

# ANÁLISE DE REDES

## Redes Aleatórias

Licenciatura em  
Ciência de Dados

1

1

## Redes Aleatórias

- \* Muitas redes reais partilham um conjunto de propriedades:
  - \* os caminhos mais curtos entre dois nodos têm poucas ligações;
  - \* têm muitos triângulos, tendo por isso coeficientes de *clustering* elevados;
  - \* apresentam heterogeneidade quando se consideram os graus dos nodos e os pesos associados às ligações.

2

2

1

## Redes Aleatórias

- \* Paul Erdős e Alfréd Rényi iniciaram o estudo da Teoria de Grafos Aleatórios, para estudar como surgem estas propriedades.
- \* Este estudo é baseado nas redes aleatórias, também conhecidas como redes de Erdős-Rényi.

3

3

## Redes Aleatórias

- \* Uma rede de Erdős-Rényi é obtida escolhendo o número de nodos ( $N$ ) e o número de ligações ( $L$ ).
- \* Para escolher as ligações, geram-se aleatoriamente  $L$  pares de nodos.

4

4

2

## Redes Aleatórias

- \* Outras redes aleatórias, propostas por E. N. Gilbert, podem ser obtidas escolhendo o número de nodos e uma probabilidade de inclusão de cada ligação ( $p$ ).

5

5

## Redes Aleatórias

- \* Neste caso, as redes serão obtidas aplicando o seguinte procedimento:
  - \* 1 - seleccione-se um par de nodos  $(i, j)$ ;
  - \* 2 - gere-se um número aleatório  $r$ , entre 0 e 1.  
Se  $r < p$  então inclui-se a ligação  $(i, j)$ ;
  - \* 3 - repetem-se os passos 1 e 2 para todos os pares de nodos.

6

6

3

## Redes Aleatórias

- \* A aplicação deste procedimento, diversas vezes, para valores fixos de  $N$  e  $p$  pode gerar redes com diferentes números de ligações.
- \* Contudo, para valores de  $N$  suficientemente grandes, espera-se que o número de ligações seja aproximado.

7

7

## Redes Aleatórias

- \* Suponha agora que se está a gerar uma rede aleatória.
- \* Inicialmente a rede é constituída apenas por nodos.
- \* À medida que se adicionam ligações, pares de nodos são ligados.
- \* Após a introdução das primeiras ligações, a rede é composta por subredes de pequena dimensão.

8

8

4

## Redes Aleatórias

- \* Quando se forma a componente gigante?
- \* Erdős e Rényi descobriram que a componente gigante forma-se quando  $\langle k \rangle = 1$ .
- \* A configuração da rede muda repentinamente quando  $\langle k \rangle$  passa de um valor inferior a 1 para igual a 1.
- \* Por outro lado, a dimensão desta componente gigante cresce rapidamente com o aumento do grau médio (para valores superiores a 1).
- \* Ver NetLogo Giant Component

9

9

## Redes Aleatórias

- \* **Densidade das Redes Aleatórias**
- \* Considere-se o método de geração de uma rede aleatória baseado na escolha do número de nodos e da probabilidade de inclusão de cada ligação (Modelo de Gilbert).
- \* Este processo de geração de uma rede aleatória é semelhante à experiência de lançar repetidas vezes uma moeda.

10

10

5

## Redes Aleatórias

- \* **Densidade das Redes Aleatórias**
- \* Suponha-se que  $p$  representa a probabilidade de obter uma cara.
- \* O número esperado de caras obtidas será dado por  $pn$ , em que  $n$  representa o número de lançamentos.

11

11

## Redes Aleatórias

- \* **Densidade das Redes Aleatórias**
- \* No caso da rede aleatória, o número de lançamentos corresponde a

$$\binom{N}{2} = \frac{N(N-1)}{2}$$

que representa o número máximo de ligações de uma rede não orientada com  $N$  nodos.

12

12

# Redes Aleatórias

- \* **Densidade das Redes Aleatórias**

- \* Então:

$$\langle L \rangle = \frac{pN(N-1)}{2}.$$

- \* Atendendo à expressão para o grau médio, tem-se:

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1).$$

13

13

# Redes Aleatórias

- \* **Densidade das Redes Aleatórias**

- \* Além disso, a densidade será dada por:

$$d = \frac{\langle L \rangle}{L_{max}} = \frac{pN(N-1)/2}{N(N-1)/2} = p.$$

- \* Como as redes reais são esparsas, uma rede aleatória adequada para o estudo de redes reais terá associada uma probabilidade pequena.

14

14

## Redes Aleatórias

- \* **Distribuição de Grau das Redes Aleatórias**
- \* A distribuição de grau de uma rede aleatória será determinada pelo cálculo das probabilidades de um qualquer nodo da rede ter  $k$  adjacentes.
- \* Seja  $i$  um nodo qualquer de uma rede com  $N$  nodos. Qualquer um dos restantes  $N - 1$  nodos pode ser adjacente de  $i$ .

15

15

## Redes Aleatórias

- \* **Distribuição de Grau das Redes Aleatórias**
- \* Para que cada um dos restantes nodos seja adjacente de  $i$ , é necessário que exista uma ligação entre  $i$  e o nodo.
- \* No processo de geração da rede aleatória, a decisão de incluir uma ligação é independente da decisão de incluir ou não incluir cada uma das restantes.

16

16



# Redes Aleatórias

## \* Distribuição de Grau das Redes Aleatórias

- \* Se a probabilidade de dois nodos serem adjacentes for  $p$  então a probabilidade de um nodo ter grau igual a  $k$  é dada por:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

17

17

# Redes Aleatórias

## \* Distribuição de Grau das Redes Aleatórias

- \* Trata-se da distribuição Binomial, de parâmetros  $(N-1)$  e  $p$ , que tem valor esperado dado por  $(N-1)p$ .
- \* Assim, para valores elevados de  $N$  e com  $Np$  constante (e não muito pequeno) pode considerar-se que  $Np \approx \langle k \rangle$ .

18

18

## Redes Aleatórias

- \* **Distribuição de Grau das Redes Aleatórias**
- \* Na distribuição Binomial os valores que acumulam maior probabilidade estão concentrados à volta do valor esperado.
- \* Além disso, o gráfico que representa as probabilidades não apresenta uma cauda longa.
- \* Podemos então concluir que nas redes aleatórias não se verifica a presença de hubs.

19

19

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Já se viu que os graus dos nodos de uma rede aleatória não diferem muito.
- \* Para estudar a existência de caminhos mais curtos com poucas ligações, considere-se uma rede conexa em que todos os nodos têm grau  $k$ .

20

20

10

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Considere-se ainda um nodo qualquer da rede.
- \* Este nodo está:
  - \* a uma distância igual a 1 de  $k$  nodos;
  - \* a uma distância igual a 2 de  $k(k - 1)$  nodos;
  - \* a uma distância igual a 3 de  $k(k - 1)^2$  nodos.

21

21

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Assim, o nodo está a uma distância igual a  $l$  de  $k(k - 1)^{l-1}$  nodos.
- \* O valor real de nodos pode ser inferior a  $k(k - 1)^{l-1}$  porque alguns dos nodos a uma distância não superior a  $l$  podem estar repetidos.

22

22

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Se o valor de  $k$  não for muito pequeno então pode considerar-se  $k \approx k - 1$  e a estimativa para o número de nodos a uma distância igual a  $l$  será aproximada por  $k^l$ .
- \* Seja  $l_{max}$  a maior das distâncias - o diâmetro da rede -, então esta distância permite alcançar todos os nodos da rede e

$$k^{l_{max}} = N.$$

23

23

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Tem-se então  $l_{max} = \log_k N = \frac{\log N}{\log k}$ .
- \* Dado que os graus dos nodos de uma rede aleatória não diferem muito, a expressão anterior constitui uma boa aproximação quando se tomam valores de  $k$  em torno de  $\langle k \rangle$ .

24

24

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Como a função logarítmica cresce lentamente, o valor de  $l_{max}$  cresce lentamente quando  $N$  aumenta.
- \* Conclui-se então que nas redes aleatórias existem caminhos mais curtos com poucas ligações.

25

25

## Redes Aleatórias

- \* **Caminhos mais curtos**
- \* Considere-se uma rede aleatória para representar a rede mundial de contactos sociais.
- \* Com base no número de Dunbar, considere-se ainda que o grau médio é 150.
- \* Tem-se que  $150^5 = 75\,937,5$  milhões, que é nove vezes superior à população mundial.
- \* Este resultado é compatível com a experiência de Stanley Milgram.

26

26

## Redes Aleatórias

- \* **Coeficiente de *Clustering***

- \* O coeficiente de *clustering* de um nodo mede a fracção de nodos adjacentes unidos por uma ligação.
- \* Uma ligação entre dois nodos adjacentes forma um triângulo.

27

27

## Redes Aleatórias

- \* **Coeficiente de *Clustering***

- \* Numa rede aleatória, gerada considerando o número de nodos e uma probabilidade  $p$  de inclusão de cada ligação, a probabilidade de dois nodos adjacentes estarem unidos é igual a  $p$ .
- \* O coeficiente de *clustering* de cada nodo pode não ser exatamente igual a  $p$ .
- \* Contudo, espera-se que a média dos coeficientes de *clustering* dos nodos seja bem aproximada por  $p$ .

28

28

## Redes Aleatórias

- \* **Coeficiente de *Clustering***

- \* Se a rede aleatória for esparsa então o valor de  $p$  será pequeno e a rede terá poucos triângulos.
- \* Se se aumentar o valor da probabilidade então o número de triângulos aumenta, assim como a densidade da rede (o que não será realista para simular redes sociais reais).

29

29

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**

- \* Com vista a obter redes aleatórias com coeficientes de *clustering* mais elevados, Duncan J. Watts e Steven H. Strogatz desenvolveram o Modelo de Mundos Pequenos (*small-world model*), também conhecido como modelo Watts-Strogatz.

30

30

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* Começaram por gerar uma rede em que cada nodo está ligado aos 4 nodos mais próximos. Esta rede apresenta um coeficiente de *clustering* de 0,5.
- \* Contudo, a distância média não é pequena, devido às distâncias entre os nodos mais afastados entre si.

31

31

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* Para reduzir a distância média, algumas ligações da rede serão substituídas. Esta substituição consiste em manter um dos nodos (um dos extremos da ligação) e modificar o outro nodo.
- \* A probabilidade de substituir cada ligação será representada por  $p$ .

32

32



## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* A substituição das ligações irá reduzir o número de triângulos, mas reduzirá algumas distâncias e a distância média.
- \* Quais os valores de  $p$  que permitem reduzir a distância média sem diminuir drasticamente o coeficiente de *clustering*?

33

33

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* A probabilidade deve ser pequena para manter muitos triângulos.
- \* Por exemplo, alguns testes sugerem que valores entre 0,01 e 0,1 permitem reduzir a distância média e manter um número significativo de triângulos.
- \* Ver NetLogo Small Worlds<sub>34</sub>

34

17

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* Outra possibilidade para a substituição de cada ligação consiste em escolher aleatoriamente os dois nodos unidos pela ligação, não mantendo nenhum dos nodos iniciais.
- \* Em vez de substituir ligações, podem ser adicionadas ligações escolhidas ao acaso.

35

35

## Redes Aleatórias

- \* **Mundos Pequenos (*Small Worlds*)**
- \* Também podem ser consideradas outras configurações iniciais. Por exemplo, uma grelha em que os nodos internos têm grau igual a 6 e os nodos na fronteira têm grau igual a 2, 3 ou 4.
- \* O modelo Watts-Strogatz não gera redes com *hubs*.

36

36

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Dada uma sequência, será possível gerar uma rede em que os graus dos nodos são os elementos da sequência?
- \* Para que exista uma rede nestas condições, a soma dos elementos da sequência terá que ser um número par.

37

37

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Uma solução simples é dada pelo Modelo de Configuração (*Configuration Model*).
- \* Suponha que se tem um conjunto de nodos e uma sequência que representa os graus dos nodos.

38

38

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Em cada nodo, desenha-se um número de linhas incidentes (*stubs*) igual ao grau do nodo.
- \* Estas linhas não representam ligações porque incidem apenas num nodo.

39

39

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* O modelo consiste em:
  - \* 1 - escolher aleatoriamente um par de linhas;
  - \* 2 - ligar as duas linhas escolhidas, obtendo-se uma ligação.
- \* Este procedimento é repetido até todas as linhas ficarem unidas.

40

40

20

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* O modelo pode originar soluções que apresentem algumas desvantagens: mais do que uma ligação entre o mesmo par de nodos e existência de lacetes (*loops*).

41

41

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Este modelo pode ser utilizado para verificar se uma propriedade de uma rede é consequência da distribuição de grau.
- \* Se a propriedade da rede em estudo também estiver presente em todas as redes geradas pelo modelo então a propriedade resulta da distribuição de grau.

42

42

21

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Para assegurar a validade desta conclusão, o modelo terá que gerar todas as redes com a distribuição de grau pretendida.
- \* Por exemplo, pretende-se averiguar se o coeficiente de *clustering* da rede resulta da distribuição de grau.

43

43

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Utilizando o Modelo de Configuração (*Configuration Model*) obtêm-se todas as redes com a distribuição de grau pretendida.
- \* Determina-se o coeficiente de *clustering* de cada uma das redes obtidas e verifica-se se é igual ao da rede inicial.

44

44

## Redes Aleatórias

- \* **Modelo de Configuração (*Configuration Model*)**
- \* Os Modelos Aleatórios Exponenciais (*Exponential Random Models*) permitem obter redes com outras propriedades ou com conjuntos de propriedades (por exemplo, redes com determinado coeficiente médio de clustering e determinada densidade).

45

45

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Todas as redes já consideradas são redes estáticas (*static networks*), porque o número de nodos não sofre alterações durante o processo de geração, apenas se acrescentam ou substituem ligações.

46

46

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Usualmente, as redes reais são dinâmicas (*dynamic networks*), no sentido em que nodos e ligações podem ser acrescentados ou removidos.
- \* A Web, a Wikipedia, o Facebook são alguns exemplos de redes dinâmicas. Apesar de algumas remoções, a dimensão destas redes tem crescido.

47

47

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Os modelos dinâmicos tipicamente incorporam alguma forma de crescimento da rede.
- \* Começa-se com uma configuração inicial, frequentemente considera-se uma clique de pequena dimensão (uma clique é uma rede ou subrede completa).
- \* Depois acrescentam-se os nodos um a um.

48

48



## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Cada novo nodo é unido aos nodos já existentes de acordo com uma regra, que caracteriza o modelo.
- \* As redes aleatórias e modelos já estudados caracterizam-se pela falta de *hubs*, porque a probabilidade de uma ligação ser escolhida não difere da probabilidade associada a qualquer outra ligação.

49

49

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O próximo modelo procura gerar redes com *hubs*.
- \* Para tal, os novos nodos serão ligados aos nodos existentes, privilegiando as ligações a alguns dos nodos.
- \* Este mecanismo é designado por ligação preferencial (*preferential attachment*).

50

50

25

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O mecanismo privilegia as ligações a nodos com maior grau.
- \* Este mecanismo baseia-se no princípio: quanto mais se tem, mais se recebe.
- \* O método de ligação preferencial (*preferential attachment*) mais conhecido para redes foi proposto por Barabási e Albert, sendo designado por modelo de Barabási-Albert ou modelo BA.

51

51

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Começa-se com uma rede completa com  $m_0$  nodos. Cada iteração consiste em dois passos:
  - \* 1 - Adiciona-se um novo nodo  $i$  à rede com  $m \leq m_0$  ligações. O parâmetro  $m$  representa o grau médio da rede inicial.
  - \* 2 - A probabilidade de unir o novo nodo a um nodo  $j$  já existente é dada por
 
$$\frac{\text{grau de } j}{\text{soma dos graus dos nodos}}.$$

52

52

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O número de iterações realizadas é escolhido de forma a obter o número de nodos pretendidos.
- \* A aplicação deste método começa com graus iguais.
- \* A introdução de novos nodos e de novas ligações leva ao aumento dos graus de alguns nodos.
- \* Desde as primeiras iterações, o grau dos primeiros nodos será maior do que o dos nodos introduzidos posteriormente.

53

53

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* Os nodos com maior grau têm maior probabilidade de serem escolhidos, logo o seu grau tende a aumentar bastante mais do que o dos restantes nodos.
- \* Assim, surgem a heterogeneidade e os *hubs*.

54

54

27

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* No modelo Ligação Preferencial (*preferential attachment*), a escolha dos nodos adjacentes de cada novo nodo, com base no grau dos nodos já existentes, é relevante para a obtenção de *hubs*.
- \* No caso de a escolha dos nodos adjacentes ser aleatória, sem ter em conta o grau, observar-se-ia um crescimento da rede sem heterogeneidade.

55

55

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O método Ligação Preferencial (*preferential attachment*) foi utilizado para modelar a população das cidades, a concentração de riqueza individual, a dimensão de empresas, a produção científica e outros fenómenos.
- \* Ver NetLogo Preferential Attachment.

56

56

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O modelo Ligação Preferencial apresenta um conjunto de desvantagens:
  - \* o padrão da distribuição de grau não se altera qualquer que seja a escolha dos parâmetros;
  - \* os *hubs* pertencem ao conjunto dos primeiros nodos; nenhum dos nodos acrescentados a partir de determinada iteração será um *hub*;
  - \* não forma muitos triângulos;

57

57

## Redes Aleatórias

- \* **Ligação Preferencial (*Preferential Attachment*)**
- \* O modelo Ligação Preferencial apresenta um conjunto de desvantagens:
  - \* só se adicionam nodos e ligações e não há remoções de nodos nem de ligações;
  - \* forma uma rede conexa.

58

58

# Redes Aleatórias

## \* Outros Modelos

- \* Uma extensão do modelo anterior utiliza uma potência do grau, deixando a preferência de ser linear. Assim, a probabilidade anterior pode ser substituída por

$$\frac{k_j^\alpha}{\sum_l k_l^\alpha}$$

59

59

# Redes Aleatórias

## \* Outros Modelos

- \* Se o valor de  $\alpha$  for inferior a 1 então as probabilidades não crescem na mesma proporção que os graus e os *hubs* tendem a desaparecer.
- \* Se, pelo contrário, o valor de  $\alpha$  for superior a 1, então os nodos com maior grau acumulam mais ligações muito mais rapidamente. Assim, m nodos serão adjacentes da maioria dos nodos.

60

60

30

# Redes Aleatórias

## \* Outros Modelos

- \* Pode concluir-se que o modelo Ligação Preferencial está dependente da preferência linear (o expoente dos graus é igual a 1) para gerar diversos *hubs*. Esta é outra desvantagem do modelo.

61

61

# Redes Aleatórias

## \* Outros Modelos

### \* *Attractiveness Model*

- \* Outra variante do modelo Ligação Preferencial modifica a probabilidade adicionando uma constante  $A$  aos graus dos nodos. A probabilidade será dada por

$$\frac{A + k_j}{\sum_l (A + k_l)}$$

62

62

## Redes Aleatórias

- \* **Outros Modelos**
- \* ***Attractiveness Model***
- \* O parâmetro  $A$  é positivo e é designado por parâmetro de atratividade.
- \* A ideia consiste em escolher nodos com base, não apenas no grau, mas também na sua atratividade (por exemplo, citações de obras).

63

63

## Redes Aleatórias

- \* **Outros Modelos**
- \* ***Attractiveness Model***
- \* Ao contrário do modelo Ligação Preferencial, este modelo pode ser aplicado a configurações iniciais com nodos com grau nulo, redes orientadas e permite obter diversas distribuições de grau.
- \* São vantagens face ao modelo Ligação Preferencial.

64

64

32



# Redes Aleatórias

- \* **Outros Modelos**

- \* ***Fitness Model***

- \* O "Fitness Model", proposto por Bianconi e Barabási, considera uma modificação no cálculo das probabilidades.
- \* Considere-se uma função que toma valores para cada um dos nodos.

65

65

# Redes Aleatórias

- \* **Outros Modelos**

- \* ***Fitness Model***

- \* O grau de cada nodo é multiplicado pelo valor da função para o nodo. O valor da função traduz o apelo do nodo.
- \* A probabilidade será dada por:

$$\frac{n_j k_j}{\sum_l (n_l k_l)}$$

66

66

33

# Redes Aleatórias

- \* **Outros Modelos**

- \* ***Fitness Model***

- \* Quanto maior for o valor, maior será o apelo. A função é designada por função de *fitness*.
- \* Este modelo permite obter diversos *hubs*, se os valores da função forem limitados.
- \* Os valores da função podem permitir a competição entre os nodos iniciais e os novos nodos.

67

67

# Redes Aleatórias

- \* **Outros Modelos**

- \* ***Fitness Model***

- \* Como exemplos de aplicação pode referir-se a *Web*, a *Wikipedia* e publicações científicas. As páginas e os artigos com mais ligações, ou mais citados, não são necessariamente os mais antigos.

68

68

## Redes Aleatórias

- \* **Outros Modelos**
- \* ***Fitness Model***
- \* Os valores da função não sofrem alterações ao longo do tempo, o que pode constituir uma desvantagem do método.

69

69

## Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (*Random Walk Model*)**
- \* As redes geradas pelo modelo BA apresentam coeficientes de *clustering* reduzidos, porque a probabilidade de um nodo receber uma ligação é proporcional ao seu grau e não tem em conta a existência de nodos adjacentes.

70

70

# Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Para aumentar o número de triângulos, é necessário um mecanismo que favoreça a introdução de ligações entre nodos adjacentes.

71

71

# Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* A formação de triângulos resultante da adição de uma ligação é designada por **fecho triádico** (*triadic closure*) e é um dos mecanismos mais relevante para a formação de ligações numa rede social.

72

72

## Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Vamos considerar a implementação mais intuitiva deste mecanismo, que é designada por **Modelo Passeio Aleatório** (*random walk model*).

73

73

## Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* A ideia consiste em unir um novo nodo não só a um já existente mas também a um ou mais nodos adjacentes deste último.

74

74

## Redes Aleatórias

- \* **Outros Modelos**
- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Considere-se uma rede qualquer de dimensão pequena. Cada iteração consiste nos passos:
  - \* 1 - Um novo nodo  $i$  é acrescentado à rede com  $m > 1$  ligações;

75

75

## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* 2 - A primeira ligação será  $(i, j)$  em que  $j$  é um nodo já existente, escolhido aleatoriamente;
- \* 3 - Cada uma das restantes ligações une  $i$  a um dos adjacentes de  $j$  com probabilidade  $p$  ou une  $i$  a um nodo escolhido aleatoriamente com probabilidade  $1 - p$ .

76

76

## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* O número de triângulos formados depende da probabilidade  $p$  considerada.
- \* Se esta probabilidade não for demasiado pequena, este modelo também vai gerar *hubs*.
- \* A probabilidade de um nodo receber ligações será proporcional ao seu grau, como no modelo BA.

77

77

## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Contudo, os novos nodos não escolhem os seus adjacentes com base no grau. A escolha do primeiro adjacente de cada novo nodo é aleatória.
- \* O fecho triádico induz a ligação preferencial.

78

78

## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Este modelo foi proposto em 2003.
- \* Contudo, em 1973, Mark S. Granovetter publicou o artigo "*The strength of weak ties*", em que estabeleceu uma relação estreita entre três componentes fundamentais das redes sociais: triângulos, peso das ligações e comunidades.

79

79

## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Introduziu o princípio do fecho triádico forte para explicar a formação das ligações nas redes sociais.

80

80



## Redes Aleatórias

- \* **Modelo Passeio Aleatório (Random Walk Model)**
- \* Neste artigo, o sociólogo argumentou que ligações fortes estão presentes em comunidades sociais, enquanto que as ligações fracas unem comunidades sociais.
- \* Estas ligações fracas permitem a circulação de "informação" na rede.

81

81

## Redes Aleatórias

- \* **Modelo Cópia (Copy Model)**
- \* O modelo Cópia consiste numa variante do modelo Passeio Aleatório, em que cada nodo é ligado a um nodo escolhido aleatoriamente ou a algum (ou a alguns) dos seus adjacentes.

82

82

41

## Redes Aleatórias

- \* **Modelo Cópia (Copy Model)**
- \* O modelo Cópia procura modelar cenários em que muitos triângulos podem resultar de cópias de contactos.

83

83

## Redes Aleatórias

- \* **Modelo Cópia (Copy Model)**
- \* Como exemplos, podem referir-se:
  - \* duplicação de genes (as cópias vão interagir com as mesmas proteínas);
  - \* conhecimento de novas obras a partir de listas bibliográficas ou lista de obras da mesma colecção;

84

84

## Redes Aleatórias

- \* **Modelo Cópia (Copy Model)**
- \* Como exemplos, podem referir-se:
  - \* criação de conteúdos na Web e cópia das hiperligações.
- \* A escolha entre um nodo e alguns dos seus adjacentes permite obter hubs mas o número de triângulos é reduzido.

85

85

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* No modelo Ligação Preferencial, frequentemente designado por BA, a escolha das ligações dos novos nodos é baseada no grau dos restantes nodos.
- \* Este modelo requer o conhecimento dos valores absolutos dos graus.

86

86

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* O modelo Ordenação permite considerar outras propriedades, além do grau.
- \* Baseia-se numa ordenação dos nodos considerando a propriedade escolhida.

87

87

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* O modelo pode começar com uma qualquer rede de pequena dimensão com  $m_0$  nodos.
- \* Uma propriedade dos nodos (grau, idade, ...) é escolhida para ordenar os nodos.

88

88

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Cada iteração consiste nos passos:
  - \* 1 - Todos os nodos são ordenados com base na propriedade. Atribuem-se os valores  $R = 1, 2, 3, \dots$  aos nodos. O nodo  $l$  da lista ordenada recebe  $R = l$ .

89

89

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Cada iteração consiste nos passos:
  - \* 2 - Adiciona-se um novo nodo  $i$  à rede com  $m \leq m_0$  novas ligações.

90

90

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Cada iteração consiste nos passos:
  - \* 3 - A probabilidade de unir o novo nodo a um nodo  $j$  já existente é dada por

$$\frac{R_j^{-\alpha}}{\sum_l R_l^{-\alpha}}$$

em que  $\alpha > 0$  é um parâmetro.

91

91

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Os nodos poderão ter que ser reordenados após cada iteração, se a propriedade escolhida depender das novas ligações adicionadas.
- \* Por exemplo, a escolha do grau obriga a reordenar os nodos.

92

92

## Redes Aleatórias

### \* **Modelo Ordenação (Rank Model)**

- \* Um nodo nas primeiras posições da lista ordenada terá maior probabilidade de receber uma nova ligação do que um dos nodos nas últimas posições da lista ordenada.
- \* Se os nodos tiverem sido ordenados pelo seu grau, os nodos com maior grau terão probabilidades associadas mais elevadas do que os com menor grau.

93

93

## Redes Aleatórias

### \* **Modelo Ordenação (Rank Model)**

- \* No entanto, a probabilidade depende da posição na ordenação e não será proporcional aos graus.
- \* A taxa de decréscimo da probabilidade depende do expoente considerado.
- \* Este modelo gera redes com distribuição de grau com caudas longas, qualquer que seja a propriedade escolhida e o expoente considerado.

94

94

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Com escolhas diferentes do expoente obtêm-se distribuições de grau com formas diferentes, o que permite reproduzir distribuições empíricas.
- \* Este modelo permite criar *hubs*, apesar de não incorporar informação detalhada sobre o sistema.

95

95

## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* Um exemplo de aplicação deste modelo é relativo à ligação de novos artigos da *Wikipedia*.
- \* Os autores de artigos procuram ligá-los a artigos relevantes.

96

96



## Redes Aleatórias

- \* **Modelo Ordenação (Rank Model)**
- \* A utilização de motores de busca permite obter uma lista de artigos ordenada por relevância.
- \* Usualmente escolhem-se os artigos da primeira página da lista ordenada.
- \* Esta escolha leva à criação de *hubs*.

97

97

## Redes Aleatórias

- \* **Algumas Aplicações de Redes Aleatórias**
- \* As redes aleatórias podem ser utilizadas para modelar e estudar alguns sistemas complexos.
- \* No caso das redes sociais, alguns estudos permitiram concluir que determinados comportamentos humanos são adoptados mais fácil e rapidamente em redes com *clusters* (ou grupos) do que em redes aleatórias.

98

98

49

## Redes Aleatórias

### \* Algumas Aplicações de Redes Aleatórias

- \* No caso da propagação de doenças, é relevante identificar as subredes em que o contágio ocorrerá mais cedo para desenvolver estratégias de prevenção e tratamento.
- \* A estrutura da rede em estudo poderá ter influência nos resultados, pelo que deve representar bem a população em estudo.

99

99

## Redes Aleatórias

### \* Algumas Aplicações de Redes Aleatórias

- \* Outras aplicações consistem no estudo:
  - \* da robustez de redes de abastecimento (eletricidade, gás, ...);
  - \* na propagação de vírus em redes;
  - \* na identificação de pontos sensíveis e fracos numa rede.

100

100

50

## Redes Aleatórias

### \* Algumas Aplicações de Redes Aleatórias

- \* No campo da neurociência, o cérebro pode ser representado por uma rede de neurónios.
- \* Há evidência de que as alterações provocadas pelo Alzheimer condicionam a actividade na rede de neurónios, estando mais próxima da actividade numa rede aleatória do que numa rede representativa de um cérebro saudável.

101

101

## Redes Aleatórias

### \* Algumas Aplicações de Redes Aleatórias

- \* Quais são as vantagens em utilizar redes aleatórias para modelar e estudar sistemas complexos?

102

102