

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

PAULO DINHNAI OGATA FILHO

Análise da Diversidade de Comportamentos em Redes Complexas: Impacto das
Características da Rede

Trabalho de conclusão de curso apresentado
ao Instituto de Física de São Carlos da
Universidade de São Paulo para obtenção do
título de Bacharel em Física Computacional.
Orientador: Prof. Gonzalo Travieso

São Carlos
2023

1 INTRODUÇÃO

Como seres sociais, vivenciamos uma constante imersão em uma ampla gama de ideias que exercem uma influência profunda sobre nossas ações e comportamentos. Essas influências englobam um vasto espectro de tópicos, desde questões políticas até manifestações artísticas, moldando nossas interações diárias. À medida que o tempo avança, essas perspectivas evoluem, contribuindo para a criação de uma rica diversidade sociocultural. Onde emergem inúmeros estilos de vida distintos e uma multiplicidade de comportamentos que não apenas incentivam a inovação, mas também desempenham um papel fundamental no progresso coletivo e na evolução da humanidade. Nesse contexto, este trabalho busca compreender a dinâmica das opiniões e comportamentos em uma sociedade e como podemos enriquecer a diversidade.

1.1 Objetivo

Neste cenário, queremos representar estruturas sociais em redes complexas e analisar o fluxo de ideias e comportamentos, verificando como diferentes características dessas redes impactam na preservação da diversidade. Para a simulação da dinâmica de comportamentos, iremos nos basear no modelo e estudo originalmente desenvolvido por (MUSSO *et al*, 2022)^[1], onde investigaram a influência de algumas características sobre as redes, como o efeito da homogeneidade de graus, a distância mínima média entre os vértices e a presença de triângulos, identificando que esses fatores exercem uma influência positiva na diversidade.

Iremos então nos aprofundar neste modelo, verificando como a diversidade se mantém em diferentes cenários, explorando características que não foram abordadas no estudo original. O foco será controlar variáveis-chave, como o grau médio de conexões entre os vértices, o tamanho da rede, a presença de agrupamentos e as estruturas de redes assortativas e disassortativas.

2 MODELO DE MUSSO E HELBING E MEDIDA DE DIVERSIDADE

2.1 Modelo de Musso e Helbing

O modelo de fluxo de ideias e comportamentos que será utilizado neste estudo funciona da seguinte maneira: consideramos uma rede de amigos representada por um grafo, em que cada vértice representa um indivíduo e cada aresta representa a relação entre esses indivíduos. Inicialmente, cada indivíduo é caracterizado por uma atribuição que identificamos como um comportamento. Em cada intervalo de tempo na rede, todos os indivíduos simultaneamente

devem adotar o comportamento de um de seus vizinhos ou criar um comportamento completamente novo, que não tenha sido previamente apresentado por ninguém. A figura 1 retrata a dinâmica.

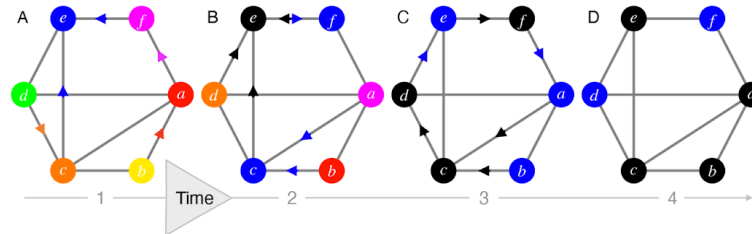


Figura 1: Representação da dinâmica do modelo proposto por Musso e Helbing. Nela vemos a dinâmica dos comportamentos (representados por cores) fluindo pelos vértices através dos intervalos de tempo, onde os vértices copiam o comportamento de um vizinho aleatório ou inventam um comportamento novo na rede. Fonte: (MUSSO *et al*, 2022)^[1], modificado.

Podemos observar da figura 1 a representação de seis indivíduos, cada um deles identificado por uma letra e possuindo uma cor que reflete seu comportamento. Esses indivíduos estão interligados, formando uma rede simples. À medida que o tempo avança, a cada iteração, percebemos que cada vértice escolhe aleatoriamente um de seus vizinhos e adota o comportamento desse vizinho. Para demonstrar esse processo, podemos tomar o exemplo do vértice *a*: no primeiro passo, copia o comportamento de *f*, e nos dois passos seguintes, adota o comportamento de *c*.

Além disso, existe a possibilidade de um vértice criar um comportamento novo que ainda não foi adotado por nenhum outro indivíduo na rede. Isso fica evidente no caso do vértice *e*, que no primeiro passo de tempo não copia o comportamento de ninguém, mas, em vez disso, inventa um novo comportamento que se espalha pela rede.

2.2 Índice de Simpson

Após a aplicação da dinâmica comportamental nas redes, ao final das iterações, realizamos o cálculo da diversidade de comportamentos na sociedade utilizando o Índice de Simpson (SIMPSON, 1949)^[2]. Este índice é amplamente empregado em estudos relacionados à ecologia e à biodiversidade que leva em consideração tanto o número total de indivíduos presentes na amostra quanto a quantidade de cada espécie. No contexto deste trabalho, o número total de indivíduos será o número de vértices presentes na rede e as espécies serão os comportamentos.

A aplicação desta medida irá nos fornecer informações sobre a heterogeneidade de uma comunidade. Para isso, inicialmente consideramos uma comunidade contendo s comportamentos e N indivíduos totais (vértices), o número de indivíduos que possuem um determinado comportamento é dado por n_j , onde j representa o j -ésimo comportamento ($j=1,2,3,...,s$), desta forma temos que a soma de todos os indivíduos de todos os comportamentos é igual a N .

$$\sum_{j=1}^s n_j = N \quad (1)$$

O cálculo do índice é realizado mediante a probabilidade de dois indivíduos, selecionados aleatoriamente e sem reposição, possuírem o mesmo comportamento. A medida é dada pela seguinte equação:

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^s n_j (n_j - 1) \quad (2)$$

onde $\frac{1}{2} \cdot n_j (n_j - 1)$ representa o número de pares retirados do mesmo grupo (vértices que apresentam o mesmo comportamento) e $\frac{1}{2} \cdot N(N - 1)$ é o número de pares que podemos obter na rede. É importante notar que os valores $\frac{1}{2}$ se cancelam na equação final. Quando isoladamente verificamos o valor que é subtraído de 1, percebemos ser a representação da probabilidade de encontrar dois indivíduos com o mesmo comportamento. Subtraindo esse valor encontrado de 1, obtemos a chance de encontrarmos indivíduos com comportamentos diferentes. Assim, quanto maior o valor obtido no Índice de Simpson, maior será a diversidade na comunidade.

Em um cenário de alta diversidade, encontramos um ambiente em que a riqueza de comportamentos, ou seja, o número de comportamentos, é elevada, e a distribuição de comportamentos entre os indivíduos é relativamente uniforme, evitando a predominância de um único comportamento. Nesse contexto, a diversidade da comunidade se manifesta de forma mais marcante. Para ilustrar o conceito, vamos aplicar a medida na Figura 1, iteração 1, onde temos $N=6$ e seis comportamentos diferentes, ou seja, $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 1$. Ao extrair o índice de Simpson nesse cenário, concluímos que a diversidade é máxima, $D=1$. Agora, ao realizarmos o mesmo na etapa 4, onde temos dois comportamentos diferentes, *azul* e *preto*, com quantidades respectivas de $n_{\text{azul}} = 2$ e $n_{\text{preto}} = 4$, obtemos um valor de diversidade $D=0.533$.

3 MEDIDAS E MODELOS DE REDES COMPLEXAS

3.1 Medidas

Para conduzir a análise da diversidade em diversos cenários, iremos introduzir diferentes redes, as quais serão caracterizadas por meio das medidas descritas a seguir.

3.1.1 Grau

O grau de um vértice se refere simplesmente à contagem de arestas que incidem sobre ele, assim como vemos na figura 2.

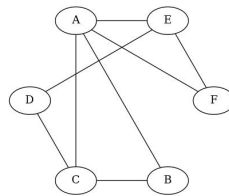


Figura 2: Representação de uma rede g , onde podemos contar a quantidade de arestas que estão presentes em cada um dos seis vértices. Fonte: elaborada pelo autor.

Podemos notar a presença de quatro arestas presentes no vértice A, sendo seu grau então igual a quatro. Para a contagem do grau de cada vértice, devemos primeiro representar a rede como uma matriz de adjacência que consiste em retratar os vértices em linhas e colunas da matriz. Quando existe uma aresta conectando dois vértices, os elementos correspondentes na matriz de adjacência são preenchidos com o valor 1; caso contrário, são preenchidos com 0. Se a rede for não direcionada, a matriz torna-se simétrica (como é o caso deste estudo). Como exemplo, representaremos a rede da figura 2 em uma matriz de adjacência na figura 3.

$$\begin{bmatrix} & a & b & c & d & e & f \\ a & 0 & 1 & 1 & 0 & 1 & 1 \\ b & 1 & 0 & 1 & 0 & 0 & 0 \\ c & 1 & 1 & 0 & 1 & 0 & 0 \\ d & 0 & 0 & 1 & 0 & 1 & 0 \\ e & 1 & 0 & 0 & 1 & 0 & 1 \\ f & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Figura 3: Representação da rede g na forma de matriz de adjacência. A primeira linha e coluna indicam os vértices, já os elementos indicam se existe conexão. Fonte: elaborada pelo autor.

Para calcularmos então o grau de cada vértice, percorremos a matriz contando as conexões

$$k_i = \sum_{j=1}^N A_{ij} \quad (3)$$

onde k_i representa o grau do vértice i , A_{ij} são os elementos da matriz de adjacência e N é o número total de vértices existentes. Já o grau médio de uma rede g ($\langle k_g \rangle$), nada mais é do que a relação da somatória dos graus médios de cada vértice pelo número de vértices.

$$\langle k_g \rangle = \frac{1}{N} \sum_{i=1}^N k_i \quad (4)$$

Desta forma, aplicando a equação 3 e 4 na matriz de adjacência da figura 3, temos os graus dos vértices iguais a $k_a = 4$, $k_b = 2$, $k_c = 3$, $k_d = 2$, $k_e = 3$ e $k_f = 2$. E para a rede toda encontramos $\langle k_g \rangle = 2.66$.

3.1.2 Distribuição de grau

A distribuição de graus em uma rede representa a probabilidade de um vértice ter um grau específico. Uma abordagem comum para caracterizar essa distribuição é por meio da função $P(k)$, que descreve essa relação e é representada pela equação 5, onde n_k representa o número de vértices com grau k , e N é o número total de vértices na rede. Quando utilizamos a notação $P(k = d)$, estamos, quantificando a probabilidade de encontrar um vértice com d conexões.

$$P(k) = \frac{n_k}{N} \quad (5)$$

Em redes aleatórias, a distribuição de graus segue uma distribuição de Poisson. Já em outras muitas redes do mundo real, a distribuição de graus segue a chamada Lei de Potência, no qual $P(k) \sim k^{-\gamma}$ para algum valor constante γ .

3.1.3 Distância

Um caminho dentro de uma rede representa uma sequência de vértices conectados que nos permite navegar de um vértice de origem i para um vértice de destino j . Em diversos cenários, iremos ter mais de um caminho para percorrer de um vértice a outro, como exemplificado na figura 4. Torna-se interessante então a utilização da distância mínima entre esses dois vértices, que por sua vez, é definida como o caminho que requer o menor número

de arestas para ser percorrido de i a j . Iremos denominar esta distância de menor extensão como $d(i, j)$.

A média dos menores caminhos da rede é representado por l , que é a média das distâncias mínimas entre todos os vértices da rede. O meio para o cálculo do menor caminho médio l é dada pela seguinte equação

$$l = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N d(i, j) \quad (6)$$

onde N representa o número de vértices na rede.

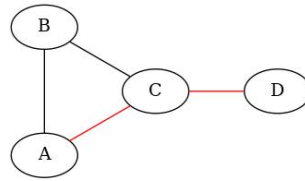


Figura 4: Exemplo de menor caminho entre dois vértices, onde vemos a existência de dois caminhos para ir do vértice A até D; sendo o mais longo A,B,C,D e o mais curto A,C,D. Fonte: elaborada pelo autor.

3.1.4 Coeficiente de Aglomeração

A análise dos agrupamentos em redes é quantificada por meio do Coeficiente de Aglomeração (C), também conhecido como o fenômeno de transitividade. Esse fenômeno se manifesta quando um nó i está conectado a um nó j e o nó j está conectado a um nó k , aumentando as probabilidades de que o nó i também esteja conectado ao nó k . Em termos simples, a transitividade indica a presença de triângulos na rede, que são conjuntos de três nós interconectados.

O coeficiente de aglomeração em torno de um vértice i é calculado como o número de conexões entre os vizinhos do vértice i dividido pelo número total possível de conexões, como na equação 7.

$$C_i = \frac{2 \cdot e_i}{k_i(k_i - 1)} \quad (7)$$

onde e_i representa o número de ligações entre os vizinhos do vértice i e k_i o grau do vértice i .

Como exemplo temos a figura 5.

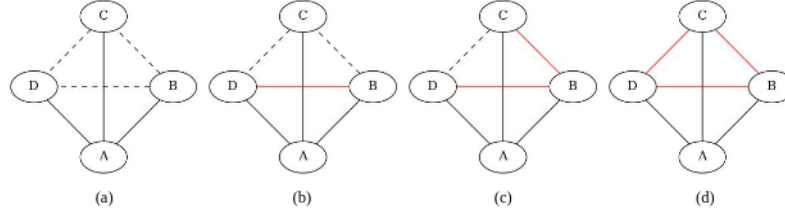


Figura 5: Cálculo do Coeficiente de Aglomeração do nó A. As linhas contínuas representam conexões existentes, enquanto as linhas tracejadas indicam possíveis conexões. Em (a), não há conexões entre os vizinhos de A, resultando em $e_A=0$ e, portanto, $C=0$. Em (b), temos $e_A=1$, resultando em $C=1/3$. Em (c), encontramos $e_A=2$, levando a $C=2/3$, e, por fim, em (d), com $e_A=3$, obtemos $C=1$. Fonte: elaborada pelo autor.

Já para o cálculo do coeficiente de aglomeração da rede podemos realizar a média dos coeficientes individuais, dado por:

$$\langle C_g \rangle = \frac{1}{N} \sum_{i=1}^N \frac{2 \cdot e_i}{k_i(k_i-1)} = \frac{1}{N} \sum_{i=1}^N C_i \quad (8)$$

aplicado na rede da figura 5 - (b) teríamos a medida $\langle C_g \rangle = 0.583$.

3.1.5 Assortatividade

A assortatividade em redes complexas é a propensão de vértices em comum a se conectarem. Tomamos como exemplo uma estrutura de rede formada por amigos, neste modelo a conexão entre os vértices é influenciada por fatores em comum, como o idioma, raça, idade e outros aspectos. Assim temos uma relação assortativa. Por outro lado, se os indivíduos tendem a se relacionar com pessoas que possuem características diferentes, chamamos de relação disassortativa (NEWMAN, 2003)^[3]. Esse conceito é fundamental para entender e explicar como algumas redes do mundo real se moldam, assim como ocorre em redes de coautoria em física, colaborações entre atores de cinema e ligações entre diretores de empresas que mostram padrões de mistura assortativa. Já por outro lado, vemos em (ARCAGNI *et al*, 2017)^[4], que redes disassortativas são exemplificadas por aquelas relacionadas à Internet, à World Wide Web (WWW), a interações de proteínas, entre outras. Neste trabalho a assortatividade de uma rede mede a tendência dos vértices de conectarem com outros de grau similares (figura 6). Se a rede tem assortatividade negativa, vértices de alto grau tendem a se conectar com vértices de baixo grau, e a rede é dita disassortativa. O coeficiente de assortatividade r é calculado conforme mostra a equação 9.

$$r = \frac{1}{\sigma_k \sigma_q} \sum_{k,q} kq(P(k,q) - P(k)P(q)) \quad (9)$$

onde $P(k, q)$ representa a relação entre o número de vezes que temos arestas saindo de um vértice de grau k e chegando num outro de grau q pelo número total de arestas presente na rede; $P(k)$, representa a relação do número de arestas cuja origem é um vértice de grau k pelo número total de arestas. O mesmo é válido para $P(q)$; k e q são os graus dos vértices de origem e destino; σ_k é o desvio padrão de $P(k)$, o mesmo é válido para σ_q . Os valores para r variam entre -1 a 1.

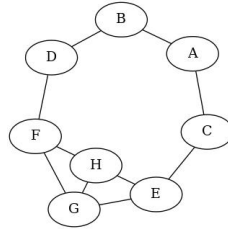


Figura 6: Representação de uma rede assortativa. Podemos identificar dois grupos de vértices; o primeiro grupo inclui os vértices A, B, C e D, cada um com grau 2, enquanto o segundo grupo engloba os vértices E, F, G e H, cada um com grau 3. A assortatividade desta rede é $r=0.583$. Fonte: elaborada pelo autor.

3.2 Modelos de redes complexas

3.2.1 Modelo de rede aleatória

O modelo foi proposto por Erdos e Rényi (ERDŐS; RÉNYI, 1959)^[5] (ERDŐS; RÉNYI, 1960)^[6], e constitui uma abordagem fundamental no estudo de redes complexas. Nesse modelo, partimos de um conjunto fixo de N vértices, cada um ligado a outros por meio de arestas não direcionadas, com uma probabilidade p de conexão constante e independente de quaisquer fatores. O grau esperado para qualquer vértice na rede é dado pela equação 10, que representa o número de tentativas de conexão de um vértice i , multiplicado pela probabilidade de sucesso dessa ligação.

$$k_i = p(N - 1) \quad (10)$$

a variável k_i representa o grau do vértice i , p é a probabilidade de conexão entre os vértices e N é o número total de vértices presentes. Para implementação desta rede contamos com um algoritmo que leva dois parâmetros, o primeiro é a quantidade de vértices N que a rede irá ter, e o segundo é a probabilidade de ligação entre os vértices p . A escolha desta rede se deve ao fato de possuir a distribuição de graus homogênea seguindo uma distribuição de Poisson. Isso ocorre devido a equação 10, onde é possível especificar um ponto onde os graus dos vértices

irão estar distribuídos, controlando assim o grau médio da rede. Para esse controle, atribuímos ao parâmetro p a relação entre o grau que queremos e a quantidade de vértices da rede.

3.2.2 Modelo Pequeno Mundo

O modelo de Watts e Strogatz (WATTS; STROGATZ, 1998)^[7] apresenta semelhanças com o modelo de Erdos-Renyi, uma vez que ambos incorporam uma probabilidade de conexão entre os vértices. No entanto, este modelo se destaca ao introduzir características que promovem conexões iniciais entre um valor v de vértices próximos, criando uma estrutura rica em triângulos. Devido a essa característica, onde cada vértice se conecta com v vizinhos, naturalmente a rede assume a forma de um anel como mostra a figura 7, para $p=0$. A probabilidade, denotada por p , na realidade, representa a probabilidade de reconexão, indicando a chance de um vértice que inicialmente estava conectado a um vizinho se reconectar a qualquer outro na rede. Isso significa que essas reconexões estabelecem atalhos reduzindo a distância mínima média (l) na rede. Desta forma, é possível notar que a variação dessa probabilidade afeta a complexidade da rede e suas características iniciais. Assim, quando $p=1$, a rede se torna totalmente aleatória, enquanto em $p=0$, a estrutura regular inicial é mantida.

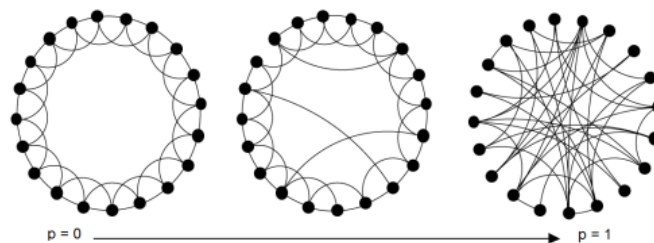


Figura 7: Rede de Watts Strogatz com conexão entre 4 vizinhos representando o aumento da complexidade alterando a probabilidade de religação p . Fonte: (WATTS; STROGATZ, 1998)^[7], modificado.

O ponto central de interesse desse modelo é explorar o espaço entre esses extremos, onde as propriedades de transitividade são mantidas, mas a distância mínima média é alterada. Recorremos então para a manipulação do parâmetro p em nossas análises, resultando na perda das características típicas iniciais da rede relacionadas à distância e ao agrupamento.

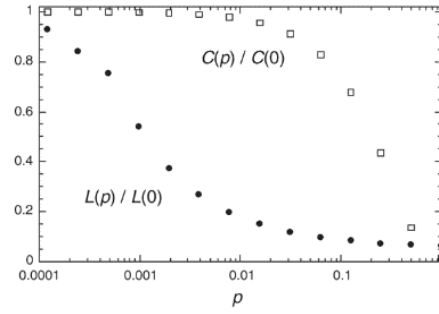


Figura 8: Relação entre a distância mínima média e o coeficiente de aglomeração da rede Watts Strogatz. Fonte: (WATTS; STROGATZ, 1998)^[7], modificado.

Podemos ver da figura 8, na faixa de valores em que p está entre 0.0001 a 0.01, a transitividade da rede sofre pouca alteração, enquanto a distância mínima média é significativamente afetada. Já para valores de p superiores a 0.01, a distância média mínima da rede diminui de forma menos acentuada, ao passo que o coeficiente de aglomeração diminui expressivamente.

Para a implementação da rede Watts-Strogatz, devemos especificar três parâmetros essenciais: N (o número total de vértices na rede), v (o número de conexões de cada vértice) e p (a probabilidade que dita as reconexões subsequentes na rede).

3.2.3 Redes livre de escala

Estudos conduzidos por (BARABÁSI *et al*, 1999)^[8] revelaram a existência de redes que exibem uma estruturação não aleatória, caracterizando-se por propriedades distintas. Nesse modelo, ocorre o fenômeno da conexão preferencial, no qual novos nós têm preferência por se conectar aos vértices mais conectados, resultando na formação de *hubs* que são vértices altamente ligados. É notável que, devido a formação de sua estrutura, este modelo de rede tende a possuir alguns hubs proeminentes. O que pode ser confirmado pela equação que descreve a distribuição de graus deste modelo, onde segue uma Lei de Potência expressa da seguinte forma:

$$P(k) \sim k^{-\gamma} \quad (11)$$

Nessa equação, $P(k)$ representa a probabilidade de um vértice possuir k conexões e γ um número real positivo que varia no geral entre 2 e 3 para redes reais. Como exemplo, temos as redes estudadas em (BARABÁSI *et al*, 1999)^[8] abordando a rede de colaboração de atores de cinema cujo valor de $\gamma = 2.3$, a World Wide Web (www) onde $\gamma = 2.1$, e as citações em publicações científicas de $\gamma = 3$.

A aplicação do algoritmo para construção desta rede, envolve dois parâmetros, N representando o número de vértices de nossa rede e m que determina a quantidade de conexões que cada novo vértice adicionado à rede deve fazer. Um aspecto significativo relacionado a esse parâmetro é a capacidade de controlar o grau médio da rede, visto que, cada novo vértice adicionado estabelece m conexões, implicando que seu grau no instante em que finaliza as ligações seja igual a m . Além disso, esse vértice contribui com o acréscimo de um grau para outros m vértices que foram ligados, resultando na adição de $2m$ graus na somatória dos graus na rede. Portanto, se adicionarmos N vértices, a soma total dos graus será igual a $N \cdot 2m$. O grau médio da rede pode então ser calculado como $\langle k_g \rangle = (N \cdot 2m) / N = 2m$.

3.2.4 Modelos de comunidades

Para essa característica de rede, escolhemos trabalhar com a rede proposta por (LANCICHINETTI *et al*, 2008)^[9], que nomearemos neste trabalho como LFR. A escolha dessa rede foi fundamentada em sua capacidade de gerar redes que se assemelham mais fielmente às características observadas em redes reais. Isso é alcançado através da incorporação da heterogeneidade nas distribuições de grau dos vértices e nos tamanhos das comunidades, tornando-as mais representativas da complexidade das redes reais.

O processo de criação das redes LFR é conduzido por meio da atribuição cuidadosa de valores para seus parâmetros-chave, incluindo os expoentes γ e β , que desempenham um papel crucial na caracterização das distribuições de grau dos nós e dos tamanhos das comunidades, respectivamente, neste trabalho utilizaremos valores típicos de redes reais abordando: $2 \leq \gamma \leq 3$ e $1 \leq \beta \leq 2$. Adicionalmente, temos a flexibilidade de ajustar parâmetros como o número de vértices N , grau médio $\langle k \rangle$ e o coeficiente de conexão intercomunidades (μ), para criar uma diversidade de cenários, configurando uma ampla gama de características encontradas em redes complexas do mundo real.

A estratégia-chave na utilização desta rede está centrada no ajuste do parâmetro de conexão entre as comunidades, denotado como μ . Pretendemos investigar uma variedade de valores desse parâmetro, uma vez que, quando μ tende a zero, a rede revela uma conectividade limitada entre as comunidades. Em contrapartida, um valor mais próximo de 1 sugere uma maior interconexão entre as comunidades. Como exemplo visual da estruturação desta rede temos a figura 9.

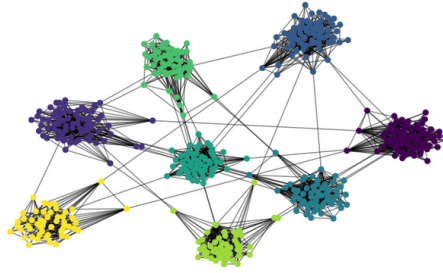


Figura 9: Representação de uma rede disposta em comunidades criada pelo algoritmo LFR. A estrutura possui 500 vértices, $\mu=0.03$, $\langle k \rangle=10$, $\gamma=3$ e $\beta=1.1$. Fonte: elaborada pelo autor.

3.2.5 Modelo de assortatividade

Para utilizar redes assortativas neste trabalho, foi utilizado como base o artigo de (Qiang Guo *et al*, 2006)^[10], onde é gerado redes de pequeno mundo com distribuição de grau livre de escala e com o coeficiente assortativo ajustável. Por simplicidade, neste trabalho iremos nos referir a essa rede como *Nass*.

Neste modelo, começamos com um conjunto inicial de vértices conectados e, em cada passo de tempo subsequente, adicionamos um novo vértice na rede. Este novo vértice é conectado aos outros em duas etapas: na primeira etapa, é conectado a vértices existentes seguindo uma probabilidade em função de seus graus; na segunda etapa, são adicionadas mais conexões a partir do novo vértice para vizinhos de vértices que foram conectados na primeira etapa, com base em uma probabilidade específica. Essa probabilidade depende de dois parâmetros: β , nomeado como o expoente preferencial, onde é utilizado para calcular a probabilidade de um vizinho ser escolhido; e P_t que é uma taxa para que a probabilidade de escolha de um vizinhos seja aceita. Esses parâmetros podem ser ajustados pelo usuário variando assim a assortatividade da rede.

Na figura 10, apresentamos os resultados obtidos por simulações onde demonstramos a eficácia no algoritmo para geração de redes tanto assortativas quanto disassortativas.

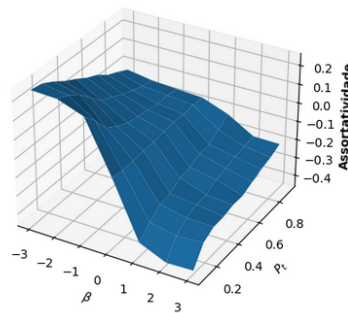


Figura 10 : Assortatividade obtida com a rede Nass. Para essa simulação utilizamos 6000 vértices, os parâmetros β e P_t foram variados na geração das redes resultando em diferentes medidas de assortatividade. Fonte: elaborada pelo autor.

4 RESULTADO

4.1 Convergência

Nesta análise, calculamos o índice de diversidade de Simpson a cada passo da dinâmica de Musso e Helbing, em todas as redes que serão utilizadas neste trabalho. Dessa forma, podemos verificar a existência de convergência para valores da diversidade.

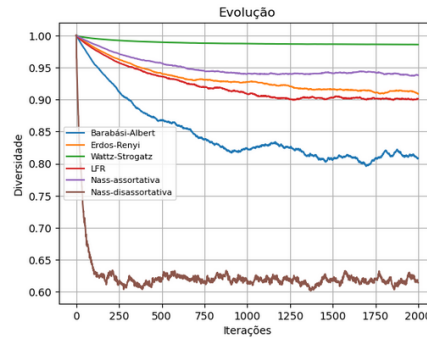


Figura 11: Evolução da diversidade em diferentes redes. Parâmetros utilizados: Barabási Albert ($N=5000$, $m=5$); Erdos Renyi ($N=5000$, $p=0.002$); LFR ($N=6000$, $\mu=0.025$, $\langle k \rangle=10$, $\gamma=3$, $\beta=1.1$); Watts Strogatz ($N=6000$, $p=0.005$, $k=4$); Nass assortativa ($N=5000$, $\beta=-3$, $P_t=0.1$); Nass disassortativa ($N=5000$, $\beta=3$, $P_t=0.1$). Fonte: elaborada pelo autor.

Com base nas simulações realizadas (figura 11), observamos que a diversidade converge para um estado estável após um período de variação. A identificação desse ponto de convergência ocorreu em até 2000 iterações, embora o tempo necessário para atingir a estabilização tenha variado de acordo com a rede utilizada e os parâmetros sob análise. Notavelmente, constatamos que, em todos os cenários analisados, 2000 iterações foram suficientes para alcançar a estabilização.

4.2 Grau médio

Para avaliar como a diversidade é influenciada pela variação dos graus médios, iremos explorar duas redes e conduzir uma comparação entre elas. Os modelos selecionados para este estudo são as redes Barabási-Albert e Erdos-Renyi, onde temos a capacidade de aplicar os graus próximos ao desejado.

Após a escolha das redes partiremos para a análise da diversidade sobre os modelos. Para isso, geramos redes de tamanhos e graus médios diferentes. A figura 12 apresenta os resultados em histogramas, representando a diversidade ao final de 2000 iterações com seu desvio padrão.

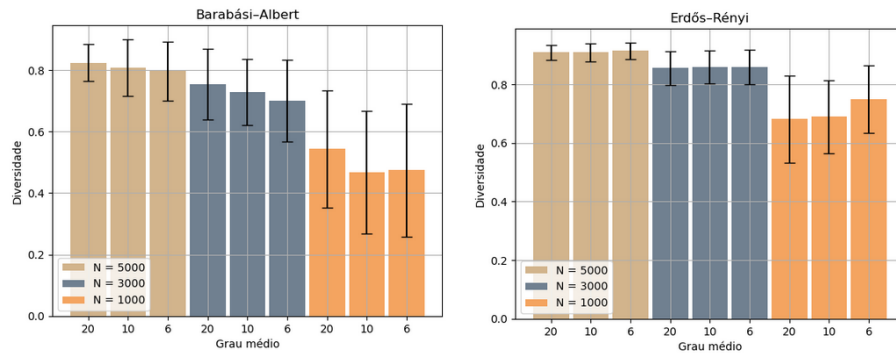


Figura 12: Diversidade obtida ao comparar as redes Barabási-Albert e Erdos-Rényi com graus médios e tamanhos de rede equivalentes por meio da aplicação do modelo de Musso e Helbing. Cada resultado foi obtido a partir de 100 evoluções e as barras de erro representam os desvios padrão obtidos. Fonte: elaborada pelo autor.

É notável, a partir da figura 12, que maiores diversidades são mantidas somente com redes maiores, visto que, mesmo ao gerar redes com graus médios substancialmente diferentes, a diversidade na rede não apresentou uma variação significativa. Em outras palavras, mesmo quando criamos redes com grau médio elevado ($\langle k \rangle = 20$) e outras com o grau médio mais baixo ($\langle k \rangle = 6$), com o objetivo de forçar uma variação significativa da diversidade, os resultados obtidos ainda se mostraram equivalentes, com o desvio padrão abrangendo as variações nas medidas. Entretanto, é importante ressaltar que ainda podemos observar algumas diferenças na diversidade obtida entre os modelos. Estas diferenças, no entanto, estão mais relacionadas à estrutura de formação das redes, um aspecto que será explorado em detalhes posteriormente.

4.3 Heterogeneidade

No contexto da análise do impacto da heterogeneidade na diversidade, revisaremos as redes Barabási-Albert e Erdos-Rényi. A escolha dessas redes se baseia em suas distintas estruturas, que resultam em distribuições de grau significativamente diferentes. A rede Barabási-Albert segue uma distribuição de lei de potência, enquanto a rede Erdos-Rényi segue uma distribuição de Poisson. Realizaremos as simulações mantendo o grau médio das redes inalterado, apenas aumentaremos seu tamanho.

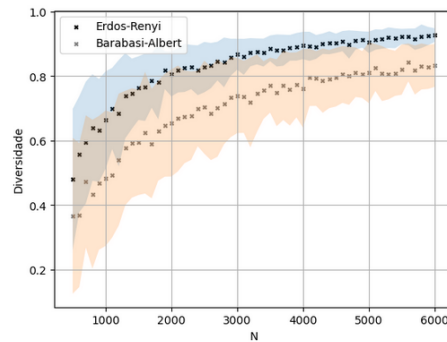


Figura 13: Comparação da diversidade nas redes Barabási Albert e Erdos Renyi com grau médio $\langle k \rangle = 10$. Simulação realizada com 100 evoluções para cada parâmetro, as faixas representam o desvio padrão. Fonte: elaborada pelo autor.

Na figura 13 temos o gráfico obtido das simulações, podemos notar novamente que a diversidade é influenciada pelo tamanho da rede, e também que, a homogeneidade de graus beneficiou a diversidade. Onde vemos, em média, a diversidade obtida na rede Erdos-Renyi apresentando resultados mais elevados em relação a Barabasi-Albert.

Como complemento podemos realizar a comparação com uma rede totalmente homogênea, utilizamos para isso o modelo de rede regular (uma rede onde todos os vértices possuem o mesmo grau). Nesta análise iremos realizar a evolução da dinâmica comportamental em redes de 10000 vértices e grau médio igual a 10.

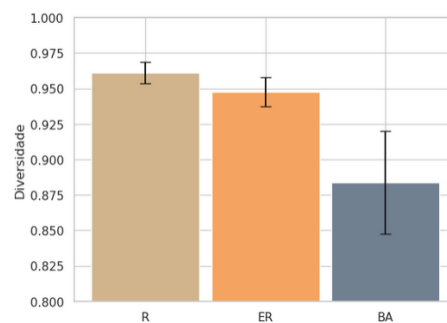


Figura 14: Comparação entre o modelo de rede Regular (R), Erdos Renyi (ER) e Barabási Albert (BA). Todas as redes possuem 10000 vértices, com grau médio igual a 10 e foram realizadas 2000 iterações. A barra de erro representa o desvio padrão. Os dados foram obtidos a partir da média de 30 redes de cada modelo. Fonte: elaborada pelo autor.

Nesse cenário, no qual temos uma rede tão homogênea quanto possível, observamos dos resultados representados na figura 14, que a diversidade é mais elevada em comparação com as redes analisadas anteriormente. Esse resultado fortalece a conclusão de que a homogeneidade da rede exerce uma influência significativa na diversidade dentro de uma rede, reforçando os resultados já obtidos por Musso e Helbing.

4.4 Distância e presença de triângulos

Neste contexto, vamos empregar a rede Watts-Strogatz, fazendo uma variação sistemática na probabilidade de reconexão da rede abordando as faixas discutidas no capítulo 3.2.2. Essa abordagem tem o propósito central de reduzir a média da distância mínima entre os nós e, ao mesmo tempo, manter a ocorrência de configurações triangulares.

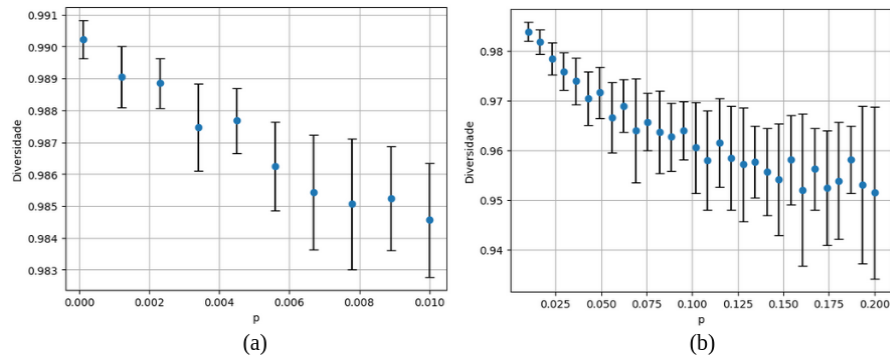


Figura 15: Em (a), a probabilidade está restrita à faixa de 0.0001 a 0.01, enquanto (b), abrange a faixa de 0.01 a 0.2. As redes são compostas por 6000 vértices, com cada vértice tendo 4 vizinhos ($\nu = 4$). Os dados foram obtidos a partir de 30 iterações para cada conjunto de parâmetros, e as barras de erro refletem o desvio padrão. Fonte: elaborada pelo autor.

No gráfico da figura 15-(a), onde p varia de 0.0001 a 0.01, é abordado uma região onde a transitividade da rede sofre pouca variação enquanto ocorre grande alteração da distância mínima média. Podemos observar que a diversidade não se manteve, ocorrendo uma diminuição de sua medida com a redução da distância l , sugerindo que maiores distâncias contribuem para uma maior diversidade. Já na figura 15-(b), onde estamos trabalhando com uma faixa de valores maiores para p , nas quais ocorre a diminuição na transitividade enquanto a distância l sofre pouca alteração, é notado uma queda ainda mais significativa na diversidade, indicando que a presença de triângulos na rede também contribui positivamente para a manutenção da diversidade.

4.5 Estrutura de comunidades

Para realizar a análise da influência da segregação de sociedades em comunidades na diversidade, empregaremos a rede LFR. A seguir, apresentamos os resultados das simulações com redes de dois tamanhos diferentes, à medida que o parâmetro μ aumenta.

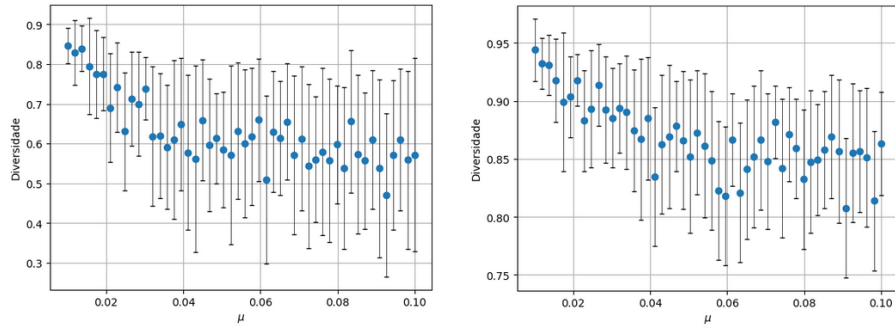


Figura 16: Dados de diversidade obtidos ao variar a conectividade entre comunidades. Os gráficos representam redes com 1000 e 5000 vértices, mantendo o grau médio fixo em 10, enquanto γ e β assumem valores de 3 e 1.1, respectivamente. Dados obtidos a partir da média de 30 evoluções. Fonte: elaborada pelo autor.

Nestes gráficos da figura 16, analisamos como a diversidade se comporta nas redes quando variamos a conectividade entre as comunidades. Para conduzir esse estudo, ajustamos o parâmetro de conexão em uma faixa que vai de 0.01 a 0.1, observando os resultados dentro desse intervalo. Através dos gráficos, podemos notar que a diversidade se manteve superior na rede de 5000 vértices, reforçando a influência positiva do tamanho da rede na diversidade. Ao explorarmos a faixa de valores de μ entre 0.01 e 0.1, identificamos mudanças significativas na rede até a proximidade de 0.04, onde ocorre uma redução da diversidade. Isso sugere que estruturas de comunidade mais robustas têm o efeito de aumentar a diversidade na rede, mas, à medida que a conectividade entre as comunidades se torna muito alta, essa diversidade é comprometida.

4.6 Correlação de graus

A rede utilizada nesta análise é a rede Nass. O objetivo aqui é examinar a maior variação de assortatividade. Para isso, podemos ver da figura 10 que ao fixarmos o parâmetro P_t em 0.1 e variarmos o parâmetro β , conseguimos obter a maior variação possível da assortatividade. Vamos então utilizar esses parâmetros em redes de diferentes tamanhos e aplicar a dinâmica de Musso e Helbing.

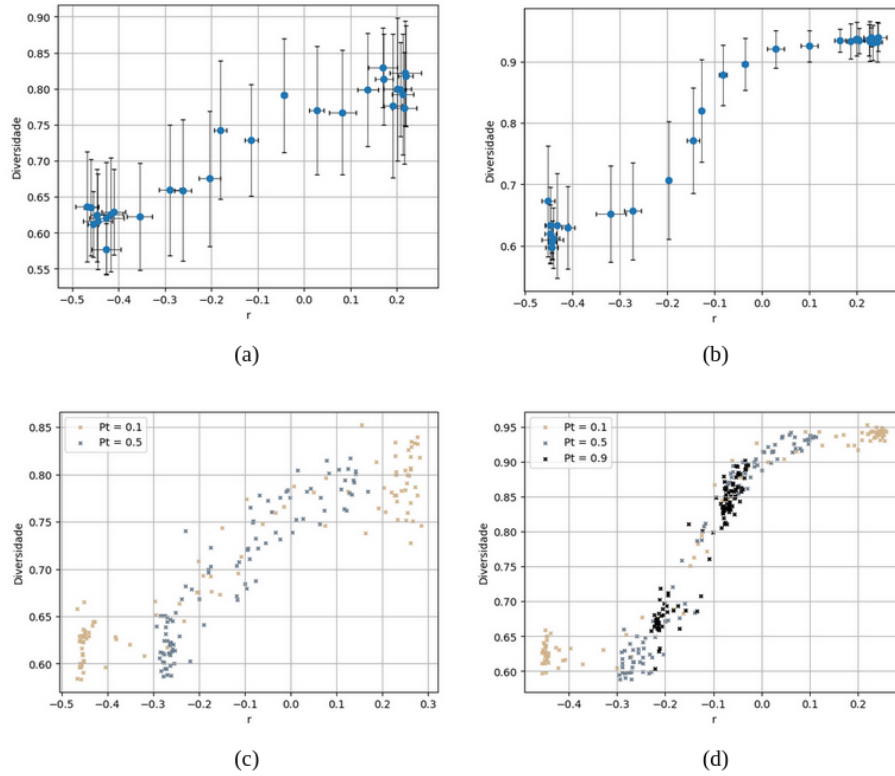


Figura 17: Dados da diversidade obtida ao variar a assortatividade da rede. Os gráficos (a) e (b) mostram a diversidade ao fixarmos P_t em 0.1 e variarmos β de -3 a 3 com $N=1000$ e 5000 vértices respectivamente. As barras de erro representam o desvio padrão da diversidade e da assortatividade das redes. Já os gráficos (c) e (d) representam os dados obtidos da diversidade para diferentes valores de P_t variando β na faixa de -3 a 3 e com 1000 e 5000 vértices. Dados obtidos a partir da média de 30 evoluções para cada parâmetro. Fonte: elaborada pelo autor.

Podemos observar na figura 17, que em redes menores, a diversidade exibe uma variação menor, e apresentando uma tendência crescente. E em redes maiores, temos uma faixa da diversidade maior, além de seus dados serem mais homogêneos. Podemos observar também a existência de uma região de transição na faixa de assortatividade entre -0.3 e 0.1, onde as variações na diversidade são mais notáveis. Tanto abaixo quanto acima desses valores, as variações na diversidade são menos pronunciadas, levando a regiões estagnadas.

Além disso, ao analisar o efeito da alteração do parâmetro P_t , podemos observar que seu aumento resulta em uma concentração na região de transição das regiões estagnadas, limitando ainda mais as variações na diversidade.

Podemos concluir, a partir dos dados obtidos, que as redes disassortativas demonstram ter mais efeito na diversidade em comparação a redes assortativas, onde é possível notar dentro da faixa de -0.3 até 0, que quanto mais disassortativas são as redes, menos diversidade temos. Já as redes assortativas demonstram ter menos efeito na diversidade, tendo sua influência

menos acentuada. No entanto, ainda observamos valores de diversidade mais elevados para as redes assortativas.

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Neste estudo, explorou-se o impacto das diversas características das redes complexas na preservação da diversidade de comportamentos e opiniões em uma sociedade. O ponto de partida foi o estudo e modelo de dinâmica comportamental originalmente proposto por (MUSSO *et al*, 2022)^[1], que promove o fluxo dos comportamentos individuais entre os membros de uma rede complexa. A investigação abrangeu as análises realizadas originalmente, como a homogeneidade de graus, a distância entre os vértices e a presença de triângulos. Além disso, cenários inexplorados anteriormente foram abordados, verificando o efeito de variáveis-chave, como o grau médio de conexões, o tamanho da rede, a presença de estruturas de rede assortativas e disassortativas, bem como a existência de agrupamentos.

Nas análises iniciais, o efeito da heterogeneidade na distribuição de graus nas redes foi observado, comparando os modelos de Barabási-Albert e Erdos-Renyi. Constatou-se que a homogeneidade de graus beneficia a diversidade, especialmente em estruturas sociais maiores. A análise também abordou como a distância mínima média e a presença de triângulos na rede afetam a diversidade, verificando que maiores distâncias e a presença de triângulos na rede contribuem positivamente para uma maior diversidade, resultados semelhantes ao estudo de Musso e Helbing.

A investigação do efeito do grau médio resultou em resultados inesperados, com uma relação fraca observada, sendo as diferenças entre os resultados obtidos das simulações menores que o desvio padrão. Mesmo ao tentar criar redes com valores elevados de ligações entre os vértices, não se obtiveram resultados significativos que indicassem uma dependência da diversidade em relação a esse parâmetro.

Ao examinar o efeito da segregação em comunidades, utilizando a rede LFR com variações no parâmetro de conectividade entre as comunidades, constatou-se que, para valores de conectividade entre as comunidades acima de 0.04, a diversidade permaneceu constante. No entanto, para valores de conectividade entre 0.01 e 0.04, observaram-se mudanças significativas na diversidade, indicando que a segregação entre comunidades beneficia a multiplicidade de comportamentos, mantendo a diversidade elevada.

A análise da correlação de graus nas redes revelou que a diversidade apresenta variações mais notáveis em uma região de transição na faixa de assortatividade entre -0.3 e 0.1. Esses resultados indicaram que estruturas sociais em que ocorre a conexão entre indivíduos com graus de conexão semelhantes auxiliam na manutenção da diversidade elevada, enquanto o

oposto, onde ocorre a mistura entre indivíduos de alto e baixo grau de conexão, inibe a diversidade.

Em todas as análises, observou-se uma diversidade mais significativa em redes de maior porte, e também identificou-se um ponto de convergência para essa diversidade. Dependendo dos parâmetros da rede, muitas vezes, um número relativamente baixo de iterações é suficiente para atingir esse estado de estabilidade na diversidade.

Portanto, as análises destacam a importância de considerar fatores como o tamanho da rede, a segregação em comunidades e a assortatividade ao analisar a dinâmica da diversidade em sistemas complexos. Essas conclusões têm implicações significativas para a compreensão de fenômenos sociais, como a propagação de informações e opiniões em redes sociais, a formação de consenso e a polarização de opiniões em sociedades interconectadas.

REFERÊNCIAS

- [1]MUSSO, Andrea; HELBING, Dirk. How networks shape diversity for better or worse. **arXiv preprint arXiv:2201.09254**, 2022.
- [2]SIMPSON, Edward H. Measurement of diversity. **nature**, v. 163, n. 4148, p. 688-688, 1949.
- [3]NEWMAN, Mark EJ. Mixing patterns in networks. **Physical review E**, v. 67, n. 2, p. 026126, 2003.
- [4]ARCAGNI, Alberto et al. Higher order assortativity in complex networks. **European Journal of Operational Research**, v. 262, n. 2, p. 708-719, 2017.
- [5]ERDŐS, P.; R&WI, A. On random graphs I. **Publ. math. debrecen**, v. 6, n. 290-297, p. 18, 1959.
- [6]ERDŐS, Paul et al. On the evolution of random graphs. **Publ. math. inst. hung. acad. sci**, v. 5, n. 1, p. 17-60, 1960.
- [7]WATTS, Duncan J.; STROGATZ, Steven H. Collective dynamics of small-world networks. **nature**, v. 393, n. 6684, p. 440-442, 1998.
- [8]BARABÁSI, Albert-László; ALBERT, Réka. Emergence of scaling in random networks. **science**, v. 286, n. 5439, p. 509-512, 1999.
- [9]LANCICHINETTI, Andrea; FORTUNATO, Santo; RADICCHI, Filippo. Benchmark graphs for testing community detection algorithms. **Physical review E**, v. 78, n. 4, p. 046110, 2008.
- [10]GUO, Qiang et al. Growing scale-free small-world networks with tunable assortative coefficient. **Physica A: Statistical Mechanics and its Applications**, v. 371, n. 2, p. 814-822, 2006.