



SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

## REDES DE CLIQUES ALEATÓRIAS: UMA APLICAÇÃO A TÍTULOS DE DISSERTAÇÕES DO PROFMAT

**Marcos Grilo**<sup>1</sup>; Inácio de Sousa Fadigas<sup>2</sup>;

<sup>1</sup> Doutor em Difusão do Conhecimento, Universidade Estadual de Feira de Santana e grilo@uefs.br;

<sup>2</sup> Doutor em Difusão do Conhecimento, Universidade Estadual de Feira de Santana e fadigas@uefs.br;

### RESUMO

Redes de coautoria, redes de atores de filmes e redes de palavras de títulos são exemplos de redes de cliques, onde os elementos básicos são topologicamente representados por grupos de vértices mutuamente conectados. Na presente pesquisa exploratória, contribuimos com pesquisas anteriores, agregando um método de construção de uma rede de cliques aleatória equivalente a uma rede real. Como aplicação, construímos uma rede semântica dos títulos das dissertações de mestrado do Programa Profmat para comparar índices clássicos e de cliques da rede semântica de cliques com aqueles da rede de cliques aleatória equivalente. A comparação permite quantificar o quanto a rede semântica de títulos se afasta da aleatoriedade de escolha das palavras. No campo das redes complexas, o método elaborado para a construção da rede de cliques aleatória permite inferências sobre o efeito mundo pequeno bem como sobre a não conformidade da lei de potência para redes de cliques. Os resultados encontrados evidenciam a forte relação entre a escolha das palavras dos títulos e a intencionalidade dos autores, de acordo com a temática do programa, influenciando na estrutura topológica das redes semânticas de cliques.

**PALAVRAS-CHAVE:** Redes de cliques; redes semânticas; rede de cliques aleatória.

<http://doi.org/10.55664/simbraredes2024.011>

### 1 INTRODUÇÃO

Redes de cliques foram introduzidas por <sup>1</sup> para designar redes formadas por títulos de artigos, cuja estrutura básica é representada por grafos completos, ou seja, grafos simples com vértices adjacentes a todos os outros (cliques). Índices para capturar propriedades inerentes às redes de cliques são propostos em <sup>2</sup> e uma extensão das propriedades para redes de cliques de tamanhos constantes encontra-se em <sup>3</sup>. A caracterização de redes usando a abordagem por cliques é encontrada em <sup>4</sup> (redes de coautoria), <sup>5</sup> (comparação de redes de títulos nos idiomas Português e Inglês), <sup>6</sup> (densidade em redes semânticas de títulos de artigos), <sup>7</sup> (robustez em redes semânticas de cliques), <sup>8</sup> (rede de arbitragem da copa FIFA de futebol masculino), <sup>9</sup> (redes de títulos de dissertações de mestrado), <sup>10</sup> (temáticas nos títulos de trabalho de conclusão de Licenciatura em Matemática), <sup>11</sup> (classificação da disciplinaridade em revistas científicas).

Para dar continuidade às pesquisas com redes de cliques, propomos um método de construção de rede de cliques aleatória equivalente, a ser usada como parâmetro para a interpretação de propriedades de redes reais. Como aplicação do método, construímos uma rede semântica de cliques formada por títulos de dissertações de mestrado profissional em Matemática (Profmat). Neste trabalho, seguimos a definição para rede semântica conforme <sup>7</sup>: é um sistema de representação de conhecimento por meio de grafos, imbuído de intenção de funcionalidade estabelecida pelo contexto, que é composto por palavras, conceitos ou entidades com significado semântico e seus relacionamentos. Exploramos basicamente propriedades de redes complexas, como distribuição de graus e efeito mundo pequeno, bem como propriedades clássicas de redes e propriedades inerentes a redes de cliques, encontradas em <sup>2</sup>.

Particularmente em relação a redes semânticas de títulos, a comparação da rede real com a rede de cliques aleatórias equivalente permite inferir, por exemplo, o quanto a escolha dos títulos está aderente ao escopo da revista, no caso de artigos, ou à temática do programa, no caso das dissertações. Pesquisas adicionais poderão estender a comparação de redes com as redes aleatórias equivalentes para outras redes de cliques, a exemplo das redes de arbitragem, redes de bancas, redes de coautoria, dentre outras.

### 2 METODOLOGIA

#### 2.1 Construção da rede de títulos de dissertações do Profmat

O tipo de grafo considerado nesta pesquisa para modelar a construção da rede semântica Profmat é finito, não-vazio, rotulado, de 1-modo, sem arestas múltiplas, sem laços, não dirigido e ponderado. Utilizamos os títulos das dissertações de mestrado como fonte de dados, coletados durante o mês de janeiro



SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

de 2023, a partir do catálogo disponibilizado no site do programa <[www.profmatsbm.org.br/dissertacoes](http://www.profmatsbm.org.br/dissertacoes)>. A coleta abrangeu o período de 2013 a 2022, compreendendo os dez primeiros anos de defesas de dissertações ocorridas no Profmat. Para cada dissertação, encontravam-se disponíveis no catálogo as seguintes informações: data de defesa; nome do aluno; título da dissertação; instituição; arquivo digital da dissertação. As informações foram organizadas em planilhas eletrônicas.

Aplicamos um tratamento manual baseado nos trabalhos de <sup>12</sup> e de <sup>13</sup>, que consistiu nas seguintes regras: 1) corrigir erros de digitação; 2) eliminar sinais de pontuação; 3) escrever siglas por extenso; 4) traduzir palavras em idioma estrangeiro para o português; 5) inspecionar o primeiro caractere de cada título de forma a garantir que cada título comece com letra maiúscula; 6) substituir símbolos e termos matemáticos por palavras reconhecidas pela ferramenta computacional de geração de redes semânticas de cliques. Armazenamos os títulos tratados em um arquivo de texto e aplicamos a ferramenta computacional desenvolvida por <sup>12</sup>, denominada de NetPal, para construir uma rede semântica de cliques na qual os vértices representam as palavras com significado intrínseco presentes nos títulos das dissertações do Profmat. Usamos um recurso do NetPal que elimina as palavras sem significado intrínseco (e.g. artigos, pronomes, preposições) e converte as demais para uma forma canônica (e.g. substantivos são convertidos para o masculino singular e verbos para o infinitivo). Uma aresta é estabelecida entre duas palavras pertencentes a um mesmo título. Neste sentido, as palavras presentes em um mesmo título conectam-se mutuamente. Dessa forma, cada título é representado por uma clique. Cada aresta entre duas palavras possui um peso que indica a quantidade de títulos nas quais ambas estão presentes simultaneamente. O resultado é uma rede semântica de cliques ponderada, baseada nos títulos de dissertações do Profmat.

A ferramenta computacional elaborada por <sup>12</sup> requisita que cada título seja alocado em uma linha com a primeira letra maiúscula. Devido a inconsistências detectadas em alguns dados, foi necessário consultar o arquivo digital de algumas dissertações. Uma inspeção manual foi realizada para garantir que cada título tivesse a primeira letra em maiúscula. Este procedimento foi necessário para que a quantidade de títulos fosse igual à quantidade de cliques. Outra questão associada à aplicação do NetPal trata da eliminação das palavras sem significado intrínseco. Especificamente, o NetPal elimina símbolos e termos matemáticos. Por isso, estabelecemos a regra de tratamento manual 6, com a finalidade de melhorar a qualidade dos dados. Elaboramos um vocabulário de controle com 854 descritores atribuindo para cada símbolo/termo matemático uma palavra que fosse classificada como *NOTFOUND* pelo NetPal. As palavras que não constam no dicionário utilizado pelo NetPal classificadas como *NOTFOUND*, não são eliminadas. Como resultado, a configuração inicial das cliques isoladas possui 6.717 cliques ( $n_q$ ) e 47.116 vértices ( $n_0$ ), e a rede resultante dos processos de justaposição e sobreposição tem 3.605 vértices ( $n$ ), denominada de rede Profmat.

## 2.2 Construção da rede aleatória equivalente

Considere que uma rede de cliques apresenta na sua configuração inicial  $n_q$  das cliques isoladas, nomeadas por  $Cl_1, \dots, Cl_{n_q}$  e com respectivos tamanhos  $q_1, \dots, q_{n_q}$ . Inicialmente, escolhe-se aleatoriamente e uniformemente,  $q_1$  vértices dentre os  $n$  vértices da rede. Conecta-se mutuamente estes  $q_1$  vértices escolhidos de modo a se obter uma clique aleatória, que denominaremos de  $ClA_1$ . Em seguida, escolhe-se aleatoriamente e uniformemente  $q_2$  vértices, dentre os  $n$  vértices. Conecta-se mutuamente estes  $q_2$  vértices de modo a se obter uma clique aleatória  $ClA_2$ . Repete-se o procedimento até obter-se uma configuração inicial de cliques aleatórias desconectadas  $ClA_1, \dots, ClA_{n_q}$  com respectivos tamanhos de cliques  $q_1, \dots, q_{n_q}$ . Tomando como estruturas básicas as  $ClA_1, \dots, ClA_{n_q}$ , podemos construir uma rede de cliques aleatória equivalente à rede de cliques original, utilizando os métodos de justaposição e de sobreposição, definidos por <sup>2</sup>. A justaposição ocorre quando duas cliques possuem apenas um vértice em comum. A sobreposição ocorre quando duas cliques possuem mais de um vértice em comum, ou seja, pelo menos uma aresta em comum. Para realizar os sorteios aleatórios na construção da rede aleatória equivalente, escrevemos um *script* usando o programa R, que é um ambiente de programação para computação estatística e gráficos, obtendo como saída um arquivo na extensão .net, compatível com o programa de redes Pajek.<sup>14</sup> Desta forma, obtemos uma rede de cliques aleatória equivalente, uma vez que tanto os  $n$  vértices quanto a sequência de tamanhos das cliques são extraídos da rede original. Vale acrescentar que a construção proposta difere daquelas usadas para os modelos clássicos de Erdős-Rényi <sup>15,16</sup> que usam como parâmetros o número de vértices e o grau médio da rede original. Também difere do modelo de <sup>17</sup> que inicia com  $n$  vértices isolados e arestas são estabelecidas aleatoriamente e uniformemente com probabilidade  $p$  para formar uma  $m$ -clique.

Usando o número de vértices e a distribuição dos tamanhos das cliques, fizemos 100 simulações de redes de cliques aleatórias conectadas equivalentes à rede Profmat. Os parâmetros da rede aleatória obtida, que denominamos de rede ProfmatCIA, correspondem à média das simulações.



### 3 RESULTADOS E DISCUSSÃO

A Tabela 1 (Anexo) apresenta a descrição dos índices usados na comparação entre as duas redes, bem como suas respectivas equações. Na Tabela 2 (Anexo), são apresentados os índices de suas equações correspondentes e na Tabela 3 os valores dos parâmetros e dos índices obtidos para as redes Profmat e ProfmatCIA.

**Tabela 3.** Índices das redes Profmat e ProfmatCIA

Índice	Rede Profmat	Rede ProfmatCIA
$n$	3.605	3.605
$n_0$	47.116	47.116
$n_q$	6.717	6.717
$\epsilon$	70.228	173.188
$\epsilon_0$	175.540	175.540
$\Delta$	0,01081	0,02666
$\Delta_{pond}$	0,02686	0,02702
$\Delta_{q_0}$	0,00015	0,00015
$\Delta_{norm}$	0,01065	0,02651
$V(\Delta)$	67,35523	167,57022
$R_v$	0,92402	0,92402
$R_\epsilon$	0,98143	0,01342
$\langle k \rangle$	38,96144	96,08246
$\langle k_p \rangle$	96,7939	97,387
$C$	0,16057	0,10419
$C_{ws}$	0,78319	0,11653
$\ell$	2,36893	2,09039
$D$	5	3,06
$D^*_{ref}$	0,11285	0,05238

A análise dos resultados encontrados da comparação dos índices das redes Profmat e ProfmatCIA mostra que o método de construção da rede aleatória equivalente (ProfmatCIA) está coerente, uma vez que os valores de  $n$ ,  $n_0$  e  $n_q$  são os mesmos, como esperado. O mesmo ocorre com o número de arestas da configuração inicial das cliques isoladas ( $\epsilon_0$ ). Consequentemente a densidade de cliques  $\Delta_{q_0}$  mantêm-se com o mesmo valor, pois só depende de  $n_0$  e  $\epsilon_0$ , bem como o fator de redução de vértices, que basicamente depende de  $n$  e  $n_0$ , pois o valor de  $n_{max}$  é muito menor que os valores de  $n$  e  $n_0$ . Por outro lado, o número de arestas  $\epsilon$  da rede ProfmatCIA é cerca de 98% do número de arestas da configuração inicial, o que é confirmado pelo valor do fator de superposição de arestas  $R_\epsilon$  (0,01342). Significa que há poucos pares de arestas repetidos na rede ProfmatCIA, o que condiz com a aleatoriedade da rede.

A densidade da rede ProfmatCIA é cerca de 2,5 vezes maior que a densidade da rede Profmat, o que é coerente com a mesma proporção entre o número de arestas  $\epsilon$  nas duas redes, uma vez que  $n$  é o mesmo em ambas as redes. Quando se leva em conta o peso nas arestas, os valores das densidades ponderadas são próximos devido ao fato de que o número de arestas não ponderadas da rede ProfmatCIA ( $\epsilon = 173.188$ ) é próximo ao número de arestas ponderadas da rede Profmat ( $\epsilon_0 = 175.540$ ). A mesma justificativa aplica-se em relação ao grau médio ponderado  $\langle k_p \rangle$ .

A relação entre os graus médios  $\langle k \rangle$  das redes Profmat e ProfmatCIA pode ser explicada devido à aleatoriedade na escolha das cliques da rede ProfmatCIA. Uma vez que há pouca repetibilidade nos pares de vértices formados aleatoriamente, há uma tendência de que os vértices na rede aleatória tenham mais conexões. Esse fato é corroborado pelos valores próximos para o grau médio ponderado, uma vez que neste caso os pesos das arestas na rede Profmat são considerados.

À guisa de uma interpretação para a rede de títulos de dissertações (rede Profmat), tomando os vértices como palavras dos títulos, conforme construção descrita na Seção 2.1, é possível inferir algumas

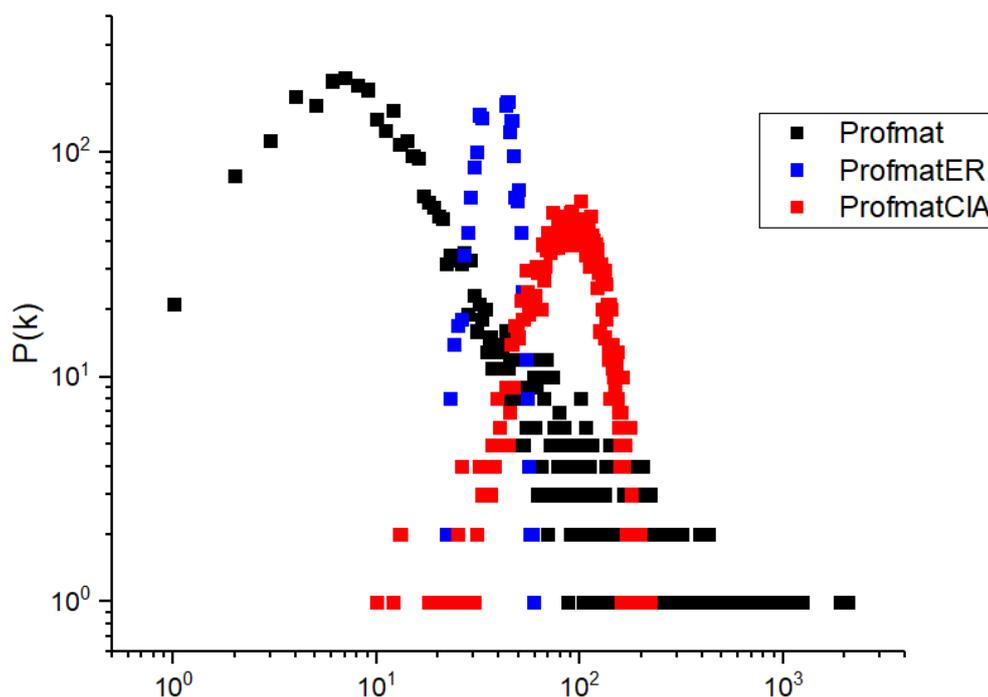
interpretações. No processo construtivo da rede ProfmatCIA, um grupo de  $q_i$  palavras tem a mesma chance de ser escolhido (aleatoriamente e uniformemente) que qualquer outro de mesmo tamanho. Por outro lado, na rede Profmat, a escolha das palavras é influenciada pelos autores, em conformidade com a temática do Programa, ou seja, a escolha das palavras pelos autores não é aleatória. Conseqüentemente, isso impacta em índices como o número de arestas  $\epsilon$ , a densidade e o grau médio, relacionados por  $\langle k \rangle = (n - 1)\Delta$ , e nos índices que estão relacionados diretamente com estes ( $R_\epsilon$ ,  $V(\Delta)$  e  $\Delta_{norm}$ ). Os valores dos índices impactados pela escolha dos autores reflete o grau de previsibilidade nas escolhas das palavras.

Na literatura de redes complexas, dizemos que uma rede apresenta o efeito mundo pequeno quando a rede real é comparada com uma rede aleatória clássica, ou seja, uma rede aleatória com o mesmo número de vértices e o mesmo grau médio.<sup>18</sup> Segundo<sup>18</sup>, uma rede é dita mundo pequeno quando valores dos caminhos mínimos  $\ell$  para ambas são próximos, mas os valores do coeficiente de aglomeração  $C_{ws}$  da rede real é muito maior que aquele da rede aleatória clássica. Para verificar a existência ou não do efeito mundo pequeno para redes de cliques, comparamos as grandezas para as redes Profmat e ProfmatCIA. Na rede Profmat, conforme mostra a Tabela 3, o valor do coeficiente de aglomeração médio  $C_{ws}$  da rede Profmat ( $C_{ws} = 0,78319$ ) é muito maior que aquele da rede ProfmatCIA ( $C_{ws} = 0,11653$ ). Entretanto, os valores dos caminhos mínimos médios ( $\ell$ ) diferem em cerca de 13%, mas ainda são compatíveis com a definição de.<sup>18</sup> O outro critério que pode ser usado para redes de cliques é devido a<sup>2</sup>, que consideram que uma apresenta uma topologia mundo pequeno quando o coeficiente de aglomeração da rede (real) de clique  $C_{clique} = C_{ws} - C \geq 0,5$  e  $0 \leq D^*_{ref} \leq 0,25$ . Para a rede Profmat temos  $C_{clique} = 0,78319 - 0,16057 = 0,62262 > 0,5$ , enquanto o diâmetro de referência normalizado  $D^*_{ref} = 0,11285 < 0,25$ . Portanto, a construção da rede de clique aleatória corrobora também com os critérios definidos por Fadigas e Pereira (2013).

A distribuição dos graus em uma rede é um indicador importante em redes complexas, pois pode caracterizar se esta distribuição segue uma lei de potência, ou seja, se a rede é livre de escala. Os estudos seminais sobre o tema são devidos a<sup>19</sup> e posteriormente<sup>20</sup> estabeleceram critérios mais rigorosos para definir se uma distribuição segue uma lei de potência.

A Figura 1 mostra os gráficos das distribuições de graus das redes Profmat, ProfmatCIA e ProfmatER. Com o propósito de ampliar a discussão quanto a esta propriedade, construímos também uma rede aleatória de Erdős-Rényi equivalente com os parâmetros  $n$  e  $\langle k \rangle$  da rede Profmat, denominada ProfmatER. Dessa forma, torna-se viável comparar o comportamento da distribuição de graus entre redes com diferentes processos de construção.

**Figura 1.** Gráfico da distribuição de graus das redes Profmat, ProfmatCIA e ProfmatER





SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

A distribuição de graus da rede Profmat segue o padrão de distribuição em diversas redes de cliques presentes na literatura de redes, a exemplo de .<sup>13,1,5</sup> A comparação da distribuição de graus das redes ProfmatCIA e ProfmatER mostra que são gráficos compatíveis com a aleatoriedade das redes, porém os gráficos da ProfmatER e ProfmatCIA são defasados em relação aos valores do grau  $k$  devido ao fato de que rede ProfmatCIA é formada por cliques, o que resulta em vértices com grau  $q_i - 1$ , onde  $q_i$  é o tamanho da clique de índice  $i$ . No modelo da rede ProfmatER, cada vértice é inserido na rede isoladamente.

O método de <sup>20</sup>, que avalia a conformidade da distribuição de graus com uma lei de potência, estabelece um certo parâmetro  $p$  tal que a lei de potência é descartada se  $p \leq 0,1$ . Para a rede Profmat encontramos  $p = 0$ . Assim, a hipótese de ser livre de escala segundo <sup>20</sup> não é plausível para a rede Profmat. A propósito de uma possível interpretação do comportamento da distribuição de graus para a rede semântica de títulos de dissertações (rede Profmat), os gráficos da Figura 1 mostram que a intencionalidade dos autores ao escolherem as palavras para compor os títulos de suas dissertações, resulta em um comportamento da distribuição de graus que se aproxima de uma lei de potência, segundo <sup>19</sup>, não confirmada pelo critério de .<sup>20</sup>

#### 4 CONSIDERAÇÕES FINAIS

Nesta pesquisa propomos um método que cria redes de cliques aleatórias equivalente a uma rede real, usando como parâmetros a quantidade de vértices e a distribuição dos tamanhos das cliques na sua configuração inicial das cliques isoladas. Como forma de aplicação, construímos uma rede de cliques formada pelas palavras dos títulos das dissertações de mestrado do Programa Profmat, bem como uma rede de cliques aleatória equivalente. Calculamos alguns índices de redes, tanto da abordagem clássica quanto da abordagem por cliques, para as duas redes: rede Profmat e rede ProfmatCIA. Como resultado importante, concluímos que a construção da rede de cliques aleatória equivalente ProfmatCIA é mais apropriada do que a rede ProfmatER para avaliar o quanto uma rede de cliques se afasta do padrão de aleatoriedade. Diferentemente do modelo Erdős-Rényi, no qual sorteiam-se grafos  $(G_{n,m})$  ou arestas  $(G_{n,p})$ , no método proposto neste trabalho, sorteamos cliques a partir de uma sequência de tamanhos de cliques. Outra contribuição consiste em definir um critério para avaliar a presença do efeito mundo pequeno em redes de cliques, compatível com àqueles consolidados na literatura de redes. Especificamente em relação à rede de títulos analisada, constatamos que a escolha das palavras dos títulos está fortemente relacionada à intencionalidade dos autores, em conformidade com a temática do programa.

Os resultados obtidos apontam para pesquisas futuras, que consistem em comparar outros tipos de redes (redes de arbitragem, redes de bancas, redes de coautoria, dentre outras) com as respectivas redes de cliques aleatórias equivalentes.

#### Agradecimentos

Os autores agradecem o apoio financeiro da Universidade Estadual de Feira de Santana por meio do programa FINAPESQ (Termo de Outorga nº 22/2022).

#### 5 REFERÊNCIAS

- 1 PEREIRA, H. B. B. et al. Semantic networks based on titles of scientific papers. **Physica A: Statistical Mechanics and its Applications**, v. 390, n. 6, p. 1192–1197, 2011.
- 2 FADIGAS, I. S.; PEREIRA, H. B. B. A network approach based on cliques. **Physica A: Statistical Mechanics and its Applications**, v. 392, n. 10, p. 2576–2587, 2013.
- 3 FADIGAS, I. S. et al. FIFA World Cup referees' networks: a constant-size clique approach. **Social Network Analysis and Mining**, v. 10, n. 1, p. 54, 2020.
- 4 ROSA, M. G. et al. Clique approach for networks: Applications for coauthorship networks. **Social Networking**, v. 3, n. 2, 2014.
- 5 HENRIQUE, T. et al. Mathematics education semantic networks. **Social Network Analysis and Mining**, v. 4, n. 1, p. 200, 2014.
- 6 PEREIRA, H.B.B. et al. Density: A measure of the diversity of concepts addressed in semantic networks. **Physica A: Statistical Mechanics and its Applications**, v. 441, 2016.
- 7 GRILO, M. et al. Robustness in semantic networks based on cliques. **Physica A: Statistical Mechanics and its Applications**, v. 472, p. 94 – 102, 2017.
- 8 FADIGAS, I.S. et al. FIFA World Cup referees' networks: a constant-size clique approach. **Social Network Analysis and Mining**, v. 10, n. 1, p. 54, 2020a.
- 9 FADIGAS, I. DE S. et al. Master's degree programs in Mathematics in Brazil: an application of networks to characterize their titles. **Obra digital: revista de comunicación**, v. 0, n. 18, 2020.



SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

- <sup>10</sup> SANTOS, V. C.; GRILO, M. Identificação de temáticas de trabalhos de conclusão de curso por meio de redes semânticas. **Revista Paranaense de Educação Matemática**, v. 9, n. 20, p. 219–235, 2020.
- <sup>11</sup> FADIGAS, I.S.; GRILO, M.; DE BARROS PEREIRA, H. B. Scientific journal disciplinary quantification and sorting using a network index. **Quality & Quantity**, 2022.
- <sup>12</sup> CALDEIRA, S. M. G. et al. The network of concepts in written texts. **The European Physical Journal B**, v. 49, p. 523–529, 2006.
- <sup>13</sup> FADIGAS, I. S. et al. Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos em educação matemática. **Educação Matemática Pesquisa**, v. 11, n. 1, p. 142–164, 2009.
- <sup>14</sup> BATAGELJ, V.; MRVAR, A. Pajek - Program for Large Network Analysis. Disponível em <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.
- <sup>15</sup> ERDÖS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, v. 6, p. 290–297, 1959.
- <sup>16</sup> ERDÖS, P.; RÉNYI, A. On the evolution of random graphs. **Bulletin Institute International Statistics**, 1961.
- <sup>17</sup> DING, Y.-M. et al. Statistical properties of random clique networks. **Frontiers of Physics**, v. 12, n. 5, p. p. 128909, 2017.
- <sup>18</sup> WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world' networks. **Nature**, v. 393, n. 6684, p. 440–442, 1998.
- <sup>19</sup> BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, n. 286, p. 509–512, 1999.
- <sup>20</sup> CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M. E. J. Power-law distributions in empirical data. **SIAM Review**, n. 04, p. 661–703, 2009.



SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

## Anexos

**Tabela 1.** Notações e descrições dos parâmetros de rede

Notação	Descrição
$n$	Número de vértices da rede (ou tamanho da rede)
$n_0$	Número de vértices na configuração inicial de cliques desconectadas
$n_q$	Número de cliques na configuração inicial de cliques desconectadas
$\epsilon$	Número de arestas da rede
$k_i$	Grau do vértice $v_i$
$k_p(i)$	Grau ponderado do vértice $v_i$
$d_{ij}$	Distância geodésica entre os vértices $v_i$ e $v_j$
$n_\Delta$	Número de triângulos da rede
$n_\vee$	Número de triplas conectadas da rede
$r_i$	Número de pares de vértices vizinhos a $v_i$ que são vizinhos entre si
$n_{max}$	Número de vértices da maior clique na configuração inicial de cliques desconectadas
$\epsilon_0$	Número de arestas na configuração inicial de cliques desconectadas
$D$	Diâmetro da rede

**Tabela 2.** Índices das redes usados na pesquisa

Índice	Definição
Densidade	$\Delta = \frac{2\epsilon}{n(n-1)}$
Densidade para rede ponderada	$\Delta_{pond} = \frac{2\epsilon_0}{n(n-1)}$
Grau médio	$\langle k \rangle = \frac{1}{n} \sum_{i=1}^n k_i$
Grau médio ponderado	$\langle k_p \rangle = \frac{1}{n} \sum_{i=1}^n k_p(i)$
Caminho mínimo médio	$\ell = \frac{2}{n(n-1)} \sum_{i \neq j}^n d_{ij}$
Transitividade	$C = \frac{3n_\Delta}{n_\vee}$
Coefficiente de aglomeração médio	$C_{ws} = \frac{1}{n} \sum_{i=1}^n \frac{2r_i}{k_i(k_i-1)}$



SimBraRedes

III Simpósio Brasileiro de Teoria  
e Ciência de Redes

Fator de redução de vértices	$R_v = \frac{n_0 - n}{n_0 - n_{max}}$
Fator de sobreposição de arestas	$R_\epsilon = \frac{\epsilon_0 - \epsilon}{\epsilon_0 - \epsilon_{max}}$
Densidade de cliques	$\Delta_{q_0} = \frac{2\epsilon_0}{n_0(n_0 - 1)}$
Densidade de cliques normalizada	$\Delta_{norm} = \frac{\Delta - \Delta_{q_0}}{1 - \Delta_{q_0}}$
Variação de densidade	$V(\Delta) = \frac{\Delta - \Delta_{q_0}}{\Delta_{q_0}}$
Diâmetro de referência normalizado	$D^*_{ref} = \frac{\ln \frac{D}{2}}{\ln \frac{n_q}{2}}$