remover o vértice k do conjunto N.

$T_{min} \cup (j,k)$ // Adicionar a aresta (j,k) ao conjunto T_{min}

Saída: Uma árvore geradora mínima de V.

Fonte: Adaptado Goldbarg e Goldbarg [2012]

Na Figura 2.7 apresentamos um exemplo de aplicação do algoritmo de Prim. Neste exemplo, o vértice origem, a partir do qual se deseja iniciar a árvore

geradora mínima, é o vértice v. Na Tabela 3 demonstramos as iterações do algoritmo.

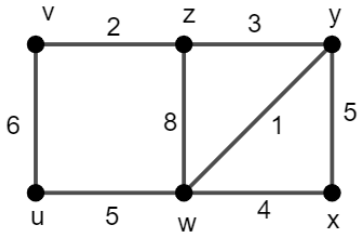


Figura 2.7: Exemplo para aplicação do algoritmo de Prim. Fonte: a autora.

Tabela 3: Iterações do algoritmo de Prim

Iteração	**Aresta Escolhida** (j, k)	**Vértice Adicionado** (k)	**Conjunto** AGM	**Conjunto** N	**Conjunto** T_{min}	**Observações**
$i = 0$	-	v	$\{v\}$	$V \setminus \{v\}$	$\emptyset$	Inicialização: o vértice v é escolhido como a raiz da AGM.
$i = 1$	(v, z)	z	$\{v, z\}$	$V \setminus \{v, z\}$	$\{(v, z)\}$	Aresta (v, z) é escolhida, pois tem o menor custo $c(v, z) = 2$
$i = 2$	(z, y)	y	$\{v, z, y\}$	$V \setminus \{v, z, y\}$	$\{(v, z), (z, y)\}$	Aresta (z, y) é escolhida com o menor custo $c(z, y) = 3$
$i = 3$	(y, w)	W	$\{v, z, y, w\}$	$V \setminus \{v, z, y, w\}$	$\{(v, z), (z, y), (y, w)\}$	Aresta (y, w) é escolhida com o menor custo $c(y, w) = 1$
$i = 4$	(y, x)	x	$\{v, z, y, w, x\}$	$V \setminus \{v, z, y, w, x\}$	$\{(v, z), (z, y), (y, w), (y, x)\}$	Aresta (y, x) é escolhida com o menor custo $c(y, x) = 4$
$i = 5$	(w, u)	u	V	$\emptyset$	$\{(v, z), (z, y), (y, w), (y, x), (w, u)\}$	Aresta (w, u) é escolhida com o menor custo $c(w, u) = 5$ completando a AGM.

O vértice v é escolhido como a raiz da AGM. Na primeira iteração, o vértice z, com o menor custo (2), é adicionado à AGM. Na segunda iteração, o vértice y, com o menor custo (3), é adicionado à AGM. Na terceira iteração, o vértice w, com o menor custo (1), é adicionado à AGM. Na quarta iteração, o vértice x, com o menor custo (4), é adicionado à AGM. Na quinta iteração, o vértice u, com o menor custo (5), é adicionado à AGM, completando a árvore geradora mínima. O resultado é uma árvore geradora mínima do grafo com os seguintes custos das arestas: $d(v\,,z) = 2,\ d(z,y) = 3,\ d(y,w) = 1,\ d(w,u) = 5,\ d(y,x) = 4$.

3 MEDIDAS DE CENTRALIDADE:

Neste capítulo, exploraremos algumas medidas de centralidade utilizadas para analisar a importância relativa de elementos em um grafo. A centralidade é um conceito fundamental na Teoria dos Grafos, pois permite identificar os vértices e arestas mais influentes ou estrategicamente posicionados no grafo. Apresentaremos a fundamentação necessária para compreensão dos conceitos de centralidade de vértices na Seção 3.1, e em seguida de arestas na Seção 3.2. Finalmente, na Seção 3.3, estabelecemos uma discussão inicial sobre a relação entre ambas as medidas, utilizando as respectivas centralidades de intermediação.

3.1 Centralidade de vértices:

No estudo de redes, é altamente relevante identificar componentes-chaves e caracterizá-los conforme a sua importância para a estrutura e dinâmica das interações. Para isso, existem diferentes conceitos e métricas focados nos vértices, como por exemplo as medidas de centralidades, que caracterizam o papel dos vértices na estrutura e dinâmica de uma rede. Neste trabalho, abordaremos as seguintes medidas de centralidades: Grau [Shaw 1964]; Proximidade [Sabidussi 1966]; e Intermediação [Freeman, 1978].

Dado um grafo $G = (V, E)$, a *centralidade de grau* de um vértice $v \in V$, $C_g(v)$, define como importante um vértice que está conectado com muitos outros vértices. Assim, a centralidade de grau de um vértice v, é dada por $C_g(v) = d(v)$ [Shaw 1964].

No caso de G ser conexo, a *centralidade de proximidade* considera como mais importante um vértice que está mais próximo dos demais vértices do grafo, ou seja, o vértice mais central é o que possui menor distância total para todos os outros vértices da rede. A centralidade de proximidade de um vértice v, denotada por $C_p(v)$, é dada pela Equação (1), onde $d(v, i)$ é a distância entre os vértices v e i [Sabidussi 1966]. Essa soma de distâncias é chamada de transmissão do vértice e será denotada como $T(v)$.

$$C_p(v) = \frac{1}{T(v)} = \frac{1}{\sum_{\substack{i \in V, \\ i \neq v}} d(v,i)} \quad , \tag{1}$$

A *centralidade de intermediação* para um grafo conexo, considera como mais importante um vértice que está frequentemente presente em caminhos mínimos. Dessa forma, a centralidade de intermediação de um vértice v mede com que frequência este vértice está no caminho mais curto entre outros pares de vértices i e j na rede. Para calcular a centralidade de intermediação de um vértice v, denotada por $C_I(v)$, utiliza-se a Equação (2) [Freeman, 1978]:

$$C_I\,(v) = \sum_{i \neq j \neq v} \frac{\sigma_{ij}(v)}{\sigma_{ij}}, \tag{2}$$

onde σ_{ij} é o número total de caminhos mínimos entre os vértices i e j, e $\sigma_{ij}(v)$ é o número de caminhos mínimos entre os vértices i e j que passam pelo vértice v.

Na Figura 3.1 apresentamos um grafo, a partir do qual estas medidas de centralidades de vértices foram calculadas. Na Tabela 4 apresentamos os valores obtidos a partir das Equações (1) e (2). É possível notar que os vértices pendentes possuem centralidade zero na medida de intermediação, pois, estes vértices não estão no caminho mais curto entre outros pares de vértices i e j da rede. Além disso, percebe-se que os vértices z e y possuem valores iguais de medida de proximidade pois estão à mesma distância total dos demais vértices do grafo. Por fim, observamos que por mais que as medidas sejam calculadas de forma diferente, neste grafo, as três apontam como mais central o vértice w.

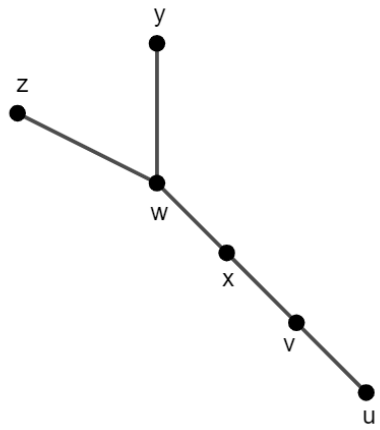


Figura 3.1: Exemplo para cálculo das medidas de centralidades apresentadas

Tabela 4: Valores de centralidade obtidos

vértices	C_g	C_p	C_I
w	*3*	$\frac{1}{8} \cong 0.12$	*7*
x	*2*	$\frac{1}{8} \cong 0.12$	*6*
v	*2*	$\frac{1}{10} \cong 0.10$	*4*
z	*1*	$\frac{1}{12} \cong 0.08$	*0*
y	*1*	$\frac{1}{12} \cong 0.08$	*0*
u	*1*	$\frac{1}{14} \cong 0.07$	*0*

3.2 Centralidade de Arestas:

A centralidade de aresta foi derivada do conceito de centralidade de intermediação de vértices, adaptando-o para avaliar a importância das arestas. Aparentemente, a primeira vez que o conceito de centralidade de aresta foi abordado foi no trabalho "*Community structure in social and biological networks*" realizado por Girvan e Newman, 2001. Este estudo buscava explorar a estrutura de comunidades em redes sociais e biológicas. Em geral, uma comunidade é um grupo de vértices que possuem muitas conexões entre si, mas poucas conexões com outros grupos. Entender essa estrutura é crucial para diversas aplicações, como a análise de redes sociais e biológicas [Girvan e Newman, 2001].

Girvan e Newman [2001] adaptaram a ideia de centralidade de intermediação de vértices para o estudo de arestas mais centrais em um grafo, considerando uma aresta mais central caso esteja presente mais frequentemente em caminhos mínimos entre pares de vértices do grafo. O procedimento para obtenção da *centralidade de intermediação de uma aresta* e, denotada por $C_{a_I}(e)$, em um grafo conexo, é análogo com o procedimento em vértices, sendo determinada pela Equação (3).

$$C_{a_I}(e) = \sum_{i \neq j} \frac{q_{ij}(e)}{q_{ij}}, \quad (3)$$

onde $i, j \in V(G)$, $q_{ij}(e)$ é o número de caminhos mínimos com extremos nos vértices i e j que passam pela aresta e, e q_{ij} é o número total de caminhos mínimos com extremos nos vértices i e j.

Diferentemente da centralidade de intermediação de vértices, em que alguns vértices podem apresentar centralidade zero, todas as arestas em um grafo conexo possuem centralidade positiva. Isso ocorre porque a centralidade de intermediação de vértices contabiliza a frequência com que um vértice está em caminhos mínimos entre pares de vértices; um vértice que não faz parte de nenhum caminho mínimo terá centralidade zero. Por exemplo, vértices folhas em uma árvore não participam dos caminhos mínimos entre outros pares de vértices, resultando em centralidade nula.

Por outro lado, a centralidade de intermediação de arestas considera os caminhos mínimos que passam por cada aresta. Em um grafo conexo, toda aresta conecta diretamente dois vértices adjacentes, de modo que o caminho mínimo entre esses dois vértices necessariamente inclui essa aresta. Assim, cada aresta é utilizada por pelo menos um caminho mínimo, garantindo que sua centralidade seja sempre maior que zero.

Segundo Girvan e Newman [2001], a análise da centralidade de arestas tem um impacto significativo na detecção de comunidades. Em uma rede que contém comunidades, os grupos de vértices (comunidades) são, em geral, ligados por um número reduzido de arestas. Todos os caminhos mais curtos que conectam vértices de comunidades distintas devem atravessar essas poucas arestas. Assim, essas arestas terão uma alta centralidade de intermediação, pois são as únicas rotas diretas entre as diferentes comunidades.

Uma outra medida de centralidade que esta pesquisa irá utilizar é a *centralidade de proximidade de aresta*. Bröhl e Lehnertz [2019] definiram a centralidade de proximidade de aresta como uma medida que avalia o quão central uma aresta é com base na soma das distâncias entre os vértices que ela conecta e todos os outros vértices da rede. Para calcular a centralidade de uma aresta e cujos extremos são os vértices a e b, utiliza-se a fórmula que leva em conta as centralidades de proximidade dos vértices a e b, representadas como

$\mathrm{C_p}(a)$ e $C_p(b)$, onde $e \in E(G)$. Formalmente, a centralidade de proximidade de uma aresta é definida pela Equação (4):

$$C_{aP}(e) = \frac{1}{T(a)+T(b)} = \frac{1}{\left(\frac{1}{\mathrm{C_p}(a)}+\frac{1}{\mathrm{C_p}(b)}\right)} = \frac{C_p(a) * C_p(b)}{C_p(a)+C_p(b)}, \quad (4)$$

Na Figura 3.2 ilustramos um grafo, a partir do qual as medidas de centralidade de intermediação e de proximidade de aresta foram calculadas para todas as arestas, e apresentada na Tabela 5. É possível notar que as arestas associadas aos vértices pendentes, como as arestas (w, y), (w, z), e (u, v), possuem valores mais baixos em ambas medidas de centralidade de aresta. Isso ocorre porque essas arestas não estão presentes em muitos caminhos mais curtos entre pares de vértices. Por outro lado, a aresta (x, w) destaca-se com o maior valor de centralidade de intermediação de aresta, 9, e maior valor de centralidade de proximidade de aresta, 0.062, indicando que essa aresta é crucial para a conexão de diferentes partes do grafo. É possível notar que o vértice w é o mais central e está presente na aresta mais central, conforme mostram as Tabelas 4 e 5. Porém, nem sempre isso ocorrerá, como será mostrado na Seção 3.3.

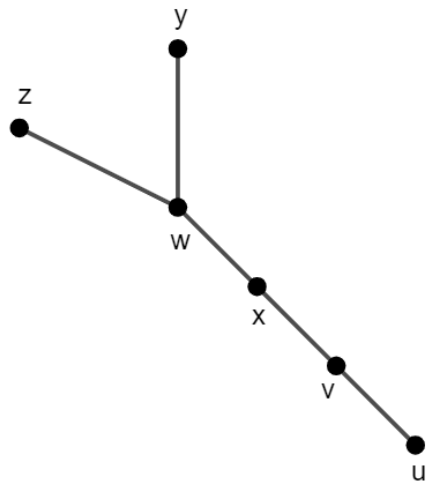


Figura 3.2: Exemplo para cálculo da medida de centralidade de aresta

Tabela 5: Valores obtidos para centralidade de intermediação e proximidade de aresta

Arestas	C_{a_I}	C_{a_P}
(x, w)	*9*	$\frac{1}{16} \cong 0.062$
(v, x)	*8*	$\frac{1}{18} \cong 0.055$
(w, z)	*5*	$\frac{1}{20} \cong 0.050$
(w y)	*5*	$\frac{1}{20} \cong 0.050$
(u, v)	*5*	$\frac{1}{24} \cong 0.041$

Posteriormente, outros estudos também propuseram novas medidas de centralidade de aresta: Kucharczuk *et* al. [2022] desenvolveram o *"PageRank for*

Edges"; Mocanu *et* al. [2018] criaram o método *"Game of Thieves (GOT");* e De Meo *et* al. [2012] introduziram a "*k-path edge centrality*". Estas não serão abordadas nesta dissertação.

As medidas de centralidade de arestas apresentadas nesta seção serão utilizadas na sua forma normalizada no Capítulo 5. As equações normalizadas serão apresentadas na Seção 5.3.

3.3 Relação entre Centralidade de intermediação de vértices e centralidade de intermediação de arestas:

Esta seção se concentra na discussão sobre a centralidade de arestas e de vértices de alguns grafos conexos e, particularmente, de árvores, visando responder a seguinte questão: “As arestas mais centrais são associadas aos vértices mais centrais em um grafo? “. Utilizaremos a medida de centralidade de intermediação como objeto dessa investigação.

Para essa discussão, todos os grafos simples conexos e não isomorfos com n vértices, onde $6 \leq n \leq 8$, e todas as árvores com n vértices, onde $6 \leq n \leq 16$, foram gerados e suas respectivas centralidades de intermediação arestas e de vértices. Nesta dissertação, todos os grafos foram gerados e todos os testes foram executados através da linguagem Python, utilizando a biblioteca NetworkX. Para gerar os grafos conexos e não isomorfos com $n \geq 9$, dificuldades de execução foram apresentadas, não sendo possível a conclusão devido a erros relacionados à falta de memória. Para árvores, o mesmo problema ocorreu para $n \geq 17$. Os grafos foram gerados em um notebook com as seguintes especificações técnicas: processador Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11GHz e 8 GB de RAM instalada. Esses fatores de hardware podem ter influenciado as limitações observadas.

No contexto da análise dos resultados, utilizamos o termo *empate* quando dois ou mais elementos de um grafo apresentam exatamente o mesmo valor de centralidade, seja no caso de vértices ou aresta, para uma dada medida. Isso significa que, ao calcular uma determinada centralidade, se dois ou mais

elementos tiverem o mesmo valor atribuído, eles são considerados igualmente centrais. A partir disso, dois subconjuntos foram definidos para cada grafo:

- **Arestas com maior centralidade:** aquelas que atingem o valor máximo de centralidade de arestas (podendo ser uma ou várias em caso de empate);
- **Vértices com maiores valores de centralidades:** definidos como os dois vértices com maiores valores de centralidade de vértices. Em caso de empate no segundo maior valor, todos os vértices com valor igual ao segundo maior também são incluídos (isto é, a lista pode conter mais de dois vértices).

Observe que enquanto o conjunto de arestas mais centrais pode ser formado por apenas um elemento, o mesmo não ocorre para o conjunto de vértices mais centrais, sendo formado por no mínimo dois elementos. Este fato está relacionado a questão central proposta inicialmente, isto é, se as arestas mais centrais são associadas aos vértices mais centrais em um grafo, e ao fato de uma aresta ser determinada por dois vértices.

Com o objetivo de investigar a relação entre as arestas com maior centralidade e os vértices com maior centralidade nos grafos, partimos de uma pergunta central: quando uma aresta é considerada mais central, seus vértices também tendem a ser os considerados mais centrais do grafo? Para explorar essa questão, propomos dois contextos de análise: o caso restrito e o caso amplo. A principal distinção entre os casos está na presença ou ausência de empates entre os valores máximos de centralidade de arestas. Nos casos restritos, há apenas uma aresta com o valor máximo de centralidade — ou seja, não há empate, o que significa que existe uma única aresta mais central no grafo. Já nos casos amplos, há empate entre duas ou mais arestas com esse valor máximo, o que caracteriza a existência de múltiplas arestas igualmente mais centrais. Em ambos os casos, considera-se que existe apenas um valor máximo de centralidade de aresta; o que os diferencia é se esse valor é único ou compartilhado entre várias arestas. A partir desses dois casos principais, propomos ainda subdivisões específicas para identificar de forma mais direta a

relação da(as) aresta(s) com valor máximo de centralidade com os vértices que atingem os valores máximos de centralidade do grafo.

Para isso definimos quatro configurações exclusivas de análise, denominadas **restrito 1-incidente**, **restrito 2-incidente**, **amplo 1-incidente** e **amplo 2-incidente.**

- **Restrito 2-incidente:**

O grafo é classificado nesta configuração quando existe apenas uma única aresta com valor máximo de centralidade e exatamente dois vértices com maior centralidade de vértices, sendo ambos os vértices extremos dessa aresta. Este é o caso mais estrito e específico.

- **Restrito 1-incidente**:

Esta configuração representa uma situação intermediária: há apenas uma aresta mais central e exatamente dois vértices mais centrais, mas apenas um dos vértices incidentes à aresta mais central está presente entre esses dois vértices. Isso indica uma associação parcial entre centralidade de arestas e vértices.

- **Amplo 2-incidente**:

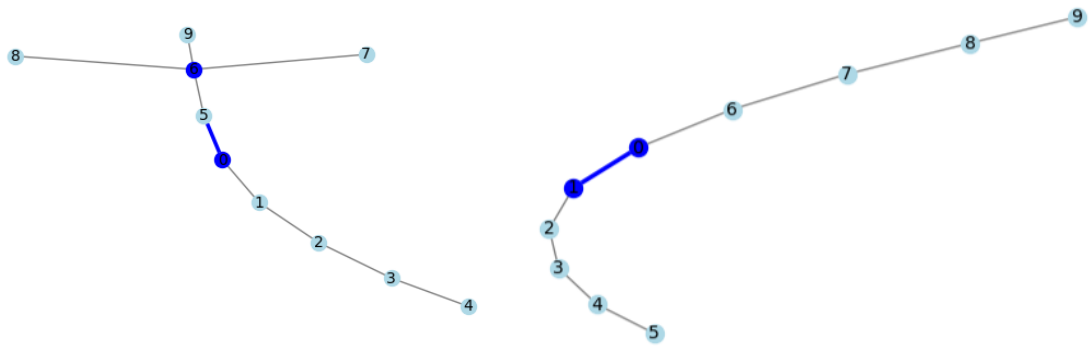
O grafo pertence a esta configuração quando todas as arestas com centralidade máxima (duas ou mais) têm ambos os extremos entre os vértices mais centrais. Aqui, aceita-se a existência de empates (ou seja, pode haver mais de dois vértices centrais), e todos os vértices incidentes nas arestas mais centrais devem estar completamente contidos dentro desse conjunto. Aqui se exige que os dois extremos de cada aresta estejam entre os vértices mais centrais.

- **Amplo 1-incidente**:

Esta configuração inclui os grafos em que pelo menos uma das arestas com maior centralidade (duas ou mais) têm apenas um dos vértices incidentes presente entre os vértices mais centrais (considerando empates).

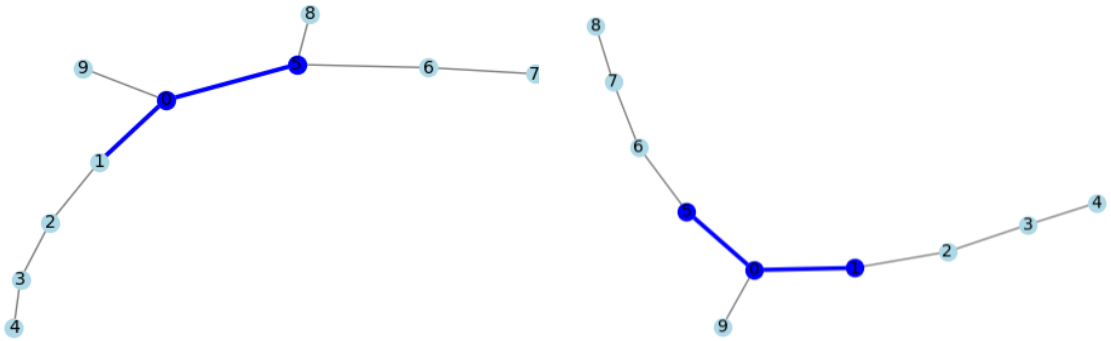
Na Figura 3.3, apresentamos um exemplo para cada configuração com destaque em azul escuro para os vértices e arestas que atingem os valores máximos de centralidade de intermediação, considerando empates.

No Figura 3.3 (a), classificado como *restrito 1-incidente*, há apenas uma aresta mais central (5,0), e dois vértices com centralidade máxima (vértices 0 e 6), dos quais apenas o vértice 0 é incidente à aresta central. Já no grafo (b), classificado como *restrito 2-incidente*, há também uma única aresta mais central (1,0) e dois vértices com maior centralidade (vértices 1 e 0), ambos incidentes na aresta mais central. No grafo (c), classificado como *amplo 1-incidente*, duas arestas compartilham o valor máximo de centralidade (1,0) e (0,5). Embora a aresta (0,5) tenha como os dois vértices incidentes os vértices com maior centralidade do grafo (vértices 0 e 5), a aresta (1,0) só possui o vértice central 0 como incidente, por isso, este grafo se enquadra na configuração *amplo 1-incidente.* Por fim, no grafo (d), classificado como *amplo 2-incidente*, também há mais de uma aresta mais central (1,0) e (0,5), porém temos três vértices que compartilham o valor máximo de centralidade (vértices 1,0 e 5), sendo estes os incidentes às duas arestas mais centrais.



(a) Restrito 1-incidente

(b) Restrito 2-incidente



(c) Amplo 1-incidente (d) Amplo 2-incidente

Figura 3.3 Exemplos de grafos para cada configuração analisada

Um ponto importante a destacar é que para fins de código, vértices com centralidade igual à 0 não são considerados. Portanto, grafos onde há somente um vértice com valor máximo e todos os outros possuem centralidade igual à 0 estão automaticamente fora das categorias analisadas, como por exemplo o grafo ilustrado na Figura 3.4.

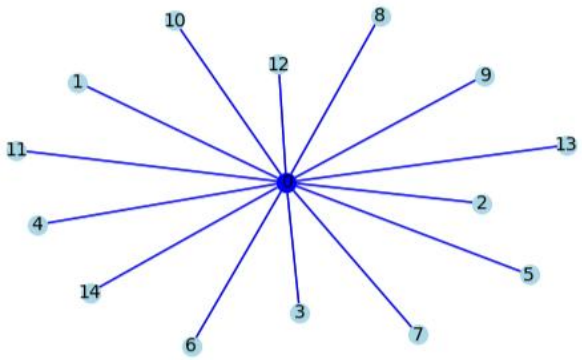


Figura 3.4: Exemplo de grafo considerado fora das categorias analisadas

Na Tabela 6 apresentamos, para os grafos simples com $6 \leq n \leq 8$, a quantidade de grafos classificados em cada uma das quatro configurações, permitindo uma avaliação da frequência relativa com que ocorrem associações fortes, parciais ou inexistentes entre vértices e arestas mais centrais. A mesma análise é realizada na Tabela 7 para árvores com $6 \leq n \leq 16$. Na última linha das Tabelas 6 e 7, por frequência relativa total nos referimos ao valor de

frequência relativa quando analisamos conjuntamente todos os grafos com $6 \leq n \leq 8$ e $6 \leq n \leq 16$, respectivamente.

Tabela 6: Resultado da análise de associações entre centralidade de intermediação de vértices e arestas nos grafos simples com $6 \leq n \leq 8$.

n	Nº de grafos conexos	Configuração restrita				Configuração ampla				Soma das frequências
		1-Incidente		2-Incidente		1-Incidente		2-Incidente		
		Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	
6	**112**	29	26%	16	14%	53	47%	10	9%	**96%**
7	**853**	348	41%	102	12%	318	38%	35	4%	**95%**
8	**11117**	6345	57%	1085	10%	2631	24%	222	2%	**93%**
Freq. Relativa Total		**55,64%**		**9,96%**		**24,85%**		**2,21%**		**92,66%**

Tabela 7: Resultado da análise de associações entre centralidade de intermediação de vértices e arestas nas árvores com $6 \leq n \leq 16$

n	Nº de árvores	Configuração restrita				Configuração ampla				Soma das frequências
		1-Incidente		2-Incidente		1-Incidente		2 Incidente		
		Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	Freq. Absoluta	Freq. Relativa	
6	**6**	0	0%	4	67%	0	0%	1	17%	**83%**
7	**11**	0	0%	5	45%	1	9%	4	36%	**91%**
8	**23**	0	0%	17	74%	1	4%	4	17%	**96%**
9	**47**	0	0%	27	57%	9	19%	10	21%	**97%**
10	**106**	9	8%	72	68%	9	8%	14	13%	**98%**
11	**235**	0	0%	152	65%	50	21%	32	14%	**100%**
12	**551**	75	14%	354	64%	56	10%	50	9%	**97%**
13	**1301**	114	9%	786	60%	272	21%	103	8%	**98%**
14	**3159**	531	17%	2047	65%	341	11%	173	5%	**98%**
15	**7741**	970	13%	4576	59%	1578	20%	392	5%	**97%**
16	**19320**	3969	21%	11850	61%	2131	11%	808	4%	**97%**
Freq. Relativa Total		**17,47%**		**61,29%**		**13,71%**		**4,90%**		**97,36%**

Tanto na Tabela 6 quanto na Tabela 7, observamos que há forte relação entre aresta mais central e vértice mais central. Quando analisamos a soma dos

percentuais em que ocorrem as configurações, os valores variam de 83% a 100%. Ou seja, para estes grafos e árvores analisados em pelo menos 83% deles temos que há incidência de pelo menos um vértice mais central a todas as arestas mais centrais.

Mais especificamente, na Tabela 6, analisamos os grafos com $n = 6, 7$ e 8 apenas, devido a limitação técnica. Nesta pequena amostra, temos que quando $n = 6$ predomina a configuração ampla 1-incidente, representando quase metade dos grafos conexos gerados. Já para n =7 e $n = 8$, observa-se a predominância da configuração restrita 1-incidente. Para estes grafos da Tabela 6, poucos são os casos em que há 2-incidentes, independentemente de ser restrito ou amplo (frequência relativa total de 9,96% e 2,21%, respectivamente).

Na Tabela 7, analisamos apenas árvores, considerando valores de n de 6 até 16. Para os menores valores de n (de 6 a 8), observa-se que a configuração restrita 2-incidente já começa a se destacar, com 67% em $n = 6$, 45% em $n = 7$ e chegando a 74% em $n = 8$. A partir de $n = 9$, essa configuração se consolida como a mais frequente, com percentuais sempre acima de 57%. Ou seja, nas árvores analisadas, a configuração em que há exatamente uma aresta mais central e dois vértices mais centrais incidentes a ela torna-se predominante conforme o número de vértices aumenta.

Por outro lado, a configuração restrita 1-incidente é menos comum nas árvores menores, não aparecendo nos casos de $n = 6$ a $n = 8$, mas se torna mais presente a partir de $n = 10$, chegando a 21% em $n = 16$. Isso indica um crescimento gradual da ocorrência de situações em que apenas um dos vértices mais centrais incide sobre a aresta mais central.

As configurações amplas, por sua vez, aparecem com menor frequência relativa total, com 13,71% para 1-incidente e 4,90% para 2-incidente. Além disso, a ocorrência da 2-incidente tende a decrescer ligeiramente com o aumento de n.

No geral, os dados da Tabela 7 demonstram que em grande parte das árvores, há associação entre vértices e arestas mais centrais. A soma dos percentuais mostra que essa associação ocorre em pelo menos 83% dos casos ($n = 6$) e chega a 100% (em $n = 11$), com frequência total de 97,63%. Isso indica

que, mesmo com o aumento do número de vértices, permanece forte a associação entre os elementos centrais da estrutura — sugerindo um comportamento estrutural recorrente nas árvores analisadas.

Na Figura 3.5, têm-se dois exemplos em grafos com 6 vértices de quando a configuração restrita 2- incidente ocorre, ou seja, há apenas uma aresta mais central, e esta é composta pelos dois vértices mais centrais. Os vértices em azul escuro são os dois mais centrais, e a aresta em azul escuro é a mais central. No exemplo (a), a aresta (3,4) possui uma centralidade de intermediação de 0.4, e os vértices mais centrais são os vértices 4 e 3 com valores de centralidade de 0.7 e 0.4, respectivamente. No exemplo (b), a aresta mais central é a (0,3) com valor de centralidade de intermediação de 0.4, e os vértices mais centrais são os 0 e 3 com valores de centralidade de 0.6 e 0.4, respectivamente.

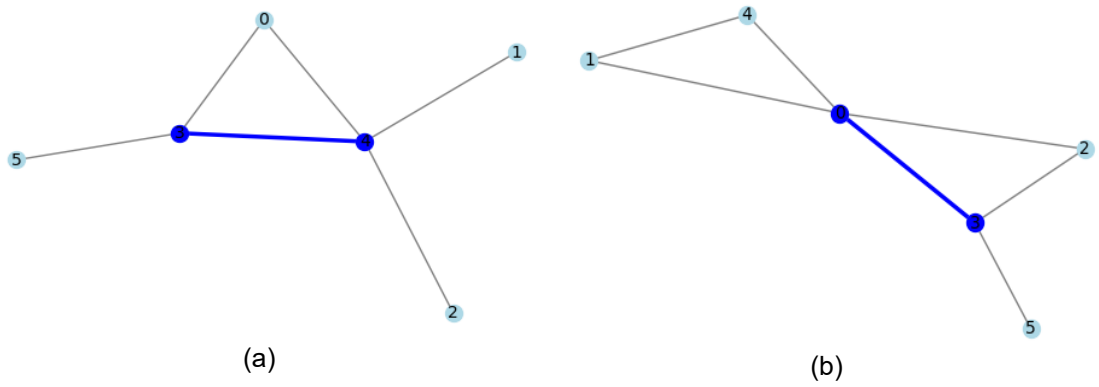
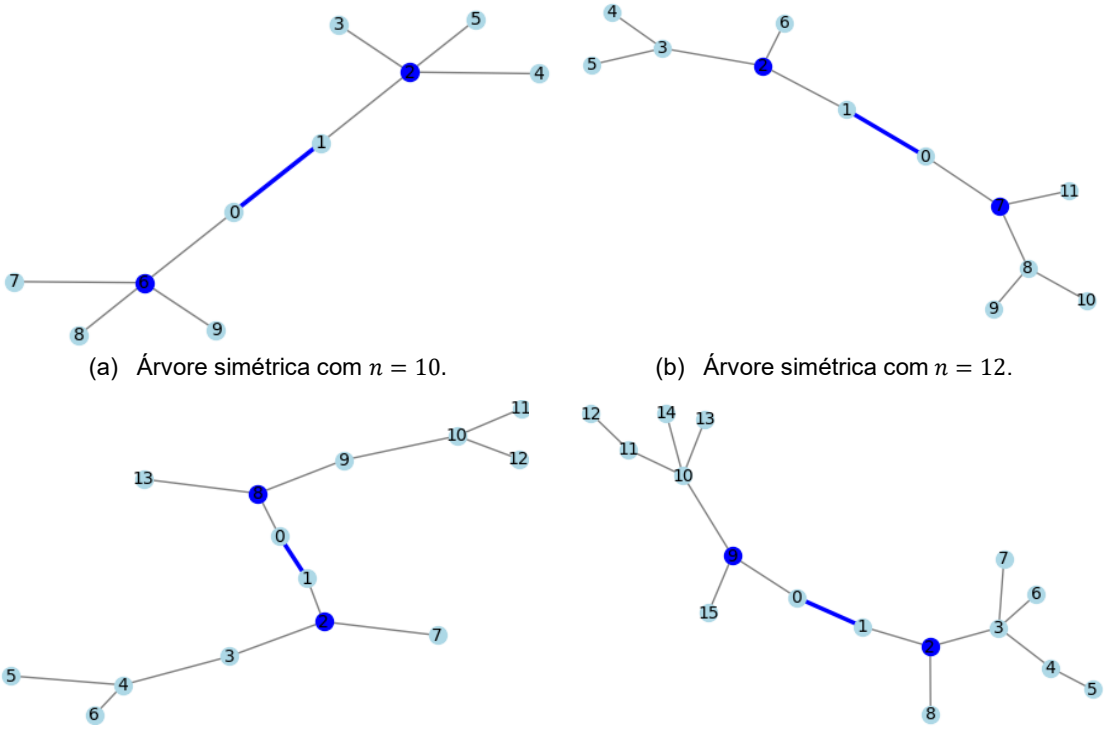


Figura 3.5:Exemplos de casos restritos em grafos simples com 6 vértices.

Alguns grafos não se enquadraram nas configurações estabelecidas, ou seja, alguma das arestas mais centrais não possuem nenhum dos vértices considerados mais centrais. Na Figura 3.6 apresentamos exemplos para os valores de $n = 10, 12, 14, 16$ onde os vértices presentes na única aresta mais central não são os dois vértices mais centrais do grafo. É possível notar, em todos os casos, que as árvores possuem uma simetria com relação a aresta mais central.



(a) Árvore simétrica com $n = 10$.

(b) Árvore simétrica com $n = 12$.

(c) Árvore simétrica com $n = 14$.

(d) Árvore simétrica com $n = 16$.

Figura 3.6:Exemplo de discordância entre centralidade de aresta de intermediação e de vértice de intermediação

Dentre esses grafos que não apresentaram associação entre vértices e arestas mais centrais, foi possível identificar algumas vassouras duplas equilibradas, cuja estrutura chama atenção justamente por sua simetria em torno da aresta mais central. Essa observação motivou o estudo apresentado no Capítulo 4, que é aprofundado nas centralidades de vértices e arestas em grafos vassoura dupla equilibradas, com o objetivo de compreender as associações entre vértices e arestas nesse tipo de grafo. As medidas de centralidade de intermediação de vértices e arestas serviram como ponto de partida, e possibilitou a identificação das vassouras duplas equilibradas. Adicionalmente, no Capítulo 4 resolvemos também analisar as medidas de proximidade de vértices e arestas.

4 RELAÇÃO ENTRE CENTRALIDADES DE VÉRTICES E ARESTAS EM VASSOURAS DUPLAS EQUILIBRADAS

Neste capítulo, analisaremos a relação entre as centralidades de vértices e arestas em um tipo específico de grafo denominado vassoura dupla equilibrada. Na Seção 4.1, apresentamos a definição formal de uma vassoura dupla equilibrada. Em seguida, nas Seções 4.2 e 4.3, investigamos duas medidas clássicas de centralidade nas vassouras duplas equilibradas: a centralidade de intermediação e a centralidade de proximidade, respectivamente para vértices e arestas.

4.1 Vassoura Dupla:

Um grafo *vassoura dupla* (em inglês, *double broom graph*) é uma árvore de n vértices obtida a partir do caminho $P_e,\ e \geq 2$, conectando $p \geq 2$ vértices a um dos vértices pendentes deste caminho e $s \geq 2$ vértices ao outro vértice pendente deste caminho. Em particular, neste capítulo trataremos dos casos em que $p = s$, e, estas árvores serão denominadas vassouras duplas equilibradas.

Dessa forma, uma árvore vassoura dupla equilibrada será composta por $n = e + 2p$ vértices. Embora existam diversas formas de rotular e representar graficamente uma vassoura dupla equilibrada, durante este Capítulo assumiremos que todas são rotuladas e representadas como na Figura 4.1.

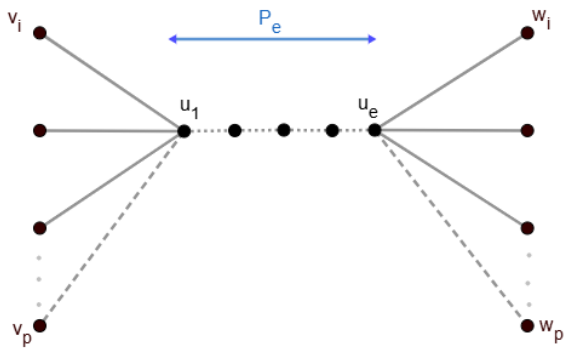


Figura 4.1: Vassoura dupla equilibrada

Os vértices $u_1, u_2, \ldots, u_e$ que compõem o caminho P_e deste grafo serão denominados vértices da espinha do grafo, enquanto todos os demais $2p$

vértices, v_i e w_i, $1 \leq i \leq p$ são vértices pendentes, conectados, respectivamente, com u_1 e u_e, os extremos do caminho P_e.

Ao longo do capítulo, também tendo como referência a Figura 4.1, usaremos, por simplicidade, os termos “esquerda” e “direita” a um vértice no seguinte contexto. Dado um vértice u_k da espinha $(u_1, u_2, \dots, u_e)$:

- Esquerda de u_k: todos os vértices na espinha que estão antes de u_k, ou seja, $u_1, \dots, u_{\{k-1\}}$, e todos os pendentes conectados a u_1.
- Direita de u_k: todos os vértices na espinha que estão depois de u_k, ou seja, $u_{\{u+1\}}, ., u_{\{e\}}$, e todos os pendentes conectados a u_e.

Por exemplo para u_1 temos que:

- Esquerda: os p vértices pendentes $v_1, v_2, \dots, v_p$.
- Direita: todos os outros vértices $u_2, u_3, \dots, u_e$, mais os pendentes $w_1, w_2, \dots\ w_p$.

4.2 Centralidade de intermediação de vértices:

Nesta seção, analisamos a centralidade de intermediação dos vértices e arestas nos grafos denominados vassouras duplas equilibradas, com o objetivo de identificar padrões que emergem à medida que se aumenta simetricamente a árvore. Encontraremos expressões matemáticas para quantificar a centralidade dos vértices e identificaremos os vértices e as arestas mais centrais.

A ideia é generalizar completamente o resultado para o caso em que o número de vértices na espinha é e e o número de vértices pendentes incidentes a u_1 e u_e é igual a p.

A espinha tem e vértices $(u_1, u_2, \dots, u_e)$. Cada extremo da espinha possui p vértices pendentes: $(v_1, v_2, \dots, v_p)$ de um lado e $(w_1, w_2, \dots, w_p)$ do outro lado. Sendo o total de vértices no grafo: $2p + e$. Buscando identificar as variações em e para dado um n fixo, temos a relação a ser utilizada nos cálculos, $e = n - 2p$.

Ao calcularmos a centralidade de intermediação de vértices, temos que devido à simetria da vassoura dupla equilibrada:

- $C_I(u_1) = C_I(u_e)$ (extremidades da espinha têm a mesma centralidade).

- $C_I(u_k) = C_I(u_{e-k+1})$, $2 \leq k \leq e-1$ (demais vértices simétricos na espinha também têm a mesma centralidade)
- $C_I(v_i) = C_I(w_j) = 0, \quad 1 \leq i, j \leq p$

Dessa forma, a centralidade de intermediação de vértices pode ser calculada por:

- $$C_I(u_1) = C_I(u_e) = \underbrace{p}_{\text{Quantidade de vértices à esquerda de } u_1} * \underbrace{(-p+n-1)}_{\text{Quantidade de vértices à direita de } u_1} + \overbrace{\binom{p}{2}}^{\text{Combinação par a par entre os vértices pendentes}} = \frac{-p^2+2np-3p}{2}, \quad (5)$$

- $$C_I(u_k) = \underbrace{(p+k-1)}_{\text{à esquerda de } u_k} * \underbrace{(-p+n-k)}_{\text{à direita de } u_k}, \text{ para cada } 2 \leq k \leq e-1, \quad (6)$$

Cabe ressaltar que os vértices pendentes possuem centralidade zero na medida de intermediação de vértices, pois, estes vértices não estão no caminho mais curto entre outros pares de vértices i e j no grafo.

Observa-se que a expressão apresentada para a centralidade dos vértices na espinha $C_I(u_k)$, $2 \leq k \leq e-1$, pode ser vista como uma função quadrática em relação a k. Como a função possui um termo quadrático negativo, sua representação gráfica é uma parábola com concavidade voltada para baixo. Assim, para determinar qual valor de k, $2 \leq k \leq e-1$, possui um maior valor de centralidade $C_I(u_k)$, basta determinarmos quando a função quadrática atinge seu maior valor

Uma forma alternativa de determinar esse ponto é calcular os zeros da função, ou seja, os valores de k para os quais $C_I(u_k) = 0$. Como a parábola é simétrica, o valor que maximiza a função pode ser encontrado tirando a média aritmética das raízes. Aplicando esse procedimento, obtemos:

$$(p + k - 1) * (n - p - k) = 0$$

$$(p + k - 1) = 0 \; ou \; (n - p - k) = 0$$

$$k_1 = 1 - p$$

$$k_2 = n - p$$

$$k_v = \frac{[(1 - p) + (n - p)]}{2} = \frac{n - 2p + 1}{2} = \frac{e + 1}{2}$$

No caso em que a espinha tem comprimento ímpar, isto é, e é um número ímpar, o valor k_v indica a posição do vértice ao longo da espinha em que a centralidade de intermediação de vértices atinge seu valor máximo dentro das restrições definidas. No caso em que o comprimento da espinha é par, embora k_v não seja um número inteiro, pela simetria da parábola, é possível identificar em quais vértices ocorrem o valor máximo de centralidade, como será discutido na Seção 4.2.2.

Observamos uma particularidade nos grafos vassoura dupla equilibradas ao comparar as configurações com números pares e ímpares de vértices na espinha. Essa distinção motivou a formulação de duas provas gerais, uma para os casos com número par de vértices e outra para os casos com número ímpar.

Ambas as provas buscam demonstrar padrões específicos na distribuição das centralidades de intermediação dos vértices e das arestas, independentemente do tamanho exato da espinha. Essa abordagem permite generalizar os resultados e destacar as diferenças estruturais fundamentais entre as duas configurações.

4.2.1. Centralidade de intermediação de vértices quando o número de vértices na espinha é ímpar:

Consideramos aqui $n \geq 7$, $e \geq 3$ e $p \geq 2$ (excluímos o caso em que o grafo é caminho ou estrela). É possível identificar que dentre os vértices da espinha, excluindo os extremos, o valor máximo da centralidade do grafo considerado, para qualquer n ímpar e qualquer p, é encontrado no vértice do

centro da espinha $\frac{e+1}{2}$. Dessa forma, a centralidade de intermediação de vértices do vértice do centro da espinha é obtida através da expressão:

$$C_I\left(u_{\frac{e+1}{2}}\right) = \left(p + \frac{e+1}{2} - 1\right) * \left(-p + n - \frac{e+1}{2}\right) = \frac{(n-1)^2}{4} \text{ (utilizando } e = n - 2p\text{)}, \qquad (7)$$

Portanto, buscando identificar os vértices mais centrais quando o número de vértices n está fixado, podemos comparar apenas a expressão da centralidade dos vértices extremos da espinha u_1 e u_e (5) com a expressão da centralidade do vértice central da espinha $\frac{e+1}{2}$ (7).

Para isso, podemos subtrair uma expressão da outra para verificar qual é maior, considerando a diferença:

$$\gamma(p) = \frac{-p^2 + 2np - 3p}{2} - \frac{(n-1)^2}{4}$$

$$\gamma(p) = \frac{1}{4}\left[-2p^2 + (4n-6)p - (n-1)^2\right]$$

Quando $\gamma(p) > 0$, os vértices do extremo da espinha tornam-se mais centrais que o vértice no meio da espinha. Quando $\gamma\ (p) < 0$ a situação oposta ocorre.

As raízes da função $\gamma(p)$ são:

$$p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 7)}\right]$$

$$p_2 = \frac{1}{2}\left[(2n-3) + \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 7)}\right]$$

Dada as condições iniciais de $n \geq 7$ e $p \geq 2$, precisamos verificar a condição para $p_1 \geq 2$.

$$p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 7)}\right] \geq 2$$

$$n \geq 7$$

Como $\gamma(p)$ possui como gráfico uma parábola com concavidade para baixo (pois o coeficiente do termo quadrático é negativo), podemos analisar o sinal da função:

- Para valores de p tais que $p_1 < p < p_2$, temos $\gamma(p) > 0$, indicando que $C_I(u_1) = C_I(u_e)$ é maior. Logo, os vértices dos extremos da espinha assumem o maior valor de centralidade.
- Para valores de p tais que $p < p_1$ ou $p > p_2$, temos $\gamma(p) < 0$, indicando que $C_I\left(u_{\frac{e+1}{2}}\right)$ é maior. Neste caso, o vértice no centro da espinha assume o maior valor de centralidade.

Na Figura 4.2 apresentamos um esboço do gráfico das duas funções de centralidades de vértices analisadas e os zeros de $\gamma(p)$. É possível observar o comportamento da centralidade de vértice. Conforme já mencionado, começamos a análise com $p_1 < p < p_2$, nesse intervalo temos a centralidade do vértice no centro da espinha sendo maior para qualquer $n \geq 7$ e ímpar. Para $p < p_1$ ou $p > p_2$, temos o maior valor de centralidade sendo dos vértices dos extremos da espinha para qualquer $n \geq 7$ e ímpar.

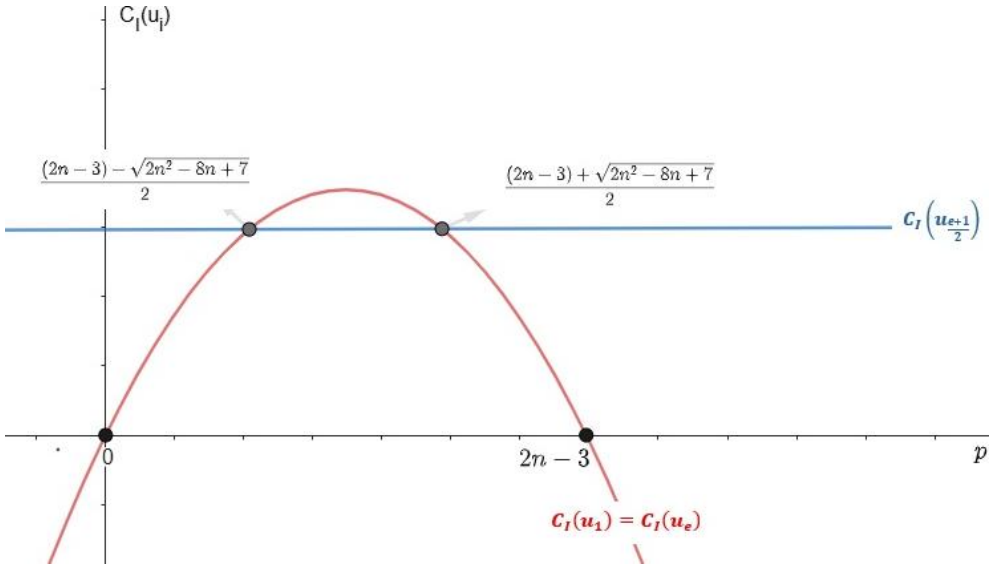


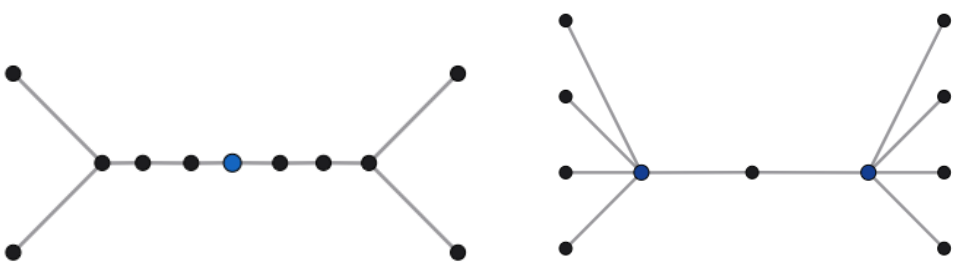
Figura 4.2: funções de centralidades de vértices analisadas e as raízes obtidas para γ

Dessa forma, ao fixarmos n, podemos determinar os valores específicos de p para os quais ocorre a mudança no comportamento da função, auxiliando na caracterização das regiões onde cada expressão domina. Na Tabela 8 apresentamos alguns valores para exemplificação. A partir dela é possível observar para cada n fixado, qual é o valor de p que altera o vértice a ser considerado com a maior centralidade. Por exemplo, para $n = 11$, temos que em $p = 2$ e $p = 3$, $\gamma(p)$ é negativo, o que significa que temos o vértice do centro da espinha com o maior valor de centralidade. A partir de $p = 4$, o sinal de $\gamma(p)$ inverte, e isso significa que a maior centralidade agora é o valor obtido nos vértices extremos da espinha.

Tabela 8: Exemplos de valores obtidos para função γ(p)

n	p	$\gamma(p)$	$e = n - 2p$
11	2	-8	7
	3	-1	5
	4	5	3
13	2	-15	9
	3	-6	7
	4	2	5
15	2	-24	11
	3	-13	9
	4	-3	7
	5	6	5

Ilustramos na Figura 4.3 essa mudança de centralidade para o grafo com $n = 11$, em que com $p = 2$ e $e = 7$ o vértice do centro da espinha é o mais central (a), porém com $p = 4$ e $e = 3$, os vértices dos extremos da espinha são os mais centrais (b), conforme indicado na Tabela 8 pela mudança de sinal da função $\gamma(p)$.



(a) (b)

Figura 4.3: Mudança na centralidade do grafo vassoura equilibrada com $n = 11$.

4.2.2. Centralidade de intermediação de vértices quando o número de vértices na espinha é par:

Nesta seção, analisaremos a centralidade de intermediação de vértices da vassoura dupla cuja espinha tem quantidade par de vértices, ou seja, os vértices da espinha são $(u_1, u_2, \dots, u_e)$, onde e é par. Consideraremos $n \geq 6$, pois desejamos trabalhar com casos não-degenerativos, excluindo quando o grafo é estrela e caminho($e \geq 2$ e $p \geq 2$).

O valor de k que representa qual vértice da espinha terá a maior centralidade é o mesmo que o encontrado na Seção 4.2. Ou seja, $k_v = \frac{e+1}{2}$.

Nota-se que para um e par, teremos um k_v não inteiro. Logo, neste caso, as duas maiores centralidades são obtidas em $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$, considerando a simetria da parábola.

$$C_I\left(u_{\frac{e}{2}}\right) = C_I\left(u_{\frac{e+2}{2}}\right) = \left(p + \frac{e}{2} - 1\right) * \left(n - p - \frac{e}{2}\right) = \frac{(n-2)n}{4}, \text{ (utilizando } e = n - 2p\text{)}, \quad (8)$$

Buscando identificar os vértices mais centrais, podemos comparar apenas a expressão da centralidade dos vértices extremos da espinha u_1 e u_e (5) com a expressão da centralidade dos vértices $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$ (8).

Para isso, podemos subtrair uma expressão da outra para verificar qual é maior, considerando a diferença:

$$\gamma(p) = \frac{-p^2 + 2np - 3p}{2} - \left(\frac{(n-2)n}{4}\right)$$

$$\gamma(p) = \frac{1}{4}\left[-2p^2 + (4n-6)p + 2n - n^2\right]$$

Quando $\gamma(p) > 0$, a primeira expressão é maior; se $\gamma(p) < 0$, a segunda expressão é maior. Estudaremos os valores de p para qualquer valor de n par fixo, de forma a determinar onde isso ocorre.

Essa expressão é uma função quadrática em relação a p, e as raízes podem ser encontradas com a fórmula geral de uma equação do segundo grau:

Logo, os valores de p que satisfazem $\gamma(p)$ são:

$$p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 9)}\right]$$

$$p_2 = \frac{1}{2}\left[(2n-3) + \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 9)}\right]$$

Dada as condições iniciais de $e \geq 2$ e $p \geq 2$, precisamos verificar a condição para $p_1 \geq 2$.

$$p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2\mathrm{n}^2 - 8\mathrm{n} + 9)}\right] \geq 2$$

$$n \geq 5 + \sqrt{5} \cong 7{,}23$$

Como n precisa ser um número inteiro, iremos considerar a solução viável a partir de $n \geq 8$. Dessa forma, o intervalo $p < p_1$ se torna inviável para o problema da vassoura dupla equilibrada com $n = 6$ ou $n = 7$.

Na Figura 4.4 apresentamos as duas funções de centralidades de vértices $C_I(u_1)$ e $C_I(u_{\frac{e+2}{2}})$ e os valores $p_1 e\, p_2$ nos quais $C_I(u_1)$= $C_I(u_{\frac{e+2}{2}})$. É possível observar o comportamento da centralidade de vértice. Para $p_1 < p < p_2$, temos que a centralidade de intermediação do vértice no centro da espinha, $C_I(u_{\frac{e+2}{2}})$, é maior do que a centralidade de intermediação dos vértices nas extremidades da espinha para qualquer $n \geq 8$ e par. Para $p < p_1$ ou $p > p_2$, temos que a centralidade de intermediação dos vértices dos extremos da espinha é maior do que a centralidade de intermediação do vértice no centro da espinha $C_I(u_{\frac{e+2}{2}})$.

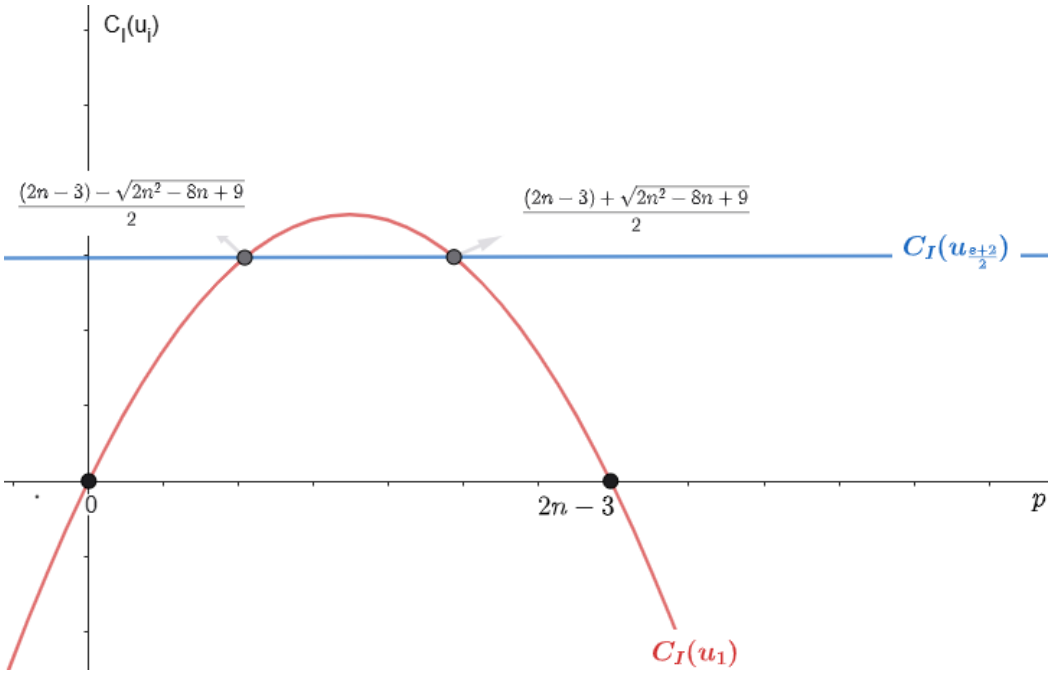


Figura 4.4: funções de centralidades de vértices analisadas e as raízes obtidas para $\gamma(p)$

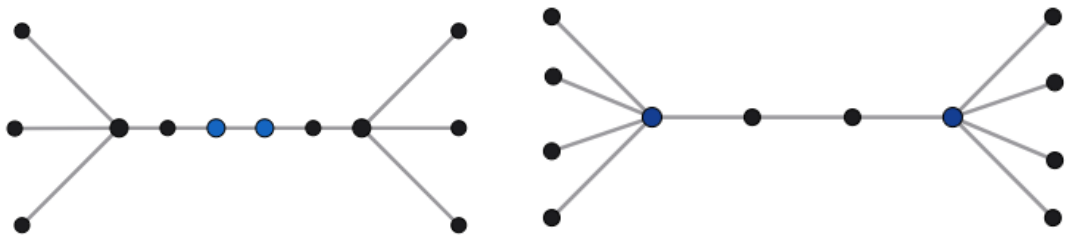
Ao fixarmos n, podemos determinar os valores específicos de p para os quais ocorre a mudança no comportamento da função $\gamma(p)$, auxiliando na caracterização das regiões onde temos a maior centralidade de intermediação de vértices: no vértice do centro da espinha ou nos vértices extremos da espinha. Na Tabela 9 apresentamos alguns valores para exemplificação, na qual podemos observar para cada n fixado, qual é o valor de p que altera o vértice a ser considerado com a maior centralidade. Por exemplo, para $n = 10$, temos que em $p = 2$, $\gamma(p)$ é negativo, o que significa que temos os dois vértices mais ao centro da espinha empatados com o maior valor de centralidade de intermediação. Em de $p = 3$ e $p = 4$, o sinal de $\gamma(p)$ é positivo, e isso significa que a maior centralidade é o obtida nos vértices extremos da espinha.

Tabela 9: Exemplos de valores obtidos para função γ

n	p	$\gamma(p)$	$e = n - 2p$
10	2	-5	6
	3	1	4
	4	6	2
12	2	-11	8
	3	-3	6

	4	4	4
14	2	-19	10
	3	-9	8
	4	0	6

Ilustramos na Figura 4.5 essa mudança de centralidade para o grafo com $n = 12$, em que com $p = 3$ $e = 6$ os dois vértices do centro da espinha são o mais centrais (a), porém com $p = 4$ e $e = 4$, os vértices dos extremos da espinha são os mais centrais (b), conforme indicado na Tabela 9 pela mudança de sinal da função $\gamma(p)$.



a) b)

Figura 4.5:Mudança na centralidade do grafo vassoura equilibrada com $n = 12$.

Um ponto interessante de ser observado é quando $\gamma(p) = 0$, como ocorreu em $n = 14$ e $p = 4$, significa que a centralidade de intermediação de vértices nos vértices extremos da espinha e nos dois vértices mais ao centro da espinha são iguais. Ou seja, nesse grafo temos quatro vértices mais centrais.

4.3 Centralidade de intermediação de arestas:

Nesta seção, desejamos estudar a centralidade de intermediação de arestas em uma vassoura dupla equilibrada, e o raciocínio utilizado é análogo ao que usamos para calcular a centralidade de intermediação de vértices na Seção 4.2. Similarmente, devido à simetria em uma vassoura dupla equilibrada temos que:

- $C_{aI}(u_i, u_{i+1}) = C_{aI}(u_{e-i}, u_{e-i+1})$ para $1 \leq i \leq e-1$
- $C_{aI}(v_i, u_1) = C_{aI}(w_j, u_e)$, $1 \leq i, j \leq p$

Não poderemos aplicar simetria quando e é par, pois temos um número ímpar de arestas neste caso, sendo a aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$ a do centro da espinha.

Recordemos que a centralidade de intermediação de uma aresta e, $C_{a_I}(e)$, em um grafo conexo, é determinada por:

$$C_{a_I}(e) = \sum_{i \neq j}^{\cdot} \frac{q_{ij(e)}}{q_{ij}},$$

onde $i, j \in V(G)$, $q_{ij(e)}$ é o número de caminhos mínimos com extremos nos vértices i e j que passam pela aresta e, e q_{ij} é o número total de caminhos mínimos com extremos nos vértices i e j.

Sabemos que em uma árvore existe um único caminho (mínimo) entre dois vértices quaisquer. Logo, o cálculo da centralidade de intermediação de uma aresta baseia-se em contar quantos caminhos mínimos passam por cada aresta. Em uma árvore, toda aresta conecta diretamente dois vértices adjacentes, de modo que o caminho mínimo entre esses dois vértices necessariamente inclui essa aresta. Assim, cada aresta é utilizada por pelo menos um caminho mínimo, garantindo que sua centralidade seja sempre maior que zero.

Por consequência, a centralidade de intermediação de uma aresta incidente a um vértice pendente é sempre $n-1$, pois, esta aresta está no caminho mais curto entre o vértice pendente e todos os demais vértices do grafo.

Dessa forma, a centralidade de intermediação das arestas pode ser calculada por:

- $C_{aI}(v_i, u_1) = C_{aI}(w_j, u_e) = n - 1, \qquad \forall\, 1 \leq i \leq p$ (9)
- $C_{aI}(u_i, u_{i+1}) = \underbrace{(p + i)}_{\text{Quantidade de vértices conectados à extremidade esquerda da aresta } (u_i, u_{i+1})} * \underbrace{(n - p - i)}_{\text{Quantidade de vértices conectados à extremidade direita da aresta } (u_i, u_{i+1})}$, para qualquer $1 \leq i \leq e - 1$, (10)

Diferentemente do que ocorre no caso da centralidade de vértices, aqui não é necessário tratar separadamente as arestas da espinha e das extremidades. As expressões (9) e (10) já contemplam todos os casos de forma direta: a primeira fornece a centralidade das arestas ligadas aos vértices pendentes, enquanto a segunda define a centralidade das arestas internas da espinha, variando de acordo com a posição i.

Análogo a centralidade dos vértices na espinha $C_I(u_k)$, observa-se que a expressão apresentada para a centralidade de intermediação das arestas na espinha $C_{aI}(u_i, u_{i+1})$, pode ser escrita como uma função quadrática em relação a i. A representação gráfica de $C_{aI}(u_i, u_{i+1})$ é uma parábola com concavidade voltada para baixo. Vamos utilizar a mesma lógica utilizada anteriormente para encontrar o valor que maximiza a função a partir da média aritmética das raízes.

$$(p + i) * (n - p - i) = 0$$

$$(p + i) = 0 \; ou \; (n - p - i) = 0$$

$$i_1 = -p$$

$$i_2 = n - p$$

$$i_v = \frac{[(-p)+(n-p)]}{2} = \frac{n-2p}{2} = \frac{e}{2}$$

4.3.1. Centralidade de intermediação de arestas quando o número de vértices da espinha é ímpar:

Observa-se que i_v é um número não inteiro, dessa forma, tendo em vista a simetria da parábola, consideramos os inteiros mais próximos $\frac{e-1}{2}$ e $\frac{e+1}{2}$, uma vez que i representa o índice dos vértices de uma aresta (u_i, u_{i+1}). Assim, temos que o maior valor de centralidade de aresta da espinha ocorre nas arestas $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e em $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$, sendo obtida pela expressão:

$$C_{a_I}\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right) = \left(p + \frac{e-1}{2}\right) * \left(n - p - \frac{e-1}{2}\right) = \frac{n^2-1}{4} \text{ (utilizando } e = n - 2p), \quad (11)$$

Portanto, buscando identificar as arestas mais centrais do grafo, comparamos a expressão da centralidade das arestas do centro da espinha (11) com a centralidade das arestas ligadas aos vértices pendentes (9). Portanto, seja $\gamma(n)$ a diferença entre $C_{aI}\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $C_{aI}(v_i, u_1)$:

$$\gamma(p) = \frac{n^2 - 1}{4} - (n - 1)$$

$$\gamma(p) = \frac{n^2 - 4n + 3}{4} = \frac{(n-3) * (n-1)}{4}$$

Temos que $\gamma(p)$ é sempre maior que zero para $n > 3$. Ao considerarmos $n \geq 7$ e ímpar, as arestas mais centrais são sempre as duas mais ao centro da espinha, $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$, como vemos em azul na Figura 4.6.

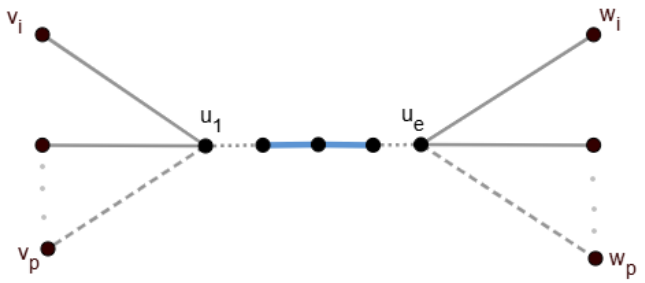


Figura 4.6: Ilustração das arestas mais centrais quando n é ímpar e $n \geq 7$.

4.3.2. Centralidade de intermediação de arestas quando o número de vértices na espinha é par:

Uma vez que i representa o índice dos vértices de uma aresta (u_i, u_{i+1}) e $i_v = \frac{e}{2}$, temos que o maior valor de centralidade de aresta na espinha, quando o número de vértices é par, ocorre na aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$, sendo obtida pela expressão:

$$C_{aI}\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right) = \left(p + \frac{e}{2}\right)\left(n - p - \frac{e}{2}\right) = \frac{n^2}{4} \text{ (utilizando } e = n - 2p), \qquad (12)$$

Portanto, buscando identificar as arestas mais centrais, podemos comparar apenas $C_{aI}\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$ com $C_{aI}(v_i, u_1) = C_{aI}(w_j, u_e)$, que representa a centralidade das arestas ligadas aos vértices pendentes.

Para isso, podemos calcular a diferença $\gamma(n)$ entre $C_{aI}\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$ e $C_{aI}(v_i, u_1)$:

$$\gamma(n) = \frac{n^2}{4} - (n-1)$$

$$\gamma(n) = \frac{n^2 - 4n + 4}{4} = \frac{(n-2)^2}{4}$$

Temos que $\gamma(n)$ é sempre maior que zero para $n > 2$. Ao considerarmos $n \geq 6$ e par a aresta mais central da vassoura dupla com quantidade par de vértices na espinha é sempre a aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$, conforme ilustrado em azul na Figura 4.7.

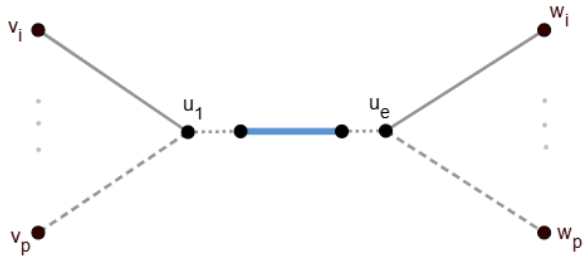


Figura 4.7: Ilustração da aresta mais central quando n é par e $n \geq 8$.

4.4 Comparando centralidade de intermediação de vértices e arestas:

Nesta seção, comparamos o comportamento das centralidades de intermediação de vértices e de arestas nos grafos vassouras duplas equilibradas, com foco na relação entre estes elementos. Inicialmente, examinamos a situação em que o número de vértices na espinha é ímpar, utilizando os resultados obtidos nas Subseções 4.2.1 e 4.3.1. Em seguida, realizamos a mesma análise para o caso em que a espinha possui número par

de vértices, utilizando os resultados obtidos nas Subseções 4.2.2 e 4.3.2, de modo a comparar os dois cenários e destacar similaridades e diferenças observadas entre eles.

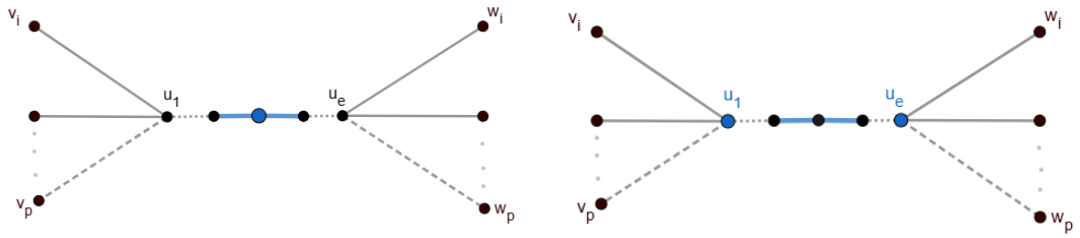
4.4.1. Número ímpar de vértices na espinha:

Para $n \geq 7$, $e \geq 3$ e $p \geq 2$, temos as seguintes análises nos grafos vassouras duplas equilibradas (que estão representadas na Figura 4.8 para cada n fixado:

1) Se $p < p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2n^2 - 8n + 7)}\right]$ ou $p > p_2 = \frac{1}{2}\left[(2n-3) + \sqrt{(2n^2 - 8n + 7)}\right]$, então o vértice no centro da espinha é o mais central (Figura 4.8 (a)).

2) Se $\frac{1}{2}\left[(2n-3) - \sqrt{(2n^2 - 8n + 7)}\right] < p < \frac{1}{2}\left[(2n-3) + \sqrt{(2n^2 - 8n + 7)}\right]$, então os vértices dos extremos da espinha têm a mesma centralidade de vértice e são aqueles com a maior centralidade no grafo (Figura 4.8 (b)).

3) Para qualquer $n \geq 7$, ambas as arestas localizadas mais ao centro da espinha são as mais centrais, $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$.



a) b)

Figura 4.8: Representação dos vértices e arestas mais centrais para e ímpar

A vassoura dupla equilibrada com quantidade ímpar de vértices na espinha apresenta dois comportamentos distintos dependendo do valor de p. Quando p está no intervalo $p_1 < p < p_2$, o grafo possui dois vértices mais

centrais e duas arestas mais centrais, sendo que esses vértices não estão associados a nenhuma das arestas mais centrais. Por esta razão este caso não é classificado em nenhuma das configurações apresentadas na Seção 3.3.

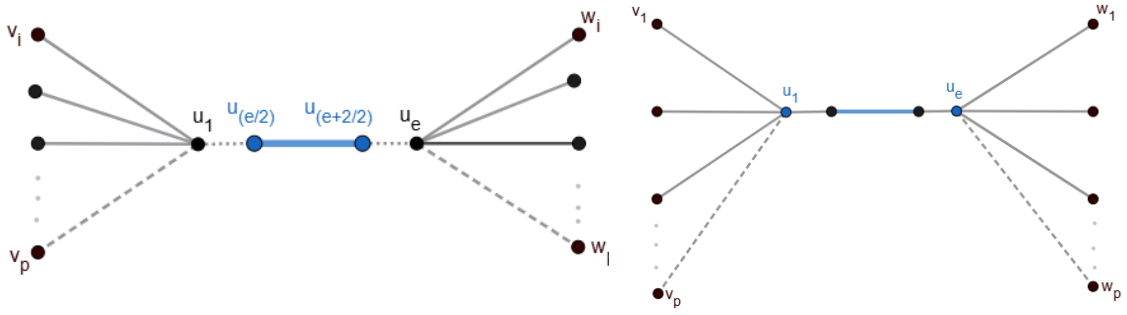
Quando $p < p_1$ ou $p > p_2$ o grafo continua apresentando duas arestas mais centrais, mas agora há apenas um vértice mais central. Nesse caso, ambas as arestas mais centrais são incidentes a esse único vértice mais central. Por isso, classificamos esse caso como caso amplo 1-incidente.

Ressaltamos que não há vassoura dupla com quantidade ímpar de vértices na espinha que se enquadre nas classificações restrito 1-incidente restrito 2-incidente, ou amplo 2-incidente.

4.4.2. Número par de vértices na espinha:

Para $n \geq 6$, $e \geq 2$ e $p \geq 2$, temos as seguintes análises nos grafos vassouras duplas equilibradas (que estão representadas na Figura 4.9) para cada n fixado:

1) Se $p < p_1 = \frac{1}{2}\left[(2n-3) - \sqrt{(2n^2 - 8n + 9)}\right.$ ou $p > p_2 = \frac{1}{2}\left[(2n-3) + \sqrt{(2n^2 - 8n + 9)}\right]$, os dois vértices no meio da espinha são os mais centrais, $u\frac{e}{2}$ e $u\frac{e+2}{2}$ (Figura 4.9(a)). Ressaltamos que $p < p_1$ é inviável para $n = 6$ e $n = 7$.
2) Se $\frac{1}{2}\left[(2n-3) - \sqrt{(2n^2 - 8n + 9)}\right] = p_1 < p < p_2 = \frac{1}{2}\left[(2n-3) + \sqrt{(2n^2 - 8n + 9)}\right]$, os vértices extremos da espinha u_1 e u_e possuem o mesmo valor de centralidade e são os vértices mais centrais do grafo (Figura 4.9 (b)).
3) Para qualquer $n \geq 8$, a aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$ localizada mais ao centro da espinha é a mais central.



a) b)

Figura 4.9:Representação dos vértices e arestas mais centrais para e par.

A vassoura dupla equilibrada com número par de vértices na espinha $(n \geq 6), e \geq 2$ e $p \geq 2$ apresenta diferentes configurações de vértices mais centrais conforme o valor de p. Quando p está no intervalo $p_1 < p < p_2$, os vértices extremos da espinha, u_1 e u_e, são os mais centrais do grafo, com valores de centralidade iguais (Figura 4.8(b)). Já quando $p < p_1$ ou $> p_2$, os vértices localizados no meio da espinha, $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$, tornam-se os mais centrais (Figura 4.8(a)).

Independentemente do valor de p, a aresta mais central é sempre a aresta localizada ao centro da espinha, isto é, $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$. Dessa forma, quando $p > p_2$, os vértices mais centrais estão associados a aresta mais central, pois a aresta mais central é incidente a ambos os vértices mais centrais. Esse caso se enquadra como caso restrito 2-incidente. Quando $p_1 < p < p_2$, os vértices mais centrais estão nas extremidades da espinha, não sendo incidentes à aresta mais central. Assim, não se verifica associação entre vértices mais centrais e a aresta mais central, nem mesmo parcialmente. Dessa forma, não há ocorrência dos casos restrito 1-incidente, amplo 1-incidente ou amplo 2-incidente.

4.5 Centralidade de proximidade de vértices:

Nesta seção, analisamos a centralidade de proximidade dos vértices e das arestas nos grafos vassouras duplas equilibradas, com o objetivo de identificar os vértices estruturalmente mais próximos dos demais e compreender como essa métrica se comporta à medida que a estrutura da árvore cresce

simetricamente. A centralidade de proximidade fornece uma medida inversamente proporcional à soma das distâncias de um vértice a todos os outros, destacando aqueles que estão mais bem posicionados em termos de alcance na rede. Consideramos uma vassoura dupla equilibrada com n vértices, onde n é um número inteiro, sendo $\boldsymbol{e}$ vértices na espinha e p vértices pendentes. Temos $n = e + 2p$. Logo, $e = n - 2p$, substituição que utilizaremos em diversos momentos desta seção.

A centralidade de proximidade dos vértices é calculada como o inverso da soma das distâncias $\sum_{i \in V} d(v, i)$, essa soma é chamada de Transmissão do vértice v. Dessa forma, temos que a centralidade de proximidade dos vértices pode ser obtida por:

$$C_P(v_i) = \frac{1}{T(v_i)}$$

onde dada um vértice v, $T(v)$ é a transmissão do vértice v.

Neste trabalho, optamos por utilizar diretamente a transmissão como medida comparativa, uma vez que ela evita o uso de números decimais nas contas, facilitando as expressões algébricas. Assim, o vértice mais central será o vértice que possui o menor valor de Transmissão.

Devido à simetria da vassoura dupla equilibrada, temos que:

- $C_P(u_1) = C_P(u_e)$ (extremidades da espinha têm a mesma centralidade).
- $C_P(u_k)$ sendo a centralidade de qualquer u_k, $1 < k < e$
- $C_P(v_i) = C_P(w_j)$, $\quad 1 \leq i, j \leq p$

A transmissão dos vértices pode ser calculada por:

Distância de u_1 até os demais vértices da espinha

- $\mathrm{T}(u_1) = \mathrm{T}(u_e) = \underbrace{1 * \mathrm{p}}_{\text{Soma da distância de } u_1 \text{ até cada folha do lado esquerdo}} + \sum_{i=1}^{e-1} i + \underbrace{[(e-1)+1] * p}_{\text{Soma da distância de } u_1 \text{ até cada folha do lado direito}} = p(1+e) + \frac{(e-1)e}{2}$

$$\mathrm{T}(u_1) = \mathrm{T}(u_e) = \frac{n^2-2np-n+4p}{2} \text{ (quando substituímos } e = n-2p\text{),} \quad (13)$$

- $T(u_k) = \underbrace{p*((k-1)+1)}_{\text{Soma da distância de } u_k \text{ até cada folha do lado esquerdo}} + \underbrace{p*((e-k)+1)}_{\text{Soma da distância de } u_k \text{ até cada folha do lado direito}} + \underbrace{\frac{(k-1)k}{2}}_{\text{Soma das distâncias de } u_k \text{ até } u_i, \text{ com } i<k} + \underbrace{\frac{(e-k)(e-k+1)}{2}}_{\text{Soma das distâncias de } u_k \text{ até } u_i, \text{ com } i>k}$

$$T(u_k) = \frac{2k^2+k(4p-2n-2)+n^2-2np+n}{2} \text{ (quando substituímos } e = n-2p\text{),} \quad (14)$$

- $T(v_i) = \underbrace{(p-1)*2}_{\text{Soma da distância de } v_i \text{ para } v_j \text{, com } i \neq j} + \underbrace{p(e+1)}_{\text{Soma da distância de } v_i \text{ para } w_i} + \underbrace{\frac{e(e+1)}{2}}_{\text{Soma da distância de } v_i \text{ para os vértices da espinha}} = \frac{e^2+2ep+e+6p-4}{2}$

$$T(v_i) = \frac{n^2-2np+n+4p-4}{2} \text{ (quando substituímos } e = n-2p\text{),} \quad (15)$$

Observa-se que a expressão apresentada para a transmissão dos vértices na espinha $T(u_k)$, sem considerar os extremos, pode ser escrita como uma função quadrática em relação a k. Como a expressão possui um termo quadrático positivo, sua representação gráfica é uma parábola com concavidade voltada para cima. Dessa forma, para determinar qual valor de k possui o menor valor para a transmissão $T(u_k)$, basta determinarmos quando a função quadrática atinge seu valor mínimo.

Como a parábola é simétrica, o k que minimiza a função pode ser encontrado através da fórmula $V_x = \frac{-b}{2a}$. Aplicando esse procedimento, obtemos:

$$T(u_k) = \frac{2k^2 + k(4p - 2n - 2) + n^2 - 2np + n}{2}$$

$$k_{min} = \frac{-2(2p-n-1)}{2*2} = \frac{-2p+n+1}{2} = \frac{e+1}{2}$$

No caso em que a espinha tem comprimento ímpar, isto é, e é um número ímpar, o valor k_{min} indica a posição do vértice ao longo da espinha em que a centralidade de proximidade de vértices atinge seu valor máximo dentro das restrições definidas. No caso em que o comprimento da espinha é par, embora k_{min} não seja um número inteiro, pela simetria da parábola, é possível identificar em quais vértices ocorrem o valor máximo de centralidade, como será discutido na Seção 4.5.2.

Como na análise por centralidade de intermediação, identificamos uma distinção entre os casos com número par e ímpar de vértices na espinha. Essa diferença motiva a formulação de duas generalizações distintas: uma para espinhas com quantidade ímpar de vértices e outra para espinhas com quantidade par de vértices.

4.5.1. Centralidade de proximidade de vértices quando número de vértices na espinha é ímpar:

Ao consideramos n ímpar, é possível identificar que dentre os vértices da espinha, excluindo os extremos, o valor máximo da centralidade do grafo considerado, para qualquer n ímpar e qualquer $p \geq 2$, é encontrado no vértice do centro da espinha $\frac{e+1}{2}$, que possui a seguinte transmissão:

$$T\left(u_{\frac{e+1}{2}}\right) = \frac{2\left(\frac{n-2p+1}{2}\right)^2 + \frac{n-2p+1}{2}\left(4p - 2n - 2\right) + n^2 - 2np + n}{2}$$

Simplificando $T\left(u_{\frac{e+1}{2}}\right)$, temos que:

$$T\left(u_{\frac{e+1}{2}}\right) = -p^2 + p + \frac{n^2-1}{4}, \tag{16}$$

Portanto, buscando identificar os vértices mais centrais, podemos comparar a expressão (13) obtida para a transmissão dos vértices extremos da

espinha u_1 e u_e com a expressão (16) obtida para a transmissão do vértice do centro da espinha $u_{\frac{e+1}{2}}$. Note que, como qualquer vértice pendente possui como único vizinho o vértice extremo da espinha, então sua transmissão necessariamente é maior do que as transmissões dos vértices dos extremos.

Para isso, podemos calcular a diferença entre $\mathrm{T}(u_1)$ e $T\left(u_{\frac{e+1}{2}}\right)$ (respectivamente, as expressões (13) e (16)) para verificar qual é maior, considerando a diferença:

$$\gamma(p) = \frac{n^2 - 2np - n + 4p}{2} - \left(-p^2 + p + \frac{n^2 - 1}{4}\right)$$

$$\gamma(p) = p^2 + (1-n)p + \frac{n^2 - 2n + 1}{4}$$

Quando $\gamma(p) > 0$, $T(u_1)$ é maior; e quando $\gamma(p) < 0$, $T\left(u_{\frac{e+1}{2}}\right)$ é maior. Estudaremos os valores de p para qualquer valor de n fixo, de forma a determinar onde isso ocorre. Os zeros da função $\gamma(p)$ são:

$$p_1 = p_2 = \frac{1}{2}\left[(n-1) \pm \sqrt{0}\right] = \frac{n-1}{2}$$

Logo, $\gamma(p)$ é sempre maior que zero para qualquer $n > 1$. Isso significa que a transmissão do vértice u_1 , $T(u_1)$ é sempre maior do que a do vértice central $u_{\frac{e+1}{2}}$, $T\left(u_{\frac{e+1}{2}}\right)$. Portanto, o vértice do centro da espinha é sempre o mais central em termos de proximidade, sem mudança de centralidade com o aumento de p, diferente do que observamos no caso da centralidade de intermediação.

4.5.2. Centralidade de proximidade de vértices quando o número de vértices na espinha é par:

Agora consideramos uma vassoura dupla com n vértices, onde n é um número inteiro par, $p \geq 2$ e $e \geq 2$. Nota-se que para um e par, teremos um k_{min} não inteiro, onde k_{min} representa o índice do vértice. Neste caso, um k_{min} não

inteiro representa que há um empate entre os valores de Transmissão dos vértices mais próximos de k_{min} , $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$.

Ou seja, é possível observar que o mínimo da transmissão ocorre nos dois vértices do meio da espinha, e ela cresce conforme k se afasta do centro. Assim, é possível identificar que dentre os vértices da espinha, excluindo os extremos, o valor máximo da centralidade de proximidade do grafo considerado, para qualquer n par e qualquer p, é encontrado nos vértices $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$.

Dessa forma, a transmissão dos dois vértices mais ao centro da espinha é obtida, já considerando que $e = n - 2p$, por:

$$T\left(u_{\frac{e}{2}}\right) = \frac{2\left(\frac{n-2p}{2}\right)^2 + \frac{n-2p}{2}\left(4p - 2n - 2\right) + n^2 - 2np + n}{2}$$

Simplificando $T\left(u_{\frac{e}{2}}\right)$, temos que:

$$T\left(u_{\frac{e}{2}}\right) = -p^2 + p + \frac{n^2}{4}, \tag{17}$$

Portanto, buscando identificar os vértices mais centrais, podemos comparar a expressão (13) obtida para a transmissão dos extremos da espinha u_1 e u_e com a expressão (17) da transmissão dos vértices $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$.

Para isso, podemos subtrair uma expressão da outra para verificar qual é maior, considerando a diferença:

$$\gamma(p) = \frac{n^2 - 2np - n + 4p}{2} - \left(-p^2 + p + \frac{n^2}{4}\right)$$

$$\gamma(p) = \left[p^2 + (1-n)p + \frac{(n^2 - 2n)}{4}\right]$$

Quando $\gamma(p) > 0$, $T(u_1)$ é maior; e quando $\gamma(p) < 0$, $T\left(u_{\frac{e+1}{2}}\right)$ é maior. Estudaremos os valores de p para qualquer valor de n fixo, de forma a determinar onde isso ocorre. Os zeros da função $\gamma(p)$ são:

$$p_1 = \frac{1}{2}[(n-1) + 1]$$

$$p_2 = \frac{1}{2}[(n-1) - 1]$$

Logo, $p_1 = \frac{n}{2} - 1 \ e \ p_2 = \frac{n}{2}$.

Dada as condições iniciais de $e \geq 2$ e $p \geq 2$, precisamos verificar a condição para $p_1 \geq 2$.

$$p_1 = \frac{n}{2} - 1 \geq 2$$

$$n \geq 6$$

Como a representação gráfica de $\gamma(p)$ é uma parábola com concavidade para cima, concluímos que, para $n \geq 6$, a função $\gamma(p)$ admite dois zeros reais distintos, $p_1 = \frac{n}{2} - 1 \ e \ p_2 = \frac{n}{2}$, que delimitam a mudança no comportamento da centralidade dos vértices. No entanto, como estamos considerando apenas valores inteiros de $p \geq 2$, e a diferença entre $p_1 \ e \ p_2$ é exatamente 1, não há valor inteiro de p entre p_1 e p_2. Isso implica que não ocorre transição de centralidade dentro do domínio considerado. Portanto, a comparação entre as transmissões mostra que, para todos os valores inteiros de p considerados nesse estudo, $T(u_1)$ sempre será maior que $T\left(u_{\frac{e+1}{2}}\right)$. Dessa forma, a centralidade de proximidade dos dois vértices mais ao centro da espinha, $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$, é maior.

4.6 Centralidade de proximidade de arestas:

Nesta seção, desejamos estudar a centralidade de proximidade de arestas em uma vassoura dupla equilibrada, e o raciocínio utilizado é análogo ao que usamos para calcular a centralidade de proximidade de vértices na Seção 4.5. Ou seja, optamos por utilizar diretamente a transmissão como medida comparativa, evitando o uso de números decimais nas contas. Assim, buscamos a aresta com o menor valor de transmissão.

Cabe relembrar a fórmula de calcular a proximidade de arestas utilizada neste estudo é dada por:

$$C_{aP}(x,y) = \frac{1}{T(x) + T(y)}$$

onde dada uma aresta (x,y), $T(x)$ é a transmissão do vértice x, e $T(y)$ a transmissão do vértice y.

Portanto, devido à simetria em uma vassoura dupla equilibrada temos que:

- $C_{aP}(u_1,u_2) = C_{aP}(u_{e-1},u_e)$
- $C_{aP}(u_i,u_{i+1}) = C_{aP}(u_{e-i},u_{e-i+1}),\ 1 \leq i \leq e$.
- $C_{aP}(v_i,u_1) = C_{aP}(w_j,u_e),\quad 1 \leq i,j \leq p$

De acordo com a definição de centralidade de proximidade das arestas, temos:

- $C_{aP}(u_1,u_2) = C_{aP}(u_{e-1},u_e) = \frac{1}{T(u_1)+T(u_2)}$

- $C_{aP}(u_k,u_{k+1}) = \frac{1}{T(u_k)+T(u_{k+1})}$, para qualquer k na espinha, excetos os extremos.

- $C_{aP}(v_i,u_1) = C_{aI}(w_j,u_e) = \frac{1}{T(v_1)+T(u_1)}$

A partir das Equações (13), (14) e (15) temos valores de transmissão dos vértices. Logo, temos que:

- $T(u_1) + T(u_2) = \frac{n^2-2np-n+4p}{2} + \frac{2(2)^2+(2)(4p-2n-2)+n^2-2np+n}{2} = n^2 - 2n + (6-2n)p + 2$

- $T(u_k) + T(u_{k+1}) = \frac{2k^2+k(4p-2n-2)+n^2-2np+n}{2} + \frac{2(k+1)^2+(k+1)(4p-2n-2)+n^2-2np+n}{2} = 2k^2 + (4p-2n)k + n^2 - 2np + 2p$

- $T(v_1) + T(u_1) = \frac{n^2-2np+n+4p-4}{2} + \frac{n^2-2np-n+4p}{2} = n^2 + (4-2n)p - 2$

Portanto, temos as seguintes centralidades de proximidade de aresta:

- $C_{aP}(u_1, u_2) = C_{aP}(u_{e-1}, u_e) = \frac{1}{T(u_1)+T(u_2)} = \frac{1}{n^2-2n+(6-2n)p+2}$, (18)

- $C_{aP}(u_k, u_{k+1}) = \frac{1}{T(u_k)+T(u_{k+1})} = \frac{1}{2k^2+(4p-2n)k+n^2-2np+2p}$, (19)

para qualquer k na espinha, excetos os extremos.

- $C_{aP}(v_i, u_1) = C_{aI}(w_j, u_e) = \frac{1}{T(v_1)+T(u_1)} = \frac{1}{n^2+(4-2n)p-2}$, (20)

É importante observar que as arestas pendentes, (v_i, u_1) e (w_j, u_e), apresentam os menores valores de centralidade de proximidade de aresta. Isso se deve ao fato de que as folhas, por estarem situadas nas extremidades do grafo e acessarem os demais vértices exclusivamente por meio da aresta pendente, possuem, em geral, transmissões elevadas. A transmissão $T(v_i)$ de um vértice folha é maximizada justamente porque todos os caminhos partindo da folha incluem necessariamente a aresta pendente e percorrem longas distâncias até os vértices do lado oposto da vassoura dupla equilibrada.

Análogo ao que foi observado para a expressão da transmissão dos vértices na espinha $T(u_k)$, observa-se que a expressão $T(u_k) + T(u_{k+1})$, é uma função quadrática em relação a k, cuja representação gráfica é uma parábola com concavidade voltada para cima. Dessa forma, vamos utilizar a mesma lógica utilizada anteriormente para calcular o valor minimiza $T(u_k) + T(u_{k+1})$, a partir de $V_x = \frac{-b}{2a}$.

$$T(u_k) + T(u_{k+1}) = 2k^2 + (4p-2n)k + n^2 - 2np + 2p$$

$$k_{min} = \frac{-(4p-2n)}{2*2} = \frac{n-2p}{2} = \frac{e}{2}$$

Similar às outras discussões, vamos separar a análise em grafos com número ímpar de vértices na espinha e número par de vértices na espinha.

4.6.1. Centralidade de proximidade de arestas quando o número de vértices na espinha é ímpar:

Ao considerarmos e ímpar, observa-se que k_{min} é um número não inteiro, dessa forma, considera-se os inteiros mais próximos $\frac{e-1}{2}$ e $\frac{e+1}{2}$. Uma vez que k representa o índice dos vértices de uma aresta (u_k, u_{k+1}), e que com e ímpar temos um número par de arestas, temos que o maior valor de centralidade de aresta ocorre em $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e em $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$, sendo obtido por:

$$C_{aP}\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right) = \frac{1}{T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)} = \frac{2}{n^2 - 4p^2 + 4p + 1}, \quad (21)$$

Portanto, buscando identificar as arestas mais centrais, podemos comparar $T(u_1) + T(u_2)$ e $T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)$. Seja $\gamma(p)$ a diferença entre as expressões:

$$\gamma(p) = (\, n^2 - 2n + (6 - 2n)p + 2\,) - (\, n^2 - 4p^2 + 4p + 1)$$

$$\gamma(p) = 4p^2 + (2 - 2n)p - 2n + 1$$

Se $\gamma(p) > 0$, então $T(u_1) + T(u_2)$ é maior que é $T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)$. Se $\gamma(p) < 0$, o contrário ocorre.

$\gamma(p)$ é uma função quadrática em relação a p, e os seus zeros são dados por:

$$p_1 = \frac{1}{8}\left[2n - 2 - \sqrt{4n^2 + 24n - 12}\right]$$

$$p_2 = \frac{1}{8}\left[2n - 2 + \sqrt{4n^2 + 24n - 12}\right]$$

Dada as condições iniciais de $e \geq 2$ e $p \geq 2$, precisamos verificar a condição para $p_1 \geq 2$.

$$p_1 = \frac{1}{8}\left[2n - 2 - \sqrt{4n^2 + 24n - 12}\right] \geq 2$$

$$n \leq 3{,}5$$

Como só consideramos valores de $n \geq 7$, só podemos considerar o intervalo onde $p > p_2$. Para $p > p_2$, temos $\gamma(p) > 0$, indicando que $T(u_1) + T(u_2)$ é sempre maior que $T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)$ para o grafo considerado. Dessa forma, o maior valor de centralidade de aresta ocorre nas duas arestas mais ao meio da espinha, nas arestas $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$.

4.6.2. Centralidade de proximidade de arestas quando o número de vértices na espinha é par:

Ao considerarmos e par, temos o $k_{min} = \frac{e}{2}$. Uma vez que k representa o índice dos vértices de uma aresta (u_k, u_{k+1}), e que com e par temos um número ímpar de arestas, temos que o maior valor de centralidade de aresta ocorre em $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$, sendo obtido por:

$$C_{aP}\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right) = \frac{1}{T\left(u_{\frac{e}{2}}\right) + T\left(u_{\frac{e}{2}+1}\right)} = \frac{2}{n^2 - 4p^2 + 4p}, \quad (22)$$

Portanto, buscando identificar as arestas mais centrais, podemos comparar $T(u_1) + T(u_2)$ e $T\left(u_{\frac{e}{2}}\right) + T\left(u_{\frac{e}{2}+1}\right)$. Seja $\gamma(p)$ a diferença entre as expressões:

$$\gamma(p) = (\, n^2 - 2n + (6 - 2n)p + 2\,) - (n^2 - 4p^2 + 4p)$$

$$\gamma(p) = 4p^2 + (2 - 2n)p - 2n + 2$$

Se $\gamma(p) > 0$, então $T(u_1) + T(u_2)$ é maior que é $T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)$. Se $\gamma(p) < 0$, o contrário ocorre.

$\gamma(p)$ é uma função quadrática em relação a p cujos zeros são dados por:

$$p_1 = \frac{1}{8}\left[2n - 2 - \sqrt{4n^2 + 24n\ - 28}\right]$$

$$p_2 = \frac{1}{8}\left[2n - 2 + \sqrt{4n^2 + 24n\ - 28}\right]$$

Dada as condições iniciais de $e \geq 2$ e $p \geq 2$, precisamos verificar a condição para $p_1 \geq 2$.

$$p_1 = \frac{1}{8}\left[2n - 2 - \sqrt{4n^2 + 24n\ - 28}\right] \geq 2$$

$$n \leq 3{,}67$$

Como só consideramos valores de $n \geq 6$, só podemos considerar o intervalo onde $p > p_2$. Ou seja, para $p > p_2$, temos $\gamma(p) > 0$, indicando que $T(u_1) + T(u_2)$ é sempre maior que $T\left(u_{\frac{e}{2}}\right) + T\left(u_{\frac{e}{2}+1}\right)$ para o grafo considerado. Dessa forma, o maior valor de centralidade de aresta ocorre na aresta mais ao meio da espinha, na aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$.

4.7 Comparando centralidade de proximidade de vértices e arestas:

Nesta seção, comparamos o comportamento das centralidades de proximidade de vértices e de arestas nos grafos vassouras duplas equilibradas, com foco na relação entre estes elementos. Inicialmente, examinamos a situação em que o número de vértices na espinha é ímpar, utilizando os resultados obtidos nas Subseções 4.5.1 e 4.6.1. Em seguida, realizamos a mesma análise para o caso em que a espinha possui número par de vértices, utilizando os resultados obtidos nas Subseções 4.5.2 e 4.6.2, de modo a comparar os dois cenários e destacar similaridades e diferenças observadas entre eles.

4.7.1. Número ímpar de vértices na espinha:

Para $n \geq 7$, $e \geq 3$ e $p \geq 2$ temos as seguintes análises para cada n fixado:

Ao analisarmos os vértices $T(u_1)$ é sempre maior do que $T\left(u_{\frac{e+1}{2}}\right)$. Portanto, o vértice do centro da espinha é sempre o mais central em termos de proximidade, sem mudança de centralidade com o aumento de p

Ao analisarmos arestas só podemos considerar o intervalo onde $p > p_2$. Para $p > p_2$, temos $\gamma(p) > 0$, indicando que $T(u_1) + T(u_2)$ é sempre maior que $T\left(u_{\frac{e-1}{2}}\right) + T\left(u_{\frac{e+1}{2}}\right)$. Dessa forma, o maior valor de centralidade de aresta ocorre nas duas arestas mais ao meio da espinha, ou seja, nas arestas $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$.

Sendo assim, tanto o vértice do meio da espinha, $u_{\frac{e+1}{2}}$, como as duas arestas mais ao centro da espinha, $\left(u_{\frac{e-1}{2}}, u_{\frac{e+1}{2}}\right)$ e $\left(u_{\frac{e+1}{2}}, u_{\frac{e+3}{2}}\right)$, são os mais centrais, o que caracteriza a configuração amplo 1- incidente, como ilustramos na Figura 4.10.

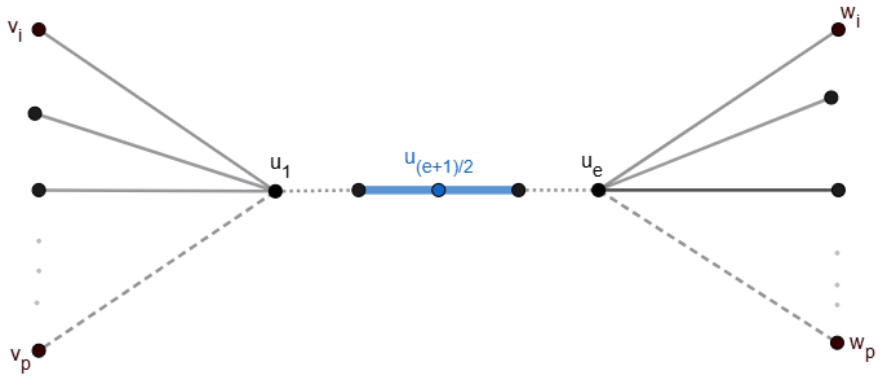


Figura 4.10: Ilustração dos vértices e arestas mais centrais para n ímpar e $n \geq 7$.

4.7.2. Número par de vértices na espinha:

Para $n \geq 6$, $e \geq 2$ e $p \geq 2$ temos as seguintes análises para cada n fixado:

Ao analisarmos os vértices temos as raízes $p_1 = \frac{n}{2} - 1$ e $p_2 = \frac{n}{2}$, que delimitam a mudança no comportamento da centralidade dos vértices. Porém,

considerando apenas valores inteiros de $p \geq 2$, não há valor inteiro de p entre p_1 e p_2. Portanto, para todos os valores inteiros de $p \geq 2$, $T(u_1)$ sempre será maior que $T\left(u_{\frac{e+1}{2}}\right)$. Dessa forma, a centralidade de proximidade dos dois vértices do centro da espinha, $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$, é sempre maior.

Ao analisarmos as arestas, só podemos considerar o intervalo onde $p > p_2$. Para $p > p_2$, temos $\gamma(p) > 0$, indicando que $T(u_1) + T(u_2)$ é sempre maior que $T\left(u_{\frac{e}{2}}\right) + T\left(u_{\frac{e}{2}+1}\right)$. Dessa forma, o maior valor de centralidade de aresta ocorre na aresta mais ao meio da espinha, a aresta $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$.

Sendo assim, tanto os dois vértices no meio da espinha, $u_{\frac{e}{2}}$ e $u_{\frac{e+2}{2}}$, como a aresta do centro, $\left(u_{\frac{e}{2}}, u_{\frac{e}{2}+1}\right)$ são os mais centrais, o que caracteriza uma configuração restrita 2- incidente, como ilustramos na Figura 4.11.

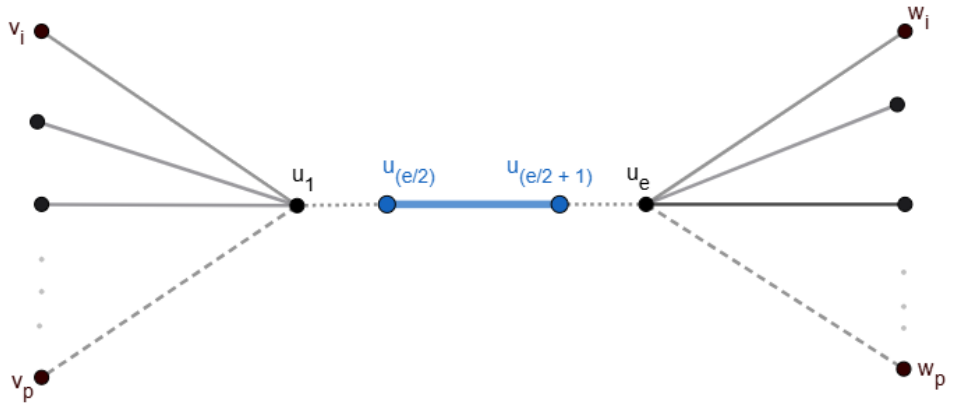


Figura 4.11: Ilustração dos vértices e das arestas mais centrais para $n \geq 6$ e n par.

5 O PROBLEMA DA ÁRVORE DE STEINER MÍNIMA E A HEURÍSTICA KMB

Neste capítulo, abordaremos o problema da árvore de Steiner mínima, uma questão clássica na Teoria dos Grafos, e a heurística proposta por Kou, Markowsky e Berman em 1981, uma das abordagens utilizadas para a construção de árvores de Steiner mínimas, conhecida como heurística KMB. Iniciaremos com uma introdução ao problema da árvore de Steiner mínima na Seção 5.1, detalhando suas principais características e relevância em diferentes contextos. Na Seção 5.2, discutiremos a heurística KMB, explicando seu funcionamento. Na Seção 5.3, apresentamos as variações dessa heurística, que incorporam conceitos de centralidade de arestas. Por fim, na Seção 5.4, apresentaremos uma comparação dos resultados obtidos por essas variações.

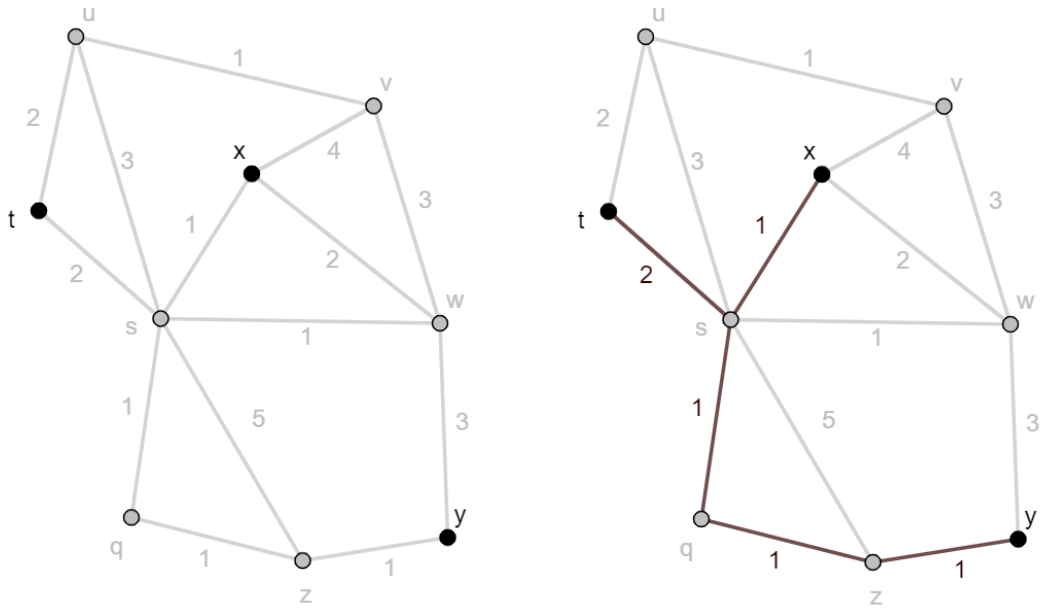
5.1 Problema da Árvore de Steiner mínima em Grafos

O Problema da Árvore de Steiner mínima é uma questão clássica de interconexão proposta por Fermat no século XVII. Inicialmente, o problema consistia em encontrar um ponto adicional em um plano euclidiano tal que a soma das distâncias deste ponto aos três pontos dados fosse mínima. Posteriormente, este problema evoluiu para o contexto de aplicações em grafos, originando o Problema de Steiner em grafos [Goldbarg e Goldbarg, 2012].

Seja $G = (V, E, c)$ um grafo não direcionado, conexo e ponderado, onde V representa o conjunto de vértices, E o conjunto de arestas e os custos são dados por uma função $c: E \rightarrow \mathbb{R}$. Considerando um subconjunto de vértices S ⊆ V, denominados vértices terminais, uma árvore de Steiner é qualquer subgrafo conexo de G sem ciclos que conecta todos os vértices terminais S e pode conter vértices auxiliares de $V - S$, denominados vértices de Steiner. O problema da árvore de Steiner mínima é encontrar a árvore de Steiner de menor custo, obtida pela soma dos custos das arestas presentes na árvore. Mais detalhes sobre este problema podem ser encontrados em Dreyfus e Wagner [1972].

Para ilustrar o problema da árvore de Steiner mínima, considere um grafo simples $G = (V, E, c)$, onde os vértices do conjunto $V = \{ q, s, t, u, v, x, w, y, z\}$

representam cidades e as arestas representam estradas entre elas, com custos correspondentes às distâncias. Deseja-se conectar um subconjunto $S = \{t, x, y\}$ de cidades (vértices terminais) da maneira mais econômica possível. Na Figura 5.1(a) exemplificamos esta situação. A árvore de Steiner mínima obtida é representada na Figura 5.1(b) e possui custo de 6, e tem como vértices de Steiner os vértices q, s e z.



a) Grafo G b) Árvore de Steiner mínima S de G

Figura 5.1:Exemplo de Árvore de Steiner mínima S de G. Adaptado de Goldbarg e Goldbarg [2012].

As árvores de Steiner possuem uma vasta gama de aplicações em diferentes áreas de estudo. Goldbarg e Goldbarg [2012] apresentam alguns exemplos para demonstrar as áreas de aplicação das Árvores de Steiner, como projeto de placas de circuitos integrados e projeto de redes de distribuição de gás natural.

Para o projeto de placas de circuitos integrados, Goldbarg e Goldbarg [2012] argumentam que é preciso conectar vários componentes eletrônicos (como resistores, capacitores etc.) da maneira mais eficiente possível. O objetivo é minimizar o comprimento total das conexões para economizar material, reduzir custos, diminuir a dissipação de energia e melhorar a confiabilidade do circuito.

Para cada componente, há várias posições possíveis onde as conexões podem ser feitas e essas posições formam um agrupamento. Dessa forma, os pontos de conexão e as possíveis trilhas que os unem são representados em um grafo e a Árvore de Steiner mínima obtida é uma rede que conecta um ponto de cada agrupamento da forma mais curta possível.

No projeto de redes de distribuição de gás natural o objetivo é minimizar os custos da construção de gasodutos, atendendo o maior número de pontos possíveis [Goldbarg e Goldbarg, 2012]. O gás natural precisa ser distribuído para vários pontos (clientes) ao longo de uma área. Esses clientes estão em diferentes locais e precisam ser conectados à central de distribuição de gás através de tubulações (gasodutos). A Árvore de Steiner mínima obtida deve conectar um conjunto de clientes (pontos de demanda), minimizando os custos de construção da rede, ao mesmo tempo que maximiza os ganhos por atender esses clientes.

5.2 Heurística KMB

Entre as diversas heurísticas propostas para a resolução do problema da árvore de Steiner mínima em grafos, a abordagem desenvolvida por Kou, Markowsky e Berman em 1981, conhecida como heurística KMB [Kou *et.*al 1981], se destaca. Esta heurística é amplamente reconhecida na literatura por sua eficiência e simplicidade na execução. A seguir, será detalhado o funcionamento da heurística KMB, juntamente com suas características.

Dado um grafo conexo não direcionado $G = (V, E, c)$, sendo c a função de ponderações do grafo. Supondo que um subconjunto $S \subseteq V$ seja dado, a heurística KMB utiliza de uma estratégia gulosa em etapas para construir uma árvore de Steiner mínima com conjunto de vértices terminais S. Seu funcionamento pode ser resumido em três fases principais: construção de um grafo completo a partir dos vértices terminais, obtenção de uma árvore geradora mínima e construção da Árvore de Steiner mínima. O pseudocódigo da heurística é apresentado a seguir.

Heurística KMB

Entrada: um grafo não direcionado $G = (V, E, c)$ e um conjunto de vértices terminais S ⊆ V .

Passo 1. Encontrar os caminhos mínimos entre $u\ e\ v, \forall\ u, v\ \in\ S$;

Passo 2. Construir um grafo completo não direcionado G_1, cujos vértices são os vértices terminais de S, atribuindo a cada aresta entre os vértices terminais $u\ e\ v$ o custo dos caminhos mínimos obtidos no Passo 1;

Passo 3. Encontrar a árvore geradora mínima de G_1, denominada MST_1. Se existir mais de uma árvore geradora mínima, escolha arbitrariamente uma;

Passo 4. Construir o subgrafo G_s, substituindo cada aresta de MST_1 pelos caminhos mínimos do grafo original G. Se existir mais de um caminho mínimo, escolha arbitrariamente um;

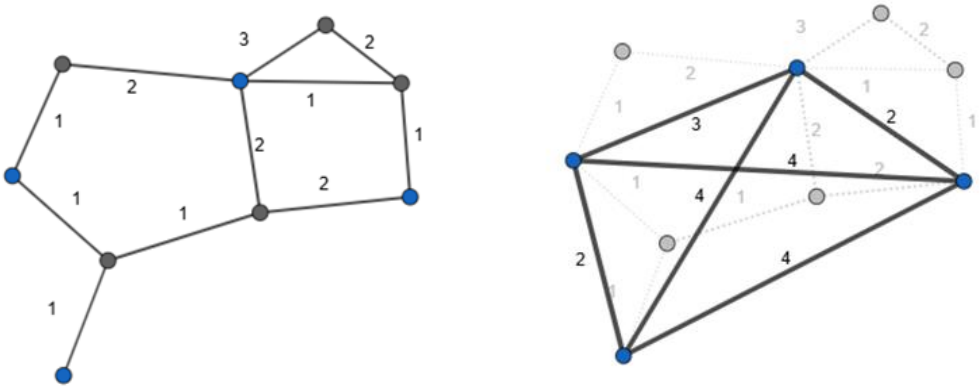
Passo 5. Encontrar a árvore geradora mínima de G_s, denominada T_s. Se existir mais de um caminho mínimo, escolha arbitrariamente um;

Passo 6. Construir a árvore de Steiner T_H de T_s deletando arestas em T_s, se necessário, de forma que todas as folhas de T_H sejam pontos terminais.

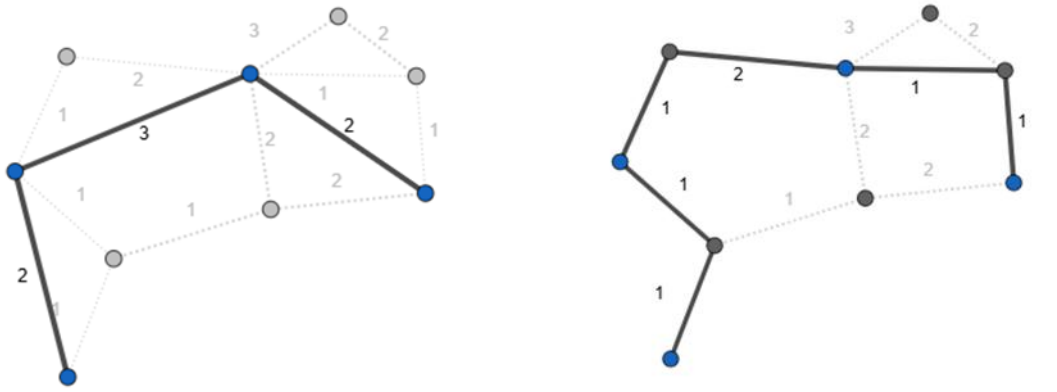
Saída: uma árvore de Steiner, T_H, para G e S.

Fonte: Adaptado de Kou *et.* al[1981]

Observa-se que, no Passo 3 e Passo 4 da heurística, critérios adicionais de escolha de árvore geradora mínima ou caminho mínimo não são apresentados. Na Figura 5.2 exemplificamos sua utilização, sendo os vértices azuis os vértices terminais. O resultado é uma árvore de Steiner mínima que abrange todos os vértices em S , contendo como folhas apenas vértices terminas, e incluindo alguns vértices de Steiner para obter um custo mínimo.



(a) Grafo Original G (b) G_1



(c) $MST1$ (d) árvore de Steiner T_H para G e S.

Figura 5.2:Exemplo de aplicação da Heurística KMB

Embora seja possível obter aproximações eficientes, a heurística KMB não possui uma prioridade quando há vários caminhos mínimos entre um mesmo par de vértices terminais, dessa forma, o custo obtido para a árvore de Steiner pode variar dependendo dos caminhos selecionados [Fujita *et al.* 2016], [Boudjelida e Lemouari 2021].

5.3 Variações da heurística KMB:

Fujita *et al.* [2016a, 2016b] propuseram alterações na heurística KMB, utilizando um parâmetro de controle para priorizar, em alguma medida, o uso de arestas que pertencem a mais caminhos mínimos entre os terminais. Para isto utilizaram o conceito de centralidade de intermediação de arestas, o qual considera que se uma aresta pertence a muitos caminhos mínimos entre pares de vértices no grafo, então ela possui uma alta centralidade. Portanto, a ideia é de que ao considerar, em conjunto ao custo associado a cada aresta, a centralidade de intermediação das arestas, a árvore de Steiner obtida possa apresentar um custo menor.

Neste trabalho, utilizou-se KMB_EB_1 e KMB_EB_2 para fazer referência as heurísticas apresentadas em Fujita *et al.* [2016a,2016b], respectivamente. Destaca-se que a primeira heurística foi utilizada posteriormente em [Boudjelida

e Lemouari 2021], também no contexto de melhorias do desempenho da heurística original KMB.

Os autores de Fujita *et al.* [2016a, 2016b] apresentaram dois métodos de construção, a partir da heurística KMB, para o problema da árvore de Steiner mínima usando centralidade de intermediação de arestas normalizada. A centralidade de intermediação normalizada de uma aresta é definida pela Equação (23) [Bröhl e Lehnertz 2022]:

$$CB(e) = \frac{\sum_{i \neq j} \frac{q_{ij(e)}}{q_{ij}}}{|V|(|V|-1)}, \quad (23)$$

onde $e \in E(G)$ é a aresta a qual se deseja obter a centralidade, $i, j \in V(G)$, $q_{ij(e)}$é o número de caminhos mínimos com extremos nos vértices i e j que passam pela aresta e, e q_{ij} é o número total de caminhos mínimos com extremos nos vértices i e j.

As novas heurísticas geram novos custos para todas as arestas do grafo G original, utilizando a centralidade de intermediação de aresta normalizada e um parâmetro de controle. A partir dos novos custos gerados, a determinação da árvore de Steiner mínima é realizada por meio da heurística KMB. A diferença entre estas duas heurísticas e a heurística KMB original baseia-se na entrada do grafo e na forma como consideram o parâmetro de controle.

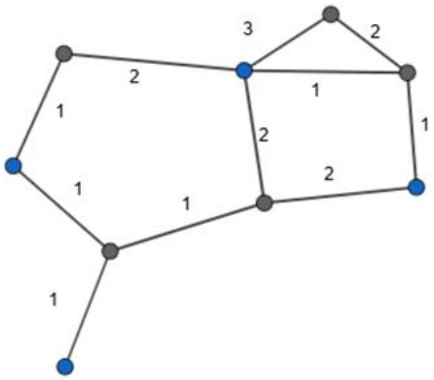
A heurística apresentada em Fujita *et al.* [2016a], aqui denominada KMB_EB_1, utiliza um percentual da aresta de custo máximo do grafo para calcular o valor do parâmetro de controle, aqui denominado de α, e a Equação (24) para o cálculo dos novos custos das arestas:

$$Custo_{novo}(e) = Custo_G(e) - \alpha\, CB(e), \quad (24)$$

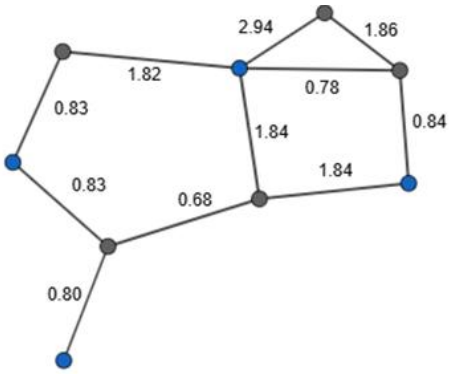
onde $Custo_G(e)$ é o custo original da aresta no grafo G, $CB(e)$ a centralidade de intermediação desta aresta e e α é um percentual do maior custo de uma aresta presente no grafo, por exemplo, em um grafo onde o maior custo de aresta seja igual a 20, e o percentual escolhido seja 30%, temos α= 6, logo, α é um número real positivo. Caso o percentual escolhido seja igual a zero, a heurística coincide

com KMB. Por outro lado, quando um percentual positivo é escolhido, é possível que a parcela $\alpha CB(e)$ seja maior que o custo original da aresta e, atribuindo um novo custo negativo a esta aresta. Em qualquer caso com $\alpha \neq 0$, quanto maior a centralidade de uma aresta, menor será o novo custo a ela associada. O novo custo pode assumir valores negativos, dependendo das porcentagens escolhidas; essa possibilidade será discutida em mais detalhes na Seção 5.4, no Tópico I.

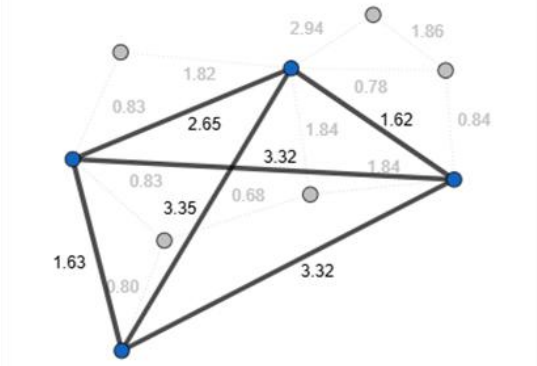
Na Figura 5.3 demonstramos os passos da heurística KMB_EB_1 para os pontos terminais marcados em azul, e utilizando o percentual de 30% para cálculo de α.



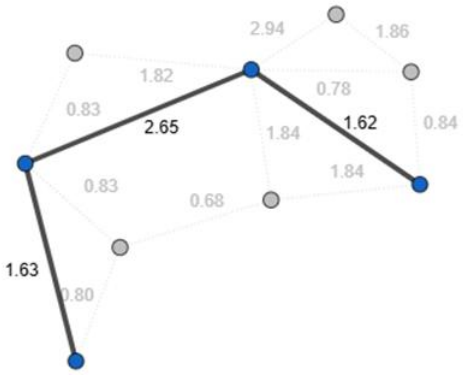
(a) Grafo original G



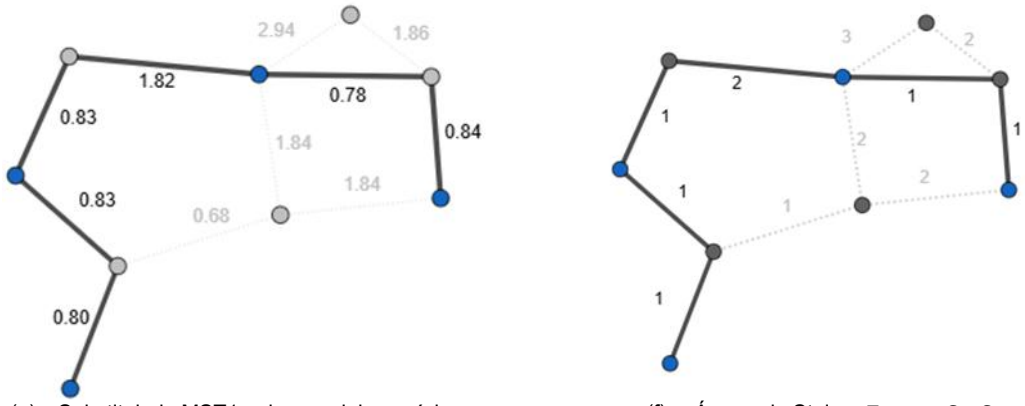
(b) Grafo com novos custos



(c) G_1



(d) MST1



(e) Substituindo MST1 pelos caminhos mínimos (f) Árvore de Steiner T_H para G e S

Figura 5.3:Exemplo de aplicação da heurística KMB_EB_1

A segunda heurística, KMB_EB_2, considera o parâmetro de controle de uma forma diferente, aqui denominado de β. Nela $\beta \in [0,1]$, sendo utilizado como um parâmetro de controle que prioriza, ora o custo original, ora a centralidade de intermediação. A Equação (25) é utilizada para o cálculo dos novos custos das arestas nesta heurística:

$$Custo_{novo}(e) = \beta\,|C(e)| + (1-\beta)\frac{1}{CB(e)}, \qquad (25)$$

Onde $|C(e)|$ é o custo da aresta e normalizado através da fórmula: $\frac{Custo_G(e)}{\gamma}$, onde γ é o custo máximo de aresta em G. Neste modo de cálculo, se β é próximo de 0, então o novo custo será muito influenciado por $\frac{1}{CB(e)}$, e quando a centralidade normalizada $CB(e)$ for alta, $\frac{1}{CB(e)}$ será pequeno, ocasionando um pequeno aumento no $Custo_{novo}(e)$. Logo, esta aresta será priorizada para obtenção da árvore de Steiner mínima. No caso $\beta = 1$, a centralidade de intermediação não é considerada para atribuição dos novos custos.

O passo a passo da heurística KMB_EB_2 é análogo ao da heurística KMB_EB_1, diferenciando apenas a forma como os novos custos das arestas são calculados e a forma como consideram o parâmetro de controle. As duas heurísticas apresentadas possuem o mesmo objetivo: obter um resultado melhor do que a heurística KMB na obtenção de uma árvore de Steiner mínima, ao considerar a centralidade de intermediação de arestas. Entretanto, as duas são formuladas com perspectivas diferentes. Enquanto a KMB_EB_1 favorece as

arestas mais centrais por meio da redução de seus custos, a KMB_EB_2 penaliza as menos centrais com um aumento de custo relativo, promovendo um caminho alternativo de priorização.

Embora ambas as heurísticas KMB_EB_1 e KMB_EB_2 difiram fundamentalmente pela forma como um parâmetro de controle é utilizado (α e β, respectivamente), maiores detalhes não são apresentados sobre possíveis intervalos de variação para estes parâmetros nos trabalhos originais e nenhuma discussão mais detalhada é feita a respeito da interferência deste parâmetro nos resultados obtidos. No entanto, este trabalho irá explorar possíveis intervalos de variação desses parâmetros e sua influência nas soluções obtidas.

Duas outras heurísticas foram propostas neste trabalho, utilizando as Equações (24) e (25) para calcular o novo custo das arestas, entretanto, alterando-se centralidade de aresta utilizada, que nomearemos de KMB_EC_1 e KMB_EC_2, respectivamente. Estas heurísticas utilizam o conceito de centralidade de aresta de proximidade (ao invés de centralidade de intermediação), onde uma aresta é considerada mais central quanto mais curtos forem os caminhos que conectam essa aresta a todas as outras arestas, [Bröhl e Lehnertz 2022]. A centralidade de proximidade normalizada da aresta e, com extremidades a e b, é definida pela Equação (26).

$$CC(e) = (|E| - 1)\frac{C_p(v_1) * C_p(v_2)}{C_p(v_1) + C_p(v_2)}, \qquad (26)$$

onde $e \in E(G)$ é a aresta que conecta os vértices $v_1, v_2 \in V(G)$ e $C_p(v_1)$ e $C_p(v_2)$ são as centralidades de proximidade não normalizadas dos vértices, dada por $C_p(u) = 1/\sum_i d(i,u)$, sendo $d(i,u)$ a distância do vértice i ao u, $\forall\, i \in V$.

5.4 Comparação dos resultados das variações da heurística KMB:

Os resultados apresentados neste trabalho consideram testes computacionais realizados a partir dos conjuntos de grafos B e C do repositório SteinLib [SteinLib 2024]. Esses testes foram conduzidos em um notebook com processador Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11GHz e 8 GB de

RAM instalada. O conjunto B do repositório é composto por 18 grafos esparsos, variando de 50 a 100 vértices e de 63 a 200 arestas. O conjunto C é formado por 20 grafos esparsos, cada um com 500 vértices, variando de 625 a 12500 arestas. Destaca-se que para as bases B e C, são conhecidos e validados os custos das árvores mínimas de cada grafo. Algumas considerações foram necessárias para a aplicação das heurísticas:

I- Embora em Fujita *et* al. [2016a] não tenham abordado essa restrição, a heurística KMB_EB_1 pode gerar custos de arestas negativos dependendo do percentual utilizado para o valor de α. Os autores utilizaram α como 30% do custo máximo de uma aresta do grafo, mas não apresentam uma justificativa para a escolha. Dessa forma, vários percentuais para cálculo de α foram testados neste trabalho. Por isso, especificamente nesta heurística, o algoritmo utilizado para obtenção dos caminhos mínimos foi o "Bellman Ford" [Bellman 1958], pois este obtém um caminho mínimo mesmo com custos negativos atribuídos as arestas, diferente de Dijkstra, e desde que não ocorram ciclos negativos (ciclo cujo somatório dos custos das arestas é negativo. Esses ciclos permitem que o caminho mais curto continue diminuindo indefinidamente ao passar por esse ciclo repetidas vezes).

II- Em KMB_EB_2, caso a centralidade de intermediação de uma aresta receba o valor zero na implementação, como aproximação, a fim de evitar esta aresta, iremos substituir o termo $\frac{1}{\mathrm{CB}(e)}$ por $\frac{1000}{\mu}$, onde μ é o menor valor das centralidades de intermediação das arestas do grafo. Essa substituição penaliza essas arestas adicionando um custo elevado, evitando que sejam escolhidas na construção da árvore de Steiner.

Considerando as heurísticas KMB_EB_1, KMB_EB_2, KMB_EC_1 e KMB_EC_2, três análises foram realizadas a partir dos testes computacionais no sentido de discutir o papel dos parâmetros de controle α e β e a eficácia das heurísticas, a saber:

a. Estudo da monotonicidade com relação ao parâmetro de controle: busca compreender o papel dos parâmetros de controle α e β nas heurísticas

KMB_EB_1 e KMB_EB_2, respectivamente, através do estudo do comportamento do parâmetro e seus respectivos resultados.

b. Comparação dos erros percentuais obtidos pelos diferentes métodos: analisa os resultados obtidos em ambas as heurísticas e os compara com o resultado obtido com a heurística tradicional KMB.

c. Comparação dos resultados obtidos com centralidade de intermediação e centralidade de proximidade: analisa os resultados utilizando as duas centralidades de arestas amplamente conhecidas na literatura, a de intermediação e a de proximidade.

a. Estudo da monotonicidade com relação ao parâmetro de controle:

Os autores de Boudjelida e Lemouari [2021] e Fujita *et al.*[2016a, 2016b] defendem que o valor do parâmetro de controle deve ser ajustado para um valor ótimo através de experimentos preliminares, pois, dependem da estrutura topológica do grafo. Dessa forma, é interessante investigar qual a influência do parâmetro de controle no custo final das árvores obtidas. Para isso, a primeira análise conduzida diz respeito a existência, ou não, de monotonicidade do custo das árvores obtidas com relação ao parâmetro de controle nas duas heurísticas apresentadas anteriormente, KMB_EB_1 e KMB_EB_2, utilizando o conjunto B.

Na Tabela 10 temos o custo das árvores utilizando a heurística KMB_EB_1, e valores de 10%, 20% e 30% para o peso de α, pois acima de 30% ciclos negativos são gerados no grafo, impossibilitando a obtenção de caminhos mínimos. É possível notar que o resultado do custo da árvore de Steiner mínima obtido por KMB_EB_1 não é monótono com relação ao parâmetro de controle, como representado na Tabela 10. No caso do grafo b18, mesmo com aumento do parâmetro de controle os resultados continuaram constantes, já no grafo b17 há redução de uma unidade no custo final a partir de aumento do parâmetro de controle.

Tabela 10: Teste da monotonicidade de KMB_EB_1 para diferentes valores de α

Grafo	Percentual p/ α	KMB_EB_1
b17	10%	135
b17	20%	135
b17	30%	134
b18	10%	223
b18	20%	223
b18	30%	223

Na Tabela 11, apresentamos o custo obtido a partir da heurística KMB_EB_2 com valores de β variando de 0,1 a 1, com acréscimos de 0,1 unidade, para analisar a monotonicidade com relação ao parâmetro de controle. Novamente, nota-se que não há comportamento monótono. Enquanto o grafo b12 apresenta algumas reduções no custo da árvore de Steiner mínima obtida conforme o parâmetro β é aumentado até 0,9, um valor maior é obtido quando β é escolhido como 0,9. Após isso, novamente há uma redução quando β é escolhido como 1. Para o grafo b02 o custo permanece constante conforme o valor de β aumenta, e novamente há um acréscimo no custo quando β é escolhido como 1. Destacamos que a escolha de $\beta = 1$ significa obter o mesmo resultado da heurística KMB.

Tabela 11: Teste da monotonicidade de KMB_EB_2 para diferentes valores de β.

Grafo	β	KMB_EB_2	Grafo	β	KMB_EB_2
b12	0,1	201	b02	0,1	86
b12	0,2	201	b02	0,2	86
b12	0,3	194	b02	0,3	86
b12	0,4	194	b02	0,4	86
b12	0,5	194	b02	0,5	86
b12	0,6	194	b02	0,6	86
b12	0,7	194	b02	0,7	86
b12	0,8	**194**	b02	0,8	86
b12	0,9	**195**	b02	0,9	86
b12	1	**174**	b02	1	90

Na Tabela 12 estão reunidos os resultados da heurística KMB_EB_1 para os valores de 30%, 50% e 70% aplicados no conjunto C, pois para percentuais maiores que 70% a heurística gera ciclos negativos. Neste caso, é possível notar que o valor do custo quando o percentual de 30% para o cálculo α (como utilizado por Fujita *et al.* [2016a]) não difere de outros percentuais, com exceção do grafo c04, cujo melhor resultado ocorre com o percentual de 70% para o cálculo de α, o que não foi mencionado pelos autores.

Tabela 12: heurística KMB_EB_1 para diferentes valores de α no conjunto C.

Grafo	Percentual p/ α	KMB_EB_1	Percentual p/ α	KMB_EB_1	Percentual p/ α	KMB_EB_1
c01	30%	88	50%	88	70%	88
c02	30%	144	50%	144	70%	144
c03	30%	782	50%	782	70%	782
c04	**30%**	**1110**	**50%**	**1109**	**70%**	**1105**
c05	30%	1598	50%	1598	70%	1598
c06	30%	60	50%	60	70%	60
c07	30%	115	50%	115	70%	115
c08	30%	531	50%	531	70%	531
c09	30%	725	50%	725	70%	725
c10	30%	1113	50%	1113	70%	1113

b. Comparação dos erros percentuais obtidos pelos diferentes métodos:

Com o objetivo de verificar os resultados obtidos pelas heurísticas KMB, KMB_EB_1 e KMB_EB_2 em um mesmo conjunto de dados, optou-se pela condução de testes utilizando os conjuntos B e C da base SteinLib. Os melhores valores de custo obtido das árvores de Steiner de cada método são apresentados na Tabela 13 e na Tabela 14. Essa comparação também utiliza o valor ótimo (OPT) indicado no repositório SteinLib. Dados os valores ótimos para cada caso, calculou-se o percentual de erro do melhor valor obtido de cada método através da Equação (27):

$$\epsilon_{\%} = \frac{|OPT - valor\ obtido|}{OPT} \qquad (27)$$

Tabela 13:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto B.

Grafo	OPT	KMB		KMB_EB_1		KMB_EB_2	
		valor	$\epsilon_{\%}$	Percentual para α	$\epsilon_{\%}$	β	$\epsilon_{\%}$
b01	82	82	0,00%	10% -30%	0%	1	0%
b02	83	90	8,43%	10% -30%	2,41%	0.1 -0.9	3,61%
b03	138	140	1,45%	10% -30%	0,72%	1	1,45%
b04	59	59	0,00%	10% -30%	0%	1	0%
b05	61	64	4,92%	10% -30%	1,64%	1	4,92%
b06	122	128	4,92%	10% -30%	4,10%	1	4,92%
b07	111	111	0,00%	10% -30%	0%	1	0%
b08	104	104	0,00%	10% -30%	0%	1	0%
b09	220	223	1,36%	10% -30%	0,45%	1	1,36%
b10	86	98	13,95%	10% -30%	13,95%	0.1 – 0.9	10,47%
b11	88	92	4,55%	10% -30%	2,27%	0.1 – 0.9	0%
b12	174	174	0,00%	10% -30%	0%	1	0%
b13	165	175	6,06%	10% -30%	6,06%	1	6,06%
b14	235	238	1,28%	10% -30%	0,85%	1	1,28%
b15	318	325	2,20%	10% -30%	1,26%	1	2,20%
b16	127	137	7,87%	10% -30%	7,87%	1	7,87%
b17	131	134	2,29%	30%	2,29%	1	1,53%
b18	218	223	2,75%	10% -30%	2,29%	1	3,67%
Média do erro			**3,65%**		**2,56%**		**2,74%**

Tabela 14:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto C

Grafo	OPT	KMB		KMB_EB_1		KMB_EB_2	
		valor	$\epsilon_{\%}$	Percentual para α	$\epsilon_{\%}$	β	$\epsilon_{\%}$
c01	85	88	3,53%	30% - 70%	3,53%	1	3,53%
c02	144	144	0,00%	30% - 70%	0,00%	1	0%
c03	754	782	3,71%	30% - 70%	3,71%	1	3,71%
c04	1079	1118	3,61%	70%	2,41%	1	3,34%
c05	1579	1603	1,52%	30% - 70%	1,20%	1	1,58%
c06	55	60	9,09%	30% - 70%	9,09%	1	9,09%
c07	102	115	12,75%	30% - 70%	12,75%	0.1 – 0.9	6,86%
c08	509	530	4,13%	30% - 70%	4,32%	1	4,13%
c09	707	728	2,97%	30% - 70%	2,55%	1	2,83%
c10	1093	1119	2,38%	30% - 70%	1,83%	1	2,10%
c11	32	37	15,63%	30% - 70%	15,63%	0.1-1	15,63%
c12	46	49	6,52%	30% - 70%	6,52%	0.1-0.9	2,17%
c13	258	278	7,75%	30% - 70%	5,04%	1	7,75%
c14	323	342	5,88%	30% - 70%	2,79%	1	6,50%
c15	556	573	3,06%	30% - 70%	0,90%	1	3,78%
c16	11	13	18,18%	30% - 70%	18,18%	1	9,09%
c17	18	19	5,56%	30% - 70%	5,56%	1	11,11%
c18	113	126	11,50%	30% - 70%	8,85%	0.1-0.9	8,85%
c19	146	159	8,90%	30% - 70%	6,85%	0.1-0.9	6,85%
c20	267	269	0,75%	30% - 70%	0,75%	1	0,75%
Média do erro			**6,37%**		**5,62%**		**5,48%**

É possível notar que em ambos os conjuntos de teste, as heurísticas que utilizam a centralidade de intermediação de aresta apresentam melhores resultados quando comparados os erros médios com a heurística KMB original.

Além disso, é possível notar que o percentual 30% para o cálculo de α na heurística KMB_EB_1 realmente obtém alguns valores ótimos de ambas as bases, entretanto, o grafo c04 do conjunto C obtém melhor resultado com o percentual de 70% para o cálculo de α, o que reforça a necessidade do teste de diferentes valores de α. Os limites máximos dos percentuais da heurística KMB_EB_1 foram obtidos testando a heurística, além desses valores ciclos negativos foram gerados, impossibilitando a obtenção de caminhos mínimos.

No Gráfico 1 e no Gráfico 2 comparamos o desempenho de ambas as heurísticas no conjunto B e conjunto C, respectivamente. Os dados representados estão na Tabela 15 e na Tabela 16. É possível notar que mesmo que a heurística KMB_EB_1 tenha um menor erro médio para o conjunto B, é a heurística KMB_EB_2 no conjunto B que obtém mais vezes o melhor valor, atingindo 6 vezes o erro zero com relação ao valor ótimo possível, de acordo ao indicado no repositório. Além disso, a heurística KMB_EB_1 apresenta um raio maior de erro, obtendo um valor máximo de erro de aproximadamente 14% no conjunto B, e 18% no conjunto C. Entretanto, não é possível inferir qual heurística que utiliza a centralidade de aresta será melhor em um contexto geral.

Gráfico 1: Comparação dos erros percentuais obtidos para as duas heurísticas com centralidade e a heurística KMB tradicional aplicadas ao conjunto B

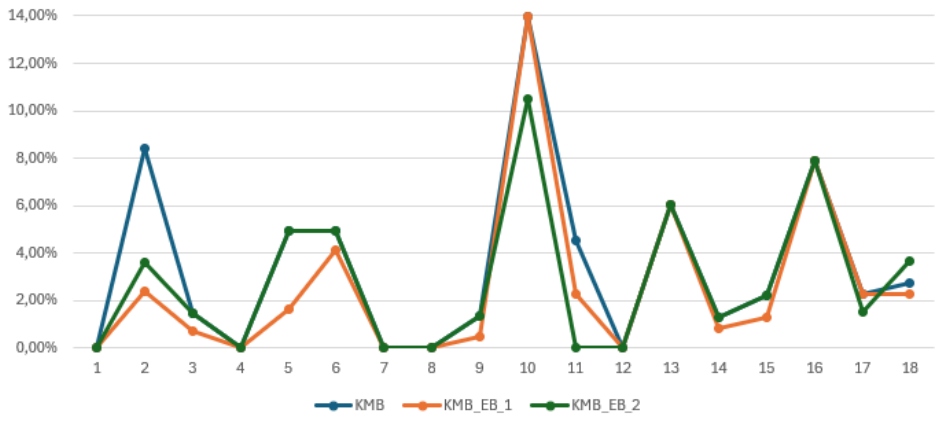
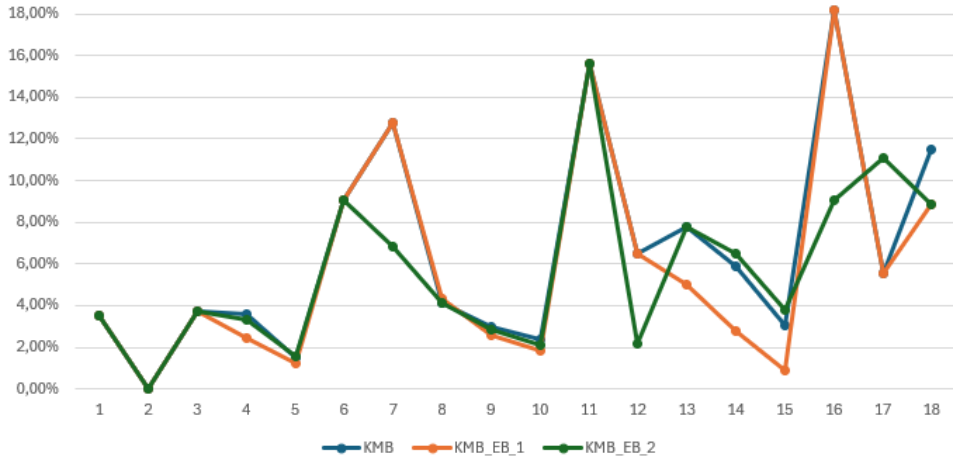


Gráfico 2: Comparação dos erros percentuais obtidos para as duas heurísticas com centralidade e a heurística KMB tradicional aplicadas ao conjunto C



c. Comparação dos resultados obtidos com centralidade de intermediação e centralidade de proximidade.

As simulações realizadas com a centralidade de aresta de intermediação foram também executadas com a centralidade de arestas de proximidade, considerando o mesmo conjunto de dados e valores possíveis para o parâmetro de controle. Denominamos como KMB_EC_1 e KMB_EC_2 as heurísticas que utilizam a centralidade de arestas por proximidade que são análogas ás heurísticas KMB_EB_1 e KMB_EB_2, respectivamente, ou seja, consideram mesmo conjunto de dados e valores possíveis para o parâmetro de controle.

Na Tabela 15 e na Tabela 16 organizamos os resultados obtidos no conjunto B e conjunto C, respectivamente, comparando a eficácia das heurísticas utilizando as duas medidas de centralidade.

Tabela 15:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto B

Grafo	OPT	KMB_EB_1		KMB_EB_2		KMB_EC_1		KMB_EC_2	
		Percentual para α	$\epsilon_{\%}$	β	$\epsilon_{\%}$	Percentual para α	$\epsilon_{\%}$	β	$\epsilon_{\%}$
b01	82	10%-30%	0%	1	0%	10%- 40%	0%	1	0%
b02	83	10%-30%	2,41%	0.1-0.9	3,61%	10%- 40%	2,41%	1	8,43%
b03	138	10%-30%	0,72%	1	1,45%	10%- 40%	0,72%	1	1,45%
b04	59	10%-30%	0%	1	0%	10%- 40%	0%	1	0%
b05	61	10%-30%	1,64%	1	4,92%	10%- 40%	1,64%	1	4,92%
b06	122	10%-30%	4,10%	1	4,92%	10%- 40%	4,10%	1	4,92%
b07	111	10%-30%	0%	1	0%	10%- 40%	0%	1	0%
b08	104	10%-30%	0%	1	0%	10%- 40%	0%	1	0%
b09	220	10%-30%	0,45%	1	1,36%	40%	0%	1	1,36%

b10	86	10%-30%	13,95%	0.1–0.9	10,47%	40%	9,30%	1	13,95%
b11	88	10%-30%	2,27%	0.1–0.9	0%	10%- 40%	2,27%	0.8-1	4,55%
b12	174	10%-30%	0%	1	0%	10%- 40%	0%	1	0%
b13	165	10%-30%	6,06%	1	6,06%	10%- 40%	6,06%	1	6,06%
b14	235	10%-30%	0,85%	1	1,28%	10%- 40%	0,85%	1	1,28%
b15	318	10%-30%	1,26%	1	2,20%	10%- 40%	1,26%	1	2,20%
b16	127	10%-30%	7,87%	1	7,87%	10%- 40%	7,87%	1	7,87%
b17	131	30%	2,29%	1	1,53%	10%- 30%	1,53%	1	1,53%
b18	218	10%-30%	2,29%	1	3,67%	10%- 40%	2,29%	1	3,67%
Média do erro			**2,56%**		**2,74%**		**2,24%**		**3,46%**

Tabela 16:Comparação dos erros percentuais obtidos para cada heurística aplicada ao conjunto C

Grafo	**OPT**	**KMB_EB_1**		**KMB_EB_2**		**KMB_EC_1**		**KMB_EC_2**	
		Percentual para α	$\epsilon_{\%}$	$\boldsymbol{\beta}$	$\epsilon_{\%}$	**Percentual para α**	$\epsilon_{\%}$	$\boldsymbol{\beta}$	$\epsilon_{\%}$
c01	85	30%-70%	3,53%	1	3,53%	10%-30%	3,53%	1	3,53%
c02	144	30%-70%	0,00%	1	0%	10%-30%	0,00%	1	0%
c03	754	30%-70%	3,71%	1	3,71%	10%-30%	3,58%	1	3,71%
c04	1079	70%	2,41%	1	3,34%	10%-30%	2,87%	1	3,34%
c05	1579	30%-70%	1,20%	1	1,58%	10%-30%	1,20%	1	1,58%
c06	55	30%-70%	9,09%	1	9,09%	10%-30%	9,09%	0.9	0%
c07	102	30%-70%	12,75%	0.1-0.9	6,86%	10%-30%	12,75%	1	12,75%
c08	509	30%-70%	4,32%	1	4,13%	10%-20%	4,52%	1	4,13%
c09	707	30%70%	2,55%	1	2,83%	10%-30%	2,69%	1	2,83%
c10	1093	30%-70%	1,83%	1	2,10%	10%-30%	1,74%	1	2,10%
c11	32	30%-70%	15,63%	0.1-1	15,63%	10%-20%	15,63%	1	15,63%
c12	46	30%-70%	6,52%	0.1-0.9	2,17%	10%-20%	6,52%	1	6,52%
c13	258	30%-70%	5,04%	1	7,75%	10%-20%	5,04%	1	7,75%
c14	323	30%-70%	2,79%	1	6,50%	30%	1,55%	1	6,50%
c15	556	30%-70%	0,90%	1	3,78%	10%-20%	1,08%	1	3,78%
c16	11	30%-70%	18,18%	1	9,09%	10%-30%	18,18%	1	9,09%
c17	18	30%-70%	5,56%	1	11,11%	10%	11,11%	1	11,11%
c18	113	30%-70%	8,85%	0.1-0.9	8,85%	10%-30%	8,85%	1	11,50%
c19	146	30%-70%	6,85%	0.1-0.9	6,85%	10%-30%	6,85%	1	8,90%
c20	267	30%-70%	0,75%	1	0,75%	10%	0,75%	1	0,75%
Média do erro		**5,62%**		**5,48%**		**5,88%**		**5,78%**	

Observa-se que para o conjunto B, a heurística KMB_EC_1 apresentou menor média de erro, ou seja, a troca pela medida de centralidade de proximidade de aresta melhorou os resultados das árvores de Steiner obtidas. Entretanto, para o conjunto C a heurística KMB_EB_2 continua obtendo melhores resultados.

Os Gráfico 3 e o Gráfico 4 comparam o desempenho das heurísticas utilizando centralidade de aresta no conjunto B e conjunto C, respectivamente. Os dados representados, encontram-se na Tabela 17 e Tabela 18. É possível notar que a heurística KMB_EC_2 possui um maior raio de erro para o conjunto

B, enquanto a heurística KMB_EC_1 possui o menor raio de erro. Entretanto para o conjunto C, o maior raio de erro é para a heurística KMB_EC_1.

Gráfico 3: Erros percentuais obtidos para as quatro heurísticas com centralidade aplicadas ao conjunto B.

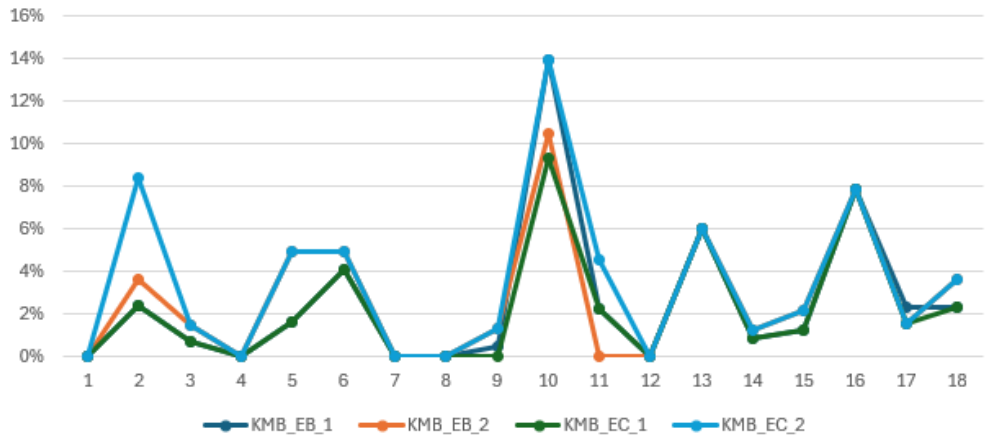
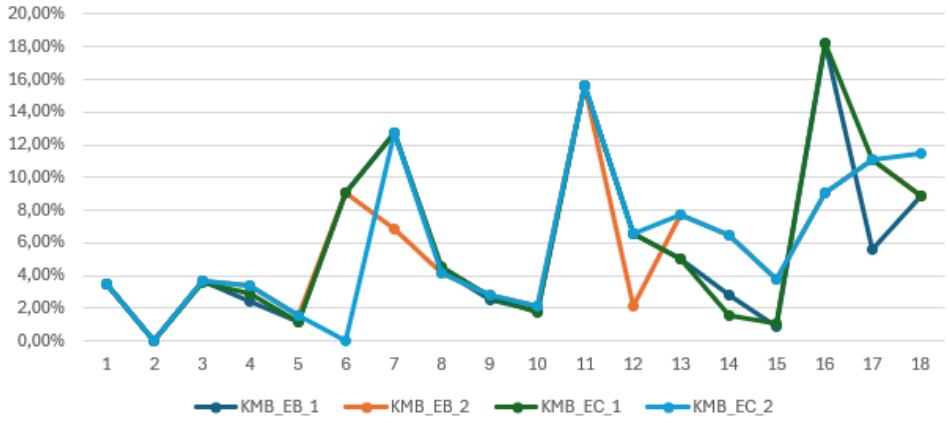


Gráfico 4: Erro percentuais obtidos para as quatro heurísticas com centralidade aplicadas ao conjunto C.



As heurísticas KMB_EB_1 e KMB_EB_2, propostas por Fujita *et* al. [2016a, 2016b], surgem como alternativas para melhorar os resultados da KMB ao incorporar a centralidade de intermediação de arestas como um parâmetro de controle. Neste estudo, utilizando os conjuntos B e C do SteinLib, os experimentos demonstraram que a KMB_EB pode, de fato, proporcionar árvores de Steiner de menor custo em comparação à heurística KMB tradicional.

Destaca-se, a partir de análises das bases B e C do repositório SteinLib, que não há relação de monotonicidade entre o parâmetro de controle e o custo

da rede obtida pelas heurísticas KMB_EB1 e KMB_EB2, com os resultados variando de acordo com topologia da rede. Além disso, ao comparar o uso da centralidade de intermediação de arestas com a centralidade de proximidade de arestas, observa-se que a eficácia das heurísticas também depende do conjunto de dados analisado. A KMB_EC_1, que utiliza centralidade de proximidade, apresentou melhores resultados no conjunto B, enquanto a KMB_EB_2 manteve-se superior no conjunto C, apresentando um menor raio de erro médio. Esses resultados indicam que não há uma solução universalmente melhor; a escolha da heurística e da centralidade de arestas pode depender das características específicas da rede em questão.

6 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo central investigar o conceito de centralidade de arestas em grafos, analisando sua relação com a centralidade de vértices e suas aplicações no Problema da Árvore de Steiner Mínima, por meio de variações da heurística KMB. A partir dessa proposta, duas linhas principais de investigação foram conduzidas: (i) uma análise experimental da associação entre vértices e arestas mais centrais, e (ii) a formulação e avaliação computacional de novas heurísticas baseadas em centralidade de arestas.

Inicialmente, foram apresentados os fundamentos da Teoria dos Grafos e as principais medidas de centralidade, com destaque para as de intermediação e proximidade, tanto para vértices quanto para arestas. A partir disso, examinou-se se as arestas mais centrais em um grafo, conforme medida de intermediação, estão associadas aos vértices também mais centrais. Essa análise revelou uma diversidade de padrões: foram encontrados grafos nos quais não há qualquer relação entre os vértices mais centrais e as arestas mais centrais, outros com associação parcial e, ainda, casos com associação total (os dois vértices mais centrais do grafo são incidentes à aresta mais central).

Dentre os resultados mais significativos, destaca-se a análise teórica detalhada sobre grafos vassoura dupla equilibrados. Nesses grafos, foi possível determinar formalmente as expressões das centralidades de intermediação e de proximidade para todos os vértices e arestas, permitindo identificar com os elementos mais centrais.

Na segunda linha de pesquisa, foram avaliadas quatro variações da heurística KMB: KMB_EB_1, KMB_EB_2, KMB_EC_1 e KMB_EC_2, sendo as duas últimas inéditas. Todas essas variações incorporam a centralidade de arestas como critério principal na escolha das ligações durante a construção da árvore de Steiner. Os experimentos computacionais realizados com instâncias do repositório *SteinLib* mostraram que, em diversos casos, essas variações apresentaram desempenhos competitivos em relação à heurística original, destacando a relevância da centralidade de arestas como um parâmetro promissor na construção de árvores de Steiner. Também foram identificadas

limitações relacionadas ao parâmetro de controle dessas heurísticas, cuja influência no desempenho foi discutida com base em testes sistemáticos.

Do ponto de vista científico, os resultados obtidos indicam que a centralidade de arestas é uma ferramenta subexplorada, mas potencialmente útil, tanto para a compreensão estrutural de grafos quanto para o aprimoramento de algoritmos de otimização combinatória. A análise empírica das relações entre centralidades de vértices e arestas, aliada à aplicação prática dessas medidas em heurísticas, representa uma contribuição original que amplia o escopo de uso dessas métricas no contexto da pesquisa operacional.

Como limitações deste trabalho, destacam-se os custos computacionais associados ao cálculo de centralidades em instâncias maiores, que restringiram o número de experimentos realizados com grafos muito densos ou com grande número de terminais. Além disso, as variações propostas para a heurística KMB ainda não foram exploradas em contextos aplicados mais amplos, como redes de comunicação, transporte ou biológicas, o que constitui uma oportunidade relevante para trabalhos futuros.

Nesse sentido, propõe-se como continuidade a reprodução do estudo de associação entre os vértices mais centrais e a aresta mais central de um grafo utilizando, por exemplo, centralidade de proximidade. Além disso, propõe-se o estudo da aplicação dessas heurísticas em problemas práticos com estrutura de rede, bem como a análise da adaptação dessas estratégias para o *grafo linha*, a fim de investigar como a centralidade de arestas pode se refletir em propriedades estruturais alternativas.

Um grafo linha é denotado por $L(G)$ e representa a adjacência entre as arestas do grafo G. Dado um grafo G, o grafo linha $L(G)$ é obtido a partir de G da seguinte forma:

a. Cada vértice de $L(G)$ representa uma aresta em G;
b. Dois vértices de $L(G)$ são adjacentes se e somente se suas arestas correspondentes compartilham um mesmo vértice em G, ou seja, são adjacentes em G.

Por fim, destaca-se que parte dos resultados deste trabalho foi aceita para publicação no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2024), o que reforça a relevância acadêmica da pesquisa desenvolvida. Espera-se que este estudo contribua não apenas para a literatura sobre o Problema da Árvore de Steiner, mas também para o avanço do entendimento sobre a importância estrutural das arestas em grafos.

REFERÊNCIAS

Bellman, R. (1958). On a Routing Problem. Quarterly of Applied Mathematics 16(1):87-90

Biggs, N. (1993). *Algebraic Graph Theory*. Cambridge University Press.

Boudjelida, A. e Lemouari, A. (2021). Solving the Delay-Constrained Least-Cost routing problem using Tabu Search with Edge Betweenness. *E3S Web of Conferences*, 229:01009.

Boaventura Netto, P. O., e Jurkiewicz, S. (2017). Grafos: Introdução e Prática. Blucher.

Borba, E. M., e Trevisan, V. (2014). Medidas de Centralidade em Grafos e Aplicações em Redes de Dados. Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, 2(1). CMAC-Sul, Curitiba-PR, 2014. DOI: 10.5540/03.2014.002.01.0088010088-1© 2014 SBMAC.

Bröhl, T., & Lehnertz, K. (2019). Centrality-based identification of important edges in complex networks. *Chaos, 29*(3), 033115.

Bröhl, T. e Lehnertz, K. (2022). A straightforward edge centrality concept derived from generalizing degree and strength. *Scientific Reports*, 12:4407.

Dijkstra, E. W., (1959). A Note on Two Problems in Connexion with Graphs. Numerische Mathematik 1:269-271.

Dreyfus, S. E.; Wagner, R. A. (1972). The Steiner problem in graphs. Networks, v. 1, p. 195-207

Freeman, L. (1977) Sociometry 40, 35–41.

Freeman, “Centrality in Social Networks Conceptual Clarification,” vol. 1, no. 1968, pp. 215–239, 1978

Ford, L. R. Jr. (1956). Network Flow Theory. The RAND Corporation, Paper P-923.

Fujita, M. *et al.*. (2016a). An effective construction algorithm for the Steiner tree problem based on edge betweenness. In *Journal of Signal Processing, v*ol. 20. No 4, PP 145 – 148.

Fujita, M. *et al.*. (2016b). A construction method for the Steiner tree problem using betweenness centrality. In *Proceedings of the 2016 International Symposium on Nonlinear Theory and Its Applications*, NOLTA2016. Yugawara, Japão.

Girvan, M., & Newman, M. E. J. (2001). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821-7826. https://doi.org/10.1073/pnas.122653799

Goldbarg, M., & Goldbarg, E. (2012). Grafos: Conceitos, algoritmos e aplicações. Elsevier Editora Ltda.

Kou *et.*al (1981). A fast algorithm for Steiner tress. In *Acta Informatica* 15. 141-145.

Kruskal, J. J. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. ACM, 7(1):48–50.

Kucharczuk, N., Wąs, T., & Skibski, O. (2022). PageRank for Edges: Axiomatic Characterization. In Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22).

Mocanu, D. C., Exarchakos, G., & Liotta, A. (2018). Decentralized Dynamic Understanding of Hidden Relations in Complex Networks. Scientific Reports, 8(1), Article 19356. DOI: 10.1038/s41598-018-19356-4.

Prim, R. C. (1957). Shortest connection networks and some generalizations. Bell System Technical Journal, 36:1389–1401.

De Meo, P., Ferrara, E., Fiumara, G., & Ricciardello, A. (2012). A novel measure of edge centrality in social networks. Knowledge-Based Systems, 30, 136-150.

Sabidussi, G. (1966). The centrality index of a graph. Psychometrika, 31:581–603.

Shaw, M. E. (1964). Communication networks, chapter 4. Academics Press, INC, New York, 1 edition.