XX ENCONTRO BAIANO DE EDUCAÇÃO MATEMÁTICA
IX FÓRUM BAIANO DAS LICENCIATURAS EM MATEMÁTICA
01 A 04 DE AGOSTO DE 2023
PAULO AFONSO - BA
ISSN 2175-1668

Ilhas em redes: um método para identificação de temáticas

Eixo Temático: Ensino Superior

Antonio Diêgo da Luz Silva. UEFS. antoniodiegoluz@gmail.com;
Bruno Rios Souza. UEFS. bruno.rios18@outlook.com.br;
Marcos Grilo. UEFS. grilo@uefs.br
Inácio de Sousa Fadigas. UEFS. fadigas@uefs.br

RESUMO

A Ciência das Redes ocupa-se da investigação do padrão das relações entre atores/objetos/entes dentro de um contexto. Os atores/objetos/entes são representados por vértices e as suas relações, por arestas ou arcos, constituindo uma rede. Neste minicurso, apresentaremos uma técnica de determinação de ilhas em redes para a identificação de temáticas. Este minicurso está dividido em duas etapas, uma teórico-expositiva e outra prática, na qual apresentaremos o método de obtenção de ilhas através do Pajek. O uso de ilhas em redes pode contribuir no desenvolvimento de trabalhos de conclusão de curso que objetivam identificar temáticas relevantes em contextos diversos, a exemplo de cursos de graduação, programas de pós-graduação, periódicos científicos, dentre outros.

Palavras-chave: Ilhas. redes semânticas. temáticas. trabalho de conclusão de curso.

INTRODUÇÃO

De acordo com Newman (2010), uma rede em sua forma mais simples, é uma coleção de pontos que se conectam por meio de linhas. Redes são representadas matematicamente por grafos, estrutura composta de vértices (pontos) e arestas (linhas). Uma rede pode ser classificada como ponderada ou valorada, quando cada aresta possui um peso ou valor, e não ponderadas, caso contrário. Em uma rede ponderada, existe uma função que fornece um valor a cada aresta.

Sub-redes podem ser obtidas a partir de uma rede ponderada e não dirigida (as arestas não possuem direção), removendo-se as arestas cujos valores estejam abaixo de um valor especificado (BATAGELJ e MRVAR, 2003; ZAVERSNIK E BATAGELJ , 2004). Tais sub-redes são denominadas ilhas. Sendo assim, Zaversnik e Batagelj (2004) definem ilhas como uma sub-rede conectada com número de vértices definido em um intervalo [k , K], com uma forte coesão interna relativa aos vértices vizinhos.

ISSN 2175-1668

Neste minicurso, mostraremos como utilizar o conceito de ilhas em redes semânticas para a identificação de temáticas. Conforme Grilo et al. (2021), a determinação de ilhas em redes semânticas possibilita a identificação de conjuntos de palavras cujas relações entre si são relevantes considerando-se o peso de suas arestas. Existe uma família de definições de redes semânticas, conforme Pereira et al. (2022), cujas aplicações aparecem em áreas como Ciência da Computação, Psicologia, Inteligência Artificial, Física, Engenharia do Conhecimento, Pesquisa Operacional, Engenharia Biomédica, Neurocomputação, Matemática, Gestão do Conhecimento, Engenharia de Alimentos, Astronomia, Poesia, Ciência Cognitiva, Fotogrametria, Educação, Podologia, Ciência da Informação, Sistemas Complexos, Saúde, Ciências das Redes, Ciências sociais, Sustentabilidade, Sensoriamento Remoto e Comunicação. Neste minicurso, seguiremos a definição de Grilo et al. (2017): uma rede semântica como um sistema de representação do conhecimento imbuído de uma intenção de funcionalidade, onde os vértices representam palavras, conceitos e entidades e as arestas representam significados semânticos dos seus relacionamentos.

Ilhas em redes semânticas podem ser usadas para a identificação de temáticas, conforme Santos e Grilo (2020) e Grilo et al. (2021). Santos e Grilo (2020) propuseram-se a identificar as temáticas mais relevantes na Licenciatura em Matemática da Uefs a partir de títulos de trabalho de conclusão de curso, enquanto Grilo et al (2021) em títulos de dissertações de mestrado em Matemática. Ainda segundo Grilo et al. (2021), a identificação de temáticas por meio de ilhas em redes é importante pois não se limita à investigação das palavras mais frequentes, pois captura as relações relevantes entre elas. Desta forma, o minicurso na qual ofertamos é uma ferramenta interessante para o desenvolvimento de investigações no campo da Educação Matemática que possam culminar em trabalhos de conclusão de curso.

Este minicurso está dividido em duas etapas: 1) Compreensão do conceito de ilhas em redes; 2) Determinação de ilhas através do Pajek.

ISSN 2175-1668

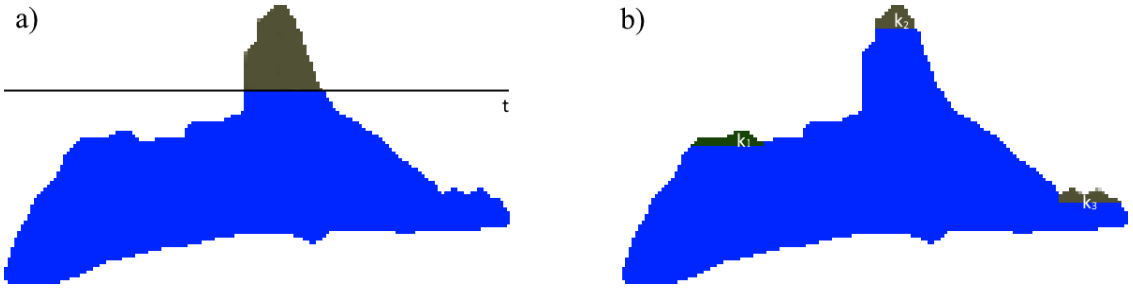
ILHAS: UMA COMPREENSÃO DO CONCEITO

Imagine a ilha de Fernando de Noronha ou a de Madagascar na África. Conforme é consabido, uma ilha é um pedaço de terra cercado por águas em toda a sua extensão territorial. No contexto da Ciência das Redes, ilhas referem-se a um grupo de vértices com forte coesão interna em relação aos seus vizinhos. De uma maneira análoga às ilhas geográficas, pode-se imaginar que o peso das arestas está associado ao nível da água. Quando mais o nível da água se eleva, ilhas podem diminuir de tamanho ou desaparecer. Ou seja, mais vértices podem ficar "submersos" pois as arestas cujos pesos estão abaixo de um certo limiar são eliminadas. Consequentemente, os vértices isolados são eliminados. Se o nível da água diminuir, as ilhas podem aumentar a sua extensão territorial e outras podem surgir. Analogamente, menos vértices ficam "submersos". Como buscamos relações relevantes, o nosso interesse aqui é o de elevar o nível da água até um ponto crítico.

A abordagem por ilhas pode ser feita de dois modos: o nível da água sobe de maneira constante; o nível da água sobe de maneira variável. Quando o nível é constante, definimos um limiar t e deletamos todas as arestas com peso menor do que t; eliminamos os vértices isolados após a deleção das arestas e, assim, obtemos as ilhas, como se fosse a água subindo de forma uniforme entre os picos de uma colina. Agora imagine que a água não vem de baixo e sim de cima, como em um dia de chuva forte e duradoura. A enxurrada, ao cair entre os montes, desce pelos picos e se concentra nas depressões existentes entre elas, se acumulando em níveis diferentes. Neste caso, não dependerá necessariamente de qual monte está mais em cima, mas sim do tamanho da porção de terra que não ficará submersa pela enxurrada. Para este método de determinação de ilhas, denominamos de nível variável. A Figura 1a) ilustra o método de determinação de ilhas em nível constante e a Figura 1b), a de nível variável.

ISSN 2175-1668

Figura 1: Determinação de ilhas. Em a), nível constante; em b), nível variável.



Fonte: Os autores.

Neste minicurso, apresentaremos a técnica de nível constante para determinação de ilhas. Esta foi a técnica utilizada por Santos e Grilo (2020) e Grilo et al. (2021) para a identificação de temáticas.

DETERMINAÇÃO DE ILHAS

Para a determinação de ilhas em redes ponderadas não dirigidas usando a técnica de nível constante, usaremos o programa gratuito Pajek, disponível em http://mrvar.fdv.uni-lj.si/pajek. O comando no Pajek é o seguinte:

Network > Create New Network > Transform > Remove > Lines with Value > lower than

Como exercício, o leitor pode encontrar ilhas na rede apresentada na Figura 2 a). As Figuras 2 b), 2 c) e 3 c) mostram as ilhas obtidas no nível constante 1, 2 e 3, respectivamente. O retângulo azul claro ilustra a subida do nível da água. Note que no nível 1, foram obtidas duas ilhas, uma com 8 vértices e outra com 2 vértices. A eliminação de arestas de peso 1 resultou em 4 vértices isolados, que foram excluídos da rede. O aumento do nível para 2, resulta em uma ilha com 5 vértices e para o nível 3, uma ilha com 3 vértices.

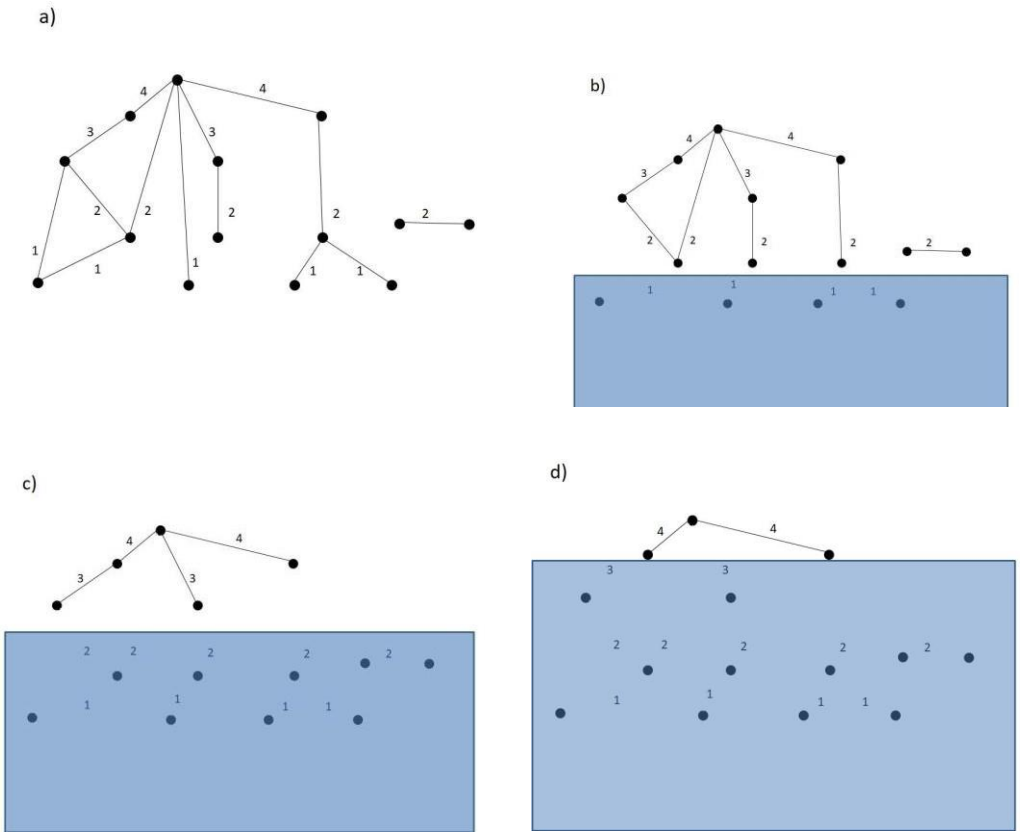
Para a identificação de temáticas, conforme Santos e Grilo (2020) e Grilo et al. (2021), pode-se construir redes semânticas usando as técnicas desenvolvidas por Caldeira (2005) e Fadigas et al. (2009). Em geral, a maior componente de uma rede semântica

ISSN 2175-1668

baseada em títulos possui pelo menos 90% do tamanho da rede (FADIGAS et al., 2009; GRILO et al., 2017). Uma componente é uma sub-rede maximal conectada. Cada par de vértices de uma componente possui um caminho que os conecta. Desta forma, conforme Santos e Grilo (2020) e Grilo et al. (2021), podemos escolher a maior componente da rede semântica para a aplicação do método de obtenção de ilhas. Para extrair a maior componente de uma rede, dois comandos devem ser efetuados:

1º comando: *Network* > *Create Partition* > *Components* > *Weak*

Figura 2: Ilhas em redes. Em a), uma rede ponderada não dirigida. Em b), ilhas obtidas no nível 1. Em c), ilha obtida no nível 2. Em d), ilha obtida no nível 3.



Fonte: Os autores.

2º comando: *Operations* > *Network + Partition* > *Extract* > *Subnetwork induced by Union of Selected Clusters*

ISSN 2175-1668

Após o uso do 2º comando, deve-se usar a própria saída do Pajek para verificar a distribuição dos tamanhos das componentes e identificar uma possível maior componente.

O primeiro comando particiona a rede em componentes. O segundo comando extrai a componente selecionada. Este método é o mesmo usado para extrair as ilhas no nível desejado. Outra questão associada a ilhas está na determinação do nível crítico. Santos e Grilo (2020) estabelece um nível crítico a partir da análise das palavras com a maior centralidade. Grilo et al. (2021) determina o nível crítico a partir da distribuição dos pesos das arestas e da quantidade de vértices da maior ilha em relação às demais. Convidamos o leitor a encontrar a(s) ilha(s) da rede semântica baseadas em títulos de TCC no nível 11, construídas por Santos e Grilo (2020).

MATERIAIS E MÉTODOS

Este minicurso tem como público alvo estudantes de licenciatura em Matemática e cursos afins. A duração do minicurso é de 3h e o número de vagas é 20. Para a realização deste minicurso, serão utilizados os seguintes materiais: primeira atividade – lápis, borracha e papel; segunda e terceira atividade - Computador ou notebook dispondo dos programas Pajek, Notepad++ (ou Bloco de Notas) e Microsoft Excel (ou Origin Lab).

Seria adequado que a realização deste minicurso ocorresse em um laboratório de informática. A metodologia deste minicurso pode ser dividida em dois momentos: a) teórico-expositiva, prezando pela interação de seus participantes com os ministrantes; b) prática, na qual os participantes realizarão as atividades sob a monitoria dos ministrantes. Na parte teórico-expositiva, contextualizaremos o minicurso, realizando uma exposição teórica sobre o conceito de redes e ilhas. Mostraremos como é o processo teórico para a determinação de ilhas. Na parte prática, inicialmente, os participantes receberão uma atividade impressa e serão convidados a determinar ilhas manualmente. Em seguida, disponibilizaremos redes em formato .net para que os participantes encontrem ilhas a partir do Pajek.

ISSN 2175-1668

CONSIDERAÇÕES FINAIS

Neste minicurso apresentamos a técnica de obtenção de ilhas em redes utilizando o software Pajek. O uso de ilhas para detectar as temáticas mais importantes em uma rede é uma análise complementar e diferente de centralidades, como por exemplo, o grau das palavras. Pode acontecer de palavras estarem ligadas a outras que não tem maior centralidade mas a relevância pode ser avaliada por intermédio do peso de suas arestas. Por fim, entendemos que a técnica mostrada neste minicurso pode ser utilizada por graduandos em Matemática como parte dos procedimentos metodológicos para trabalhos de conclusão de curso.

REFERÊNCIAS

BATAGELJ, V.; MRVAR, A. **Analysis of large networks Islands**. . Em: PROCEEDINGS OF PAJEK WORKSHOP AT XXVI SUNBELT CONFERENCE. 2003.

CALDEIRA, S. M. G. **Caracterização de Signos Linguísticos: um modelo baseado no aparelho psíquico de Freud**. Dissertação. Faculdade Visconde de Cairu, 2005.

FADIGAS, I. DE S. et al. Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos de divulgação em educação matemática. **Educação Matemática Pesquisa,** v. 11, n. 1, 2009.

GRILO, M. et al. Identificação de temáticas em programas de mestrado em matemática: uma abordagem por ilhas de redes. Em: **Construção do conhecimento em organizações na perspectiva da modelagem de processos industriais e sistemas complexos**. Compartilhando Saberes. São Paulo: Pimenta Cultural, 2021. v. 3.

GRILO, M. et al. Robustness in semantic networks based on cliques. **Physica A: Statistical Mechanics and its Applications**, v. 472, p. 94–102, 2017.

NEWMAN, M. E. J. **Networks: An Introduction**. Oxford University Press, 2010.

PEREIRA, H. B. DE B. et al. Systematic review of the “semantic network” definitions. **Expert Systems with Applications**, v. 210, p. 118455, 2022.

ISSN 2175-1668

SANTOS, V. C. DOS; GRILO, M. Identificação de temáticas de trabalhos de conclusão de curso por meio de redes semânticas. **Revista Paranaense de Educação Matemática; Vol. 9, No 20 (2020)**, 2020.

ZAVERŠNIK, M.; BATAGELJ, V. **Islands**. Slides from Sunbelt XXIV, Portoroz, Slovenia. **Anais**...2004.