



PUC Minas

Pontifícia Universidade Católica de Minas Gerais
Departamento de Ciência da Computação
Redes Complexas

Redes Complexas

(assortatividade)

Prof.: Felipe Domingos
felipe@pucminas.br

Homofilia

- “Tendência de um pessoa se associar a outra com propriedades similares.”
(gosto, sexo, meio social, ...)

Homofilia

- “Tendência de um nó se conectar a outro com propriedades similares.”
(grau, coeficiente de agrupamento,
localização,...)

Homofilia

- Uma propriedade comum em muitas redes complexas é seu crescimento seguir uma lei de potência.
- Isso ocorre pois certos nós com maiores recursos podem atrair mais recurso ainda:
 - ◆ ricos ficam mais ricos
 - ◆ populares ficam mais populares
 - ◆ ...

Homofilia

- Esse fenômeno é conhecido como **ligação preferencial** e diz que ao surgir um novo nó na rede, este tem uma preferência em se conectar com outros nós com determinada propriedade.
 - ◆ Ex.: se conectar com nós de maior grau

Homofilia

- Duas situações podem ser observadas em redes reais:
 - ◆ Os nós de maior grau preferem se conectar com nós similares (assortativos ou homofílicos);
 - ◆ Os nós de maior grau preferem se conectar com nós diferentes (disassortativos ou heterofílicos).

Homofilia

- Em redes reais pode se pensar em:
 - ◆ Pessoas em uma rede social se conectam a pessoas com mesma popularidade;
 - ◆ Portais agregadores na Internet contém links para sites com poucos links (geradores de conteúdo).

Assortatividade

- A **ASSORTATIVIDADE** r mede justamente se uma rede tem preferência por determinadas conexões ou não.
- Esse valor varia de -1 a 1 e representa:
- $r > 0$: os nós tendem a se conectar com outros nós de grau similar.
- $r < 0$: os nós tendem a se conectar com outros nós de graus diferentes (grau alto com grau baixo).
- $r \sim 0$: os nós não tem preferência.

Assortatividade

- O coeficiente de ASSORTATIVIDADE de uma rede é calculada com a correlação de Pearson entre os graus dos pares de nos conectados entre si

$$r = \frac{1}{\sigma_o \sigma_d} \sum_{k_1, k_2} k_1 k_2 (P(k_1, k_2) - P(k_1)P(k_2))$$

Assortatividade

- $P(k_1, k_2)$ representa quantas vezes ocorre uma aresta em que o nó de origem tem grau = k_1 e o nó de destino tem grau = k_2 dividido pelo número de arestas (ou $2 * |E|$ caso o grafo seja não direcionado)
- $P(k_1)$ representa quantas arestas tem nó de origem com grau igual a k_1 dividido pelo número de arestas
- $P(k_2)$ representa quantas arestas tem nó de destino com grau igual a k_2 dividido pelo número de arestas

Assortatividade

$$P(k_1, k_2) = \frac{|(v, u) \in E | grau(v) = k_1 \wedge grau(u) = k_2|}{|E|}$$

$$P(k_1) = \frac{|(v, u) \in E | grau(v) = k_1|}{|E|}$$

$$P(k_2) = \frac{|(v, u) \in E | grau(u) = k_2|}{|E|}$$

Assortatividade

- Desvio-padrão da distribuição do grau dos nós de origem das arestas

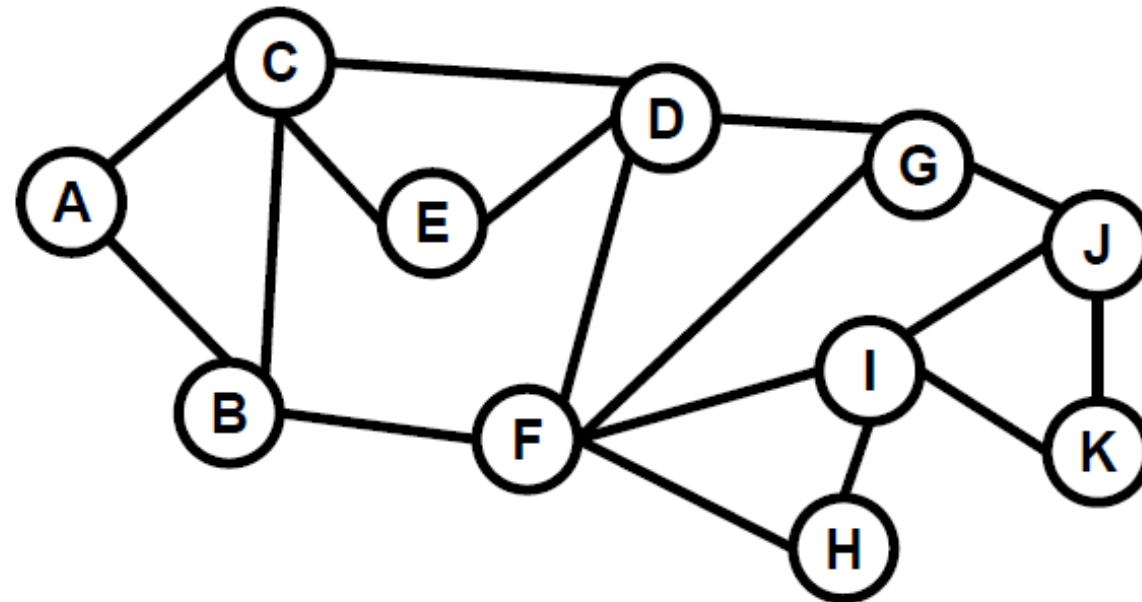
$$\sigma_o = \sqrt{\sum_{k_1} k_1^2 P(k_1) - \left(\sum_{k_1} k_1 P(k_1) \right)^2}$$

- Desvio-padrão da distribuição do grau dos nós de destino das arestas

$$\sigma_d = \sqrt{\sum_{k_2} k_2^2 P(k_2) - \left(\sum_{k_2} k_2 P(k_2) \right)^2}$$

Assortatividade

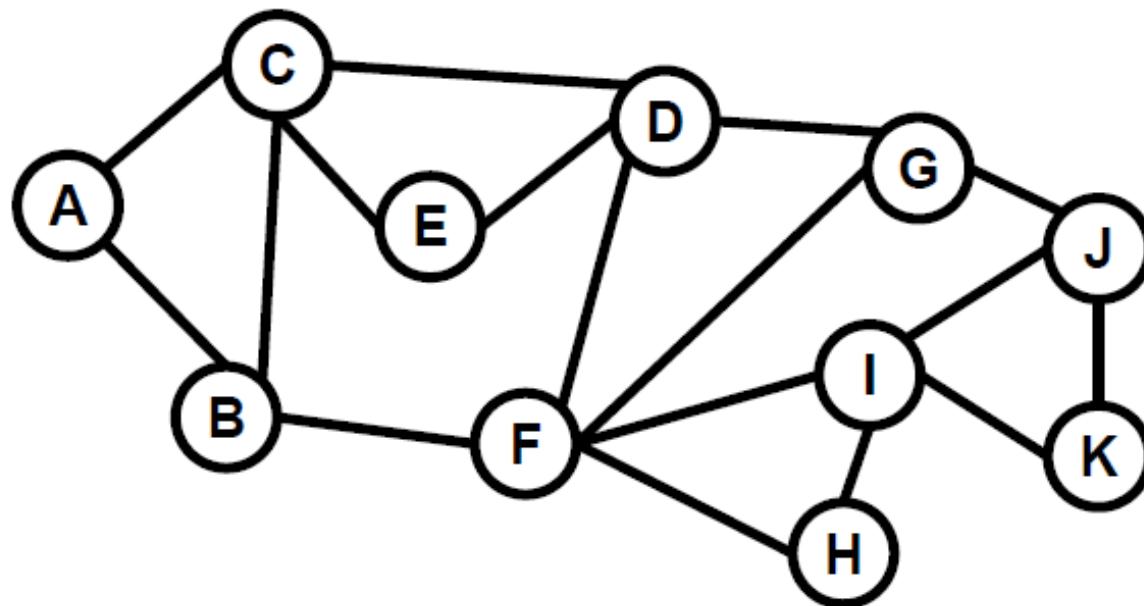
- Qual a assortatividade desta rede:



Assortatividade

- Os graus de entrada de cada nó são:

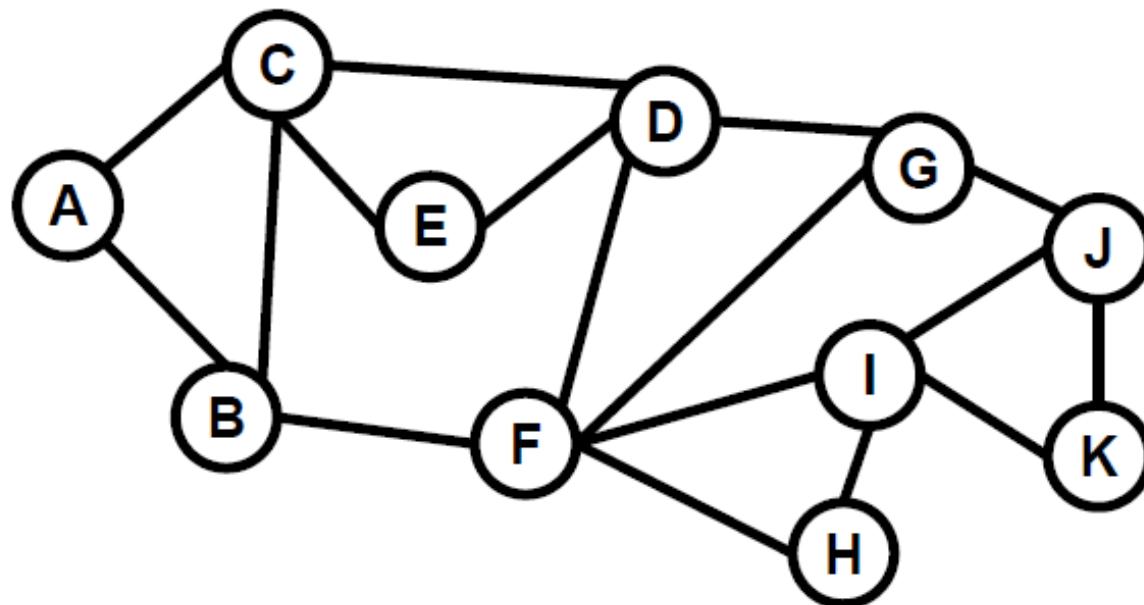
A	2
B	3
C	4
D	4
E	2
F	5
G	3
H	2
I	4
J	3
K	2



Assortatividade

- Os graus de saída de cada nó são:

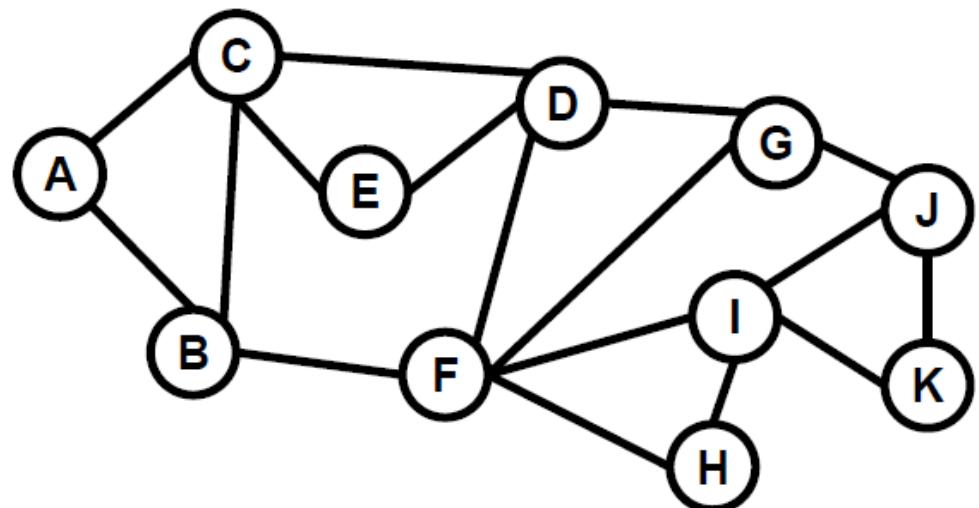
A	2
B	3
C	4
D	4
E	2
F	5
G	3
H	2
I	4
J	3
K	2



Assortatividade

- Enumerando, temos nós com grau igual a 2, 3, 4 e 5. Vamos relacionar quantas arestas interligam o grau x ao grau y:

	2	3	4	5
2	0	2	5	1
3	2	2	3	2
4	5	3	2	2
5	1	2	2	0

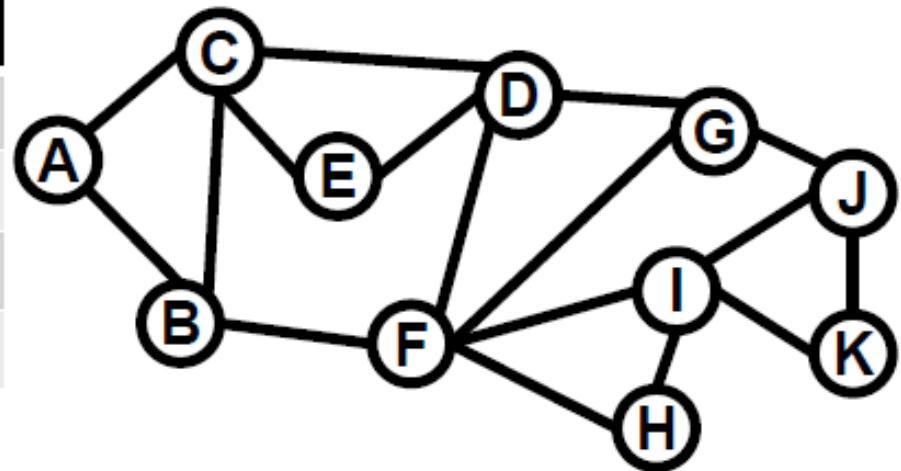


- Temos um total de 34 arestas ($2 \cdot 17$, pois é não-direcionado).

Assortatividade

- Dividimos os valores da tabela por 34 para obtermos $P(k_1, k_2)$:

$P(k_1, k_2)$	2	3	4	5
2	0,00	0,06	0,15	0,03
3	0,06	0,06	0,09	0,06
4	0,15	0,09	0,06	0,06
5	0,03	0,06	0,06	0,00



Assortatividade

- A soma das linhas resulta em $P(k_1)$ e a soma das colunas resulta em $P(k_2)$:

$P(k_1, k_2)$	2	3	4	5	$P(k_1)$
2	0,00	0,06	0,15	0,03	0,24
3	0,06	0,06	0,09	0,06	0,27
4	0,15	0,09	0,06	0,06	0,36
5	0,03	0,06	0,06	0,00	0,15
$P(k_2)$	0,24	0,27	0,36	0,15	1

Assortatividade

- Com esses dados podemos calcular:

$$R(k_1, k_2) = k_1 k_2 (P(k_1, k_2) - P(k_1)P(k_2))$$

P(k1,k2)	2	3	4	5	P(k1)
2	0,00	0,06	0,15	0,03	0,24
3	0,06	0,06	0,09	0,06	0,27
4	0,15	0,09	0,06	0,06	0,36
5	0,03	0,06	0,06	0,00	0,15
P(k2)	0,24	0,27	0,36	0,15	1

R(k1,k2)	2	3	4	5
2	-0,23	-0,03	0,51	-0,06
3	-0,03	-0,12	-0,09	0,29
4	0,51	-0,09	-1,11	0,12
5	-0,06	0,29	0,12	-0,56

Assortatividade

- Com esses dados podemos calcular:

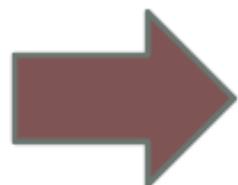
$$\sum_{k1,k2} R(k_1, k_2) = \sum_{k1,k2} k_1 k_2 (P(k_1, k_2) - P(k_1)P(k_2)) \\ = -0,54$$

R(k1,k2)	2	3	4	5
2	-0,23	-0,03	0,51	-0,06
3	-0,03	-0,12	-0,09	0,29
4	0,51	-0,09	-1,11	0,12
5	-0,06	0,29	0,12	-0,56

Assortatividade

- O desvio-padrão deve ser calculado em relação aos graus e quantas vezes eles ocorrem:

	2	3	4	5
2	0	2	5	1
3	2	2	3	2
4	5	3	2	2
5	1	2	2	0
SOMA:	8	9	12	5

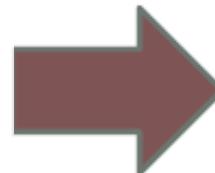


$$\begin{aligned} \text{MÉDIA} = \\ 2*8+3*9+4*12+5*5 / 34 \\ = 3,41 \end{aligned}$$

Assortatividade

- O desvio-padrão deve ser calculado em relação aos graus e quantas vezes eles ocorrem:

	2	3	4	5
2	0	2	5	1
3	2	2	3	2
4	5	3	2	2
5	1	2	2	0
SOMA:	8	9	12	5



$$\begin{aligned} \text{VARIÂNCIA} = & \\ & 8*(2-3,41)^2+9*(3-3,41)^2+1 \\ & 2*(4-3,41)^2+5*(5-3,41)^2 / \\ & 34= \\ & 34,23 / 34 = \\ & 1,0067 \end{aligned}$$

$$\begin{aligned} \text{DESVIO-PADRÃO} = & \\ & \sqrt{1,0067} = 1,0034 \end{aligned}$$

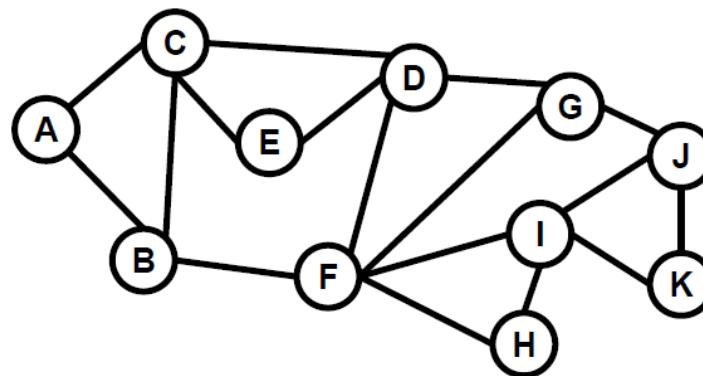
Assortatividade

- Finalmente, o cálculo do coeficiente de assortatividade será:

$$\begin{aligned} r &= \frac{1}{\sigma_o \sigma_d} \sum_{k_1, k_2} k_1 k_2 (P(k_1, k_2) - P(k_1)P(k_2)) = \\ &= \frac{1}{1,0034 \cdot 1,0034} \cdot -0,54 = -0,5363 \end{aligned}$$

Assortatividade

- Esse valor é apenas aproximado, pois arredondamos os cálculos para duas casas decimais.



- O valor correto da assortatividade dessa rede é -0,58; que ainda indica uma disassortatividade, ou heterofilia.

Assortatividade em redes reais

Network	<i>n</i>	<i>r</i>
Physics coauthorship (a)	52 909	0.363
Biology coauthorship (a)	1 520 251	0.127
Mathematics coauthorship (b)	253 339	0.120
Film actor collaborations (c)	449 913	0.208
Company directors (d)	7 673	0.276
Internet (e)	10 697	-0.189
World-Wide Web (f)	269 504	-0.065
Protein interactions (g)	2 115	-0.156
Neural network (h)	307	-0.163
Marine food web (i)	134	-0.247
Freshwater food web (j)	92	-0.276

SOCIAIS

Biológicas
e
Tecnológicas

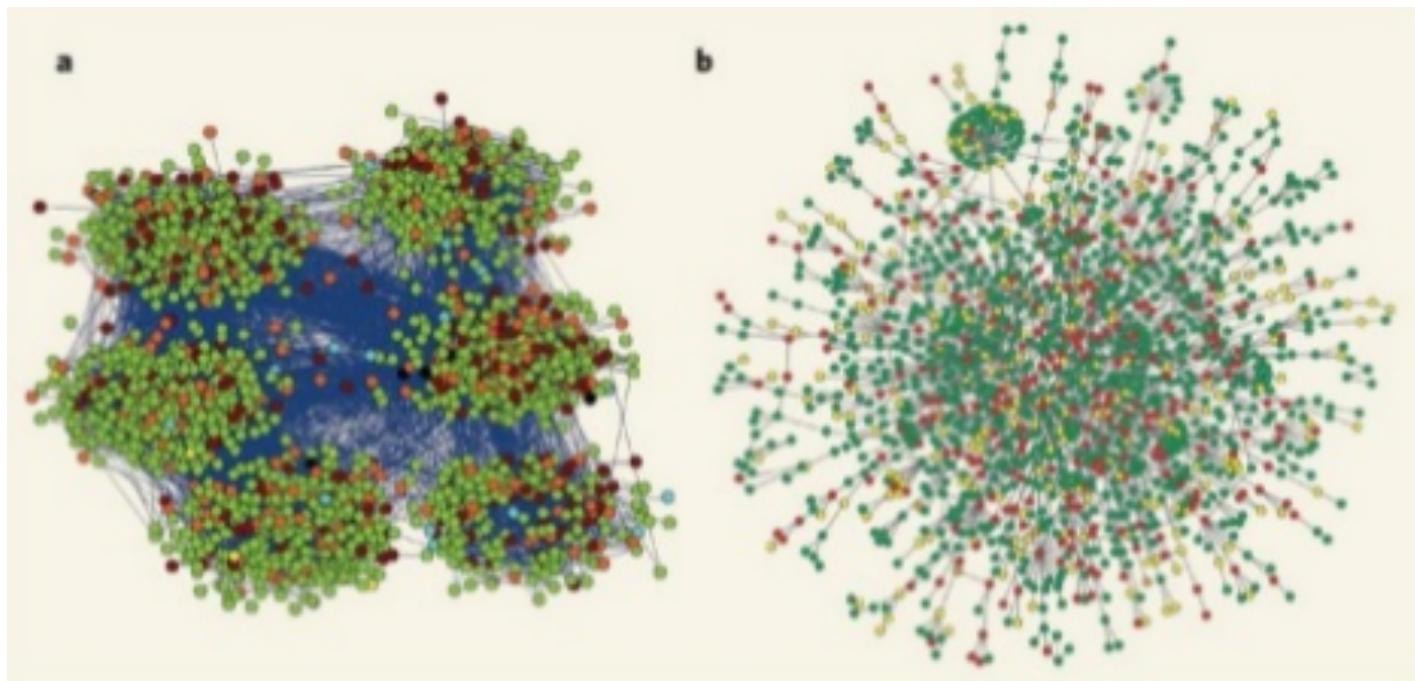
Outras Assortatividades

- A assortatividade também pode ser mensurada em relação a outros aspectos da rede, não apenas grau.
- Em um estudo verificou-se que a rede de conversas no twitter (respostas de um usuário para outro) tem assortatividade positiva em relação aos tweets “alegres”.
- Catherine A. Bliss, Isabel M. Kloumann, Kameron Decker Harris, Christopher M. Danforth, Peter Sheridan Dodds: Twitter reciprocal reply networks exhibit assortativity with respect to happiness CoRR abs/1112.1010: (2011)

Outras Assortatividades

- Além disso pode-se medir assortatividade através de qualquer tipo de cálculo de correlação.
- Vimos a assortatividade com correlação de Pearson, que mede uma correlação linear.
- Outra medida muito utilizada é a correlação de *Spearman*, que mede a tendência de que, se uma variável aumenta, a outra também aumenta, sem necessidade de ter uma correlação linear.

Exemplos



Exemplos

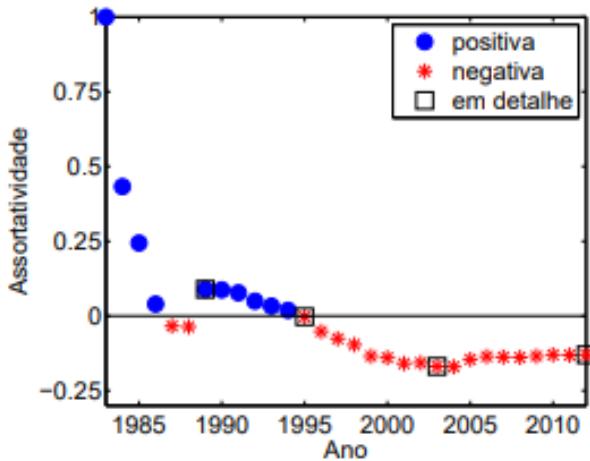


Figura 5. A assortatividade da rede ao longo do tempo. Observa-se que a rede torna-se disassortativa com o tempo, indicando que nós de grau baixo têm maior tendência de se conectarem a nós de grau alto. O comportamento do $\langle k_{nn} \rangle_k$ das redes, representado pelos pontos marcados como “em detalhe”, é mostrado na figura 2

Assortatividade e Tolerância a Falhas

- Em um artigo um pesquisador fez o seguinte experimento, gerando:
 - N redes aleatórias que apresentavam assortatividade ($r>0$)
 - redes aleatórias que apresentavam disassortatividade ($r<0$)
 - N redes aleatórias que não apresentavam assortatividade ($r=0$)

Assortatividade e Tolerância a Falhas

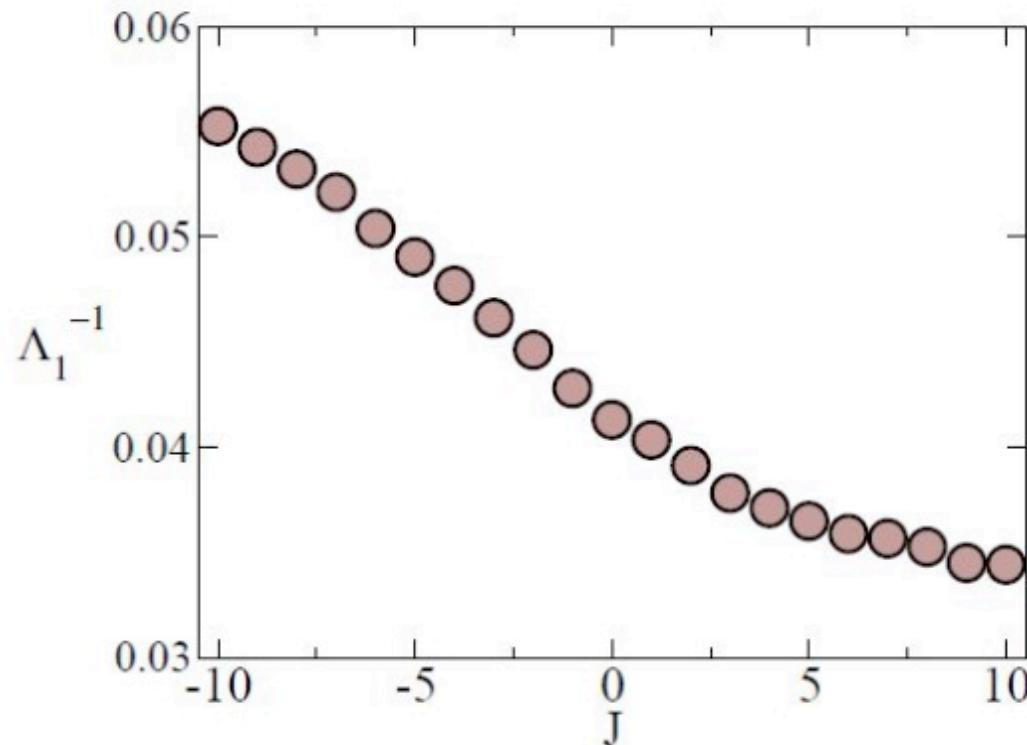
- Adotou o seguinte procedimento:
- com probabilidade p (variando entre 0,1 e 0,9) marcava cada uma das arestas como escolhida (ou não).
- as arestas escolhidas representam arestas que vão difundir certa informação ou serão removidas por falha ou ataque.

Assortatividade e Tolerância a Falhas

- Verificou-se o **ponto crítico** de cada rede:
- probabilidade em que a remoção das arestas selecionadas implica em quase **desconexão da rede**
- probabilidade em que a passagem de informação pelas arestas selecionadas implica que **quase todos os nós receberam a informação**

Assortatividade e Tolerância a Falhas

- Ponto crítico desconexão



Assortatividade e Tolerância a Falhas

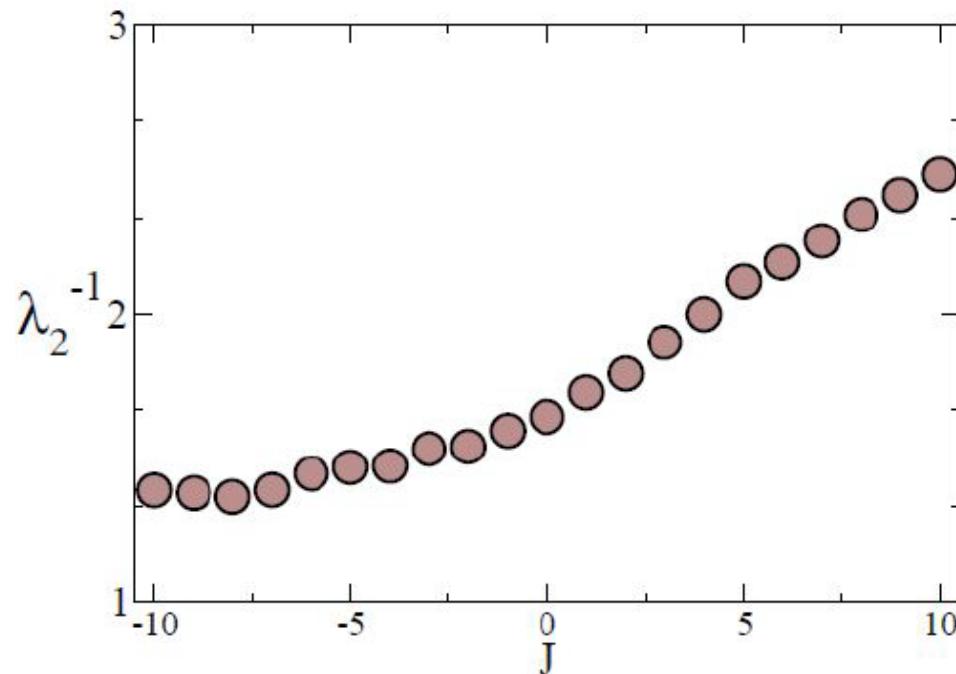
- **Redes assortativas:**
- Tem um ponto crítico mais baixo, é desconectada mais facilmente.
- Não precisa de muito esforço para transmitir e espalhar informações.
- É desconectada rapidamente em ataques aleatórios.

Assortatividade e Tolerância a Falhas

- **Redes disassortativas:**
- Tem um ponto crítico mais alto, demora mais para desconectar.
- Tem um custo maior para a transmissão de informação.
- Resistentes a ataques aleatórios.

Assortatividade e Tolerância a Falhas

- Velocidade do espalhamento da informação em função da assortatividade



Assortatividade e Tolerância a Falhas

Redes assortativas:

- A velocidade de transmissão é menor, portanto temos mais tempo para medidas preventivas.

Redes disassortativas:

- A velocidade de transmissão é maior, em caso de espalhamento de informação será difícil conter a mesma.

Assortatividade e Tolerância a Falhas

- As Redes Sociais são, em geral, **assortativas!!!**
- Logo elas são capazes de transmitir as doenças mais facilmente, são mais difíceis de imunizar (vacinar poucos nós para desconectar a transmissão), porém existe um tempo maior para o processo de imunização.

Assortatividade e Tolerância a Falhas

- As Redes Tecnológicas são, em geral, **disassortativas!!!**
- Logo falhas no sistema dificilmente atinge boa parte da rede, são mais fáceis de identificar os pontos mais importantes para proteger mas, se um ataque direcionado ocorrer, a desconexão é bem mais difícil.

Saber mais....

- Noh, J.D. Percolation transition in networks with degree-degree correlation. *Physical Review E*, v.76, n.2. 2007.
- Newman, M.E.J. Assortative mixing in networks. *Physical Review Letters*, v. 89, n. 20. 2002.
- G. D'Agostino, A. Scala, V. Zlatic, G. Caldarelli. Robustness and Assortativity for Diffusion-like Processes in Scale-free Networks, 2012