Instituto de Ciências Matemáticas e de Computação

| Universidade de São Paulo |

Introdução às Redes Complexas

Lucas Antiqueira

Disciplina: SCC216 - Modelagem Computacional em Grafos

Docente: Profa. Dra. Rosane Minghim

21/05/2013

Roteiro da Aula

1. Contexto geral

2. Um pouco de história

3. O "*boom*" das redes complexas
   - Inserção dos modelos WS e BA

4. Outros conceitos
   - Algumas medidas de redes complexas

5. Exemplos de aplicações
   - Textos e biologia

Contexto geral

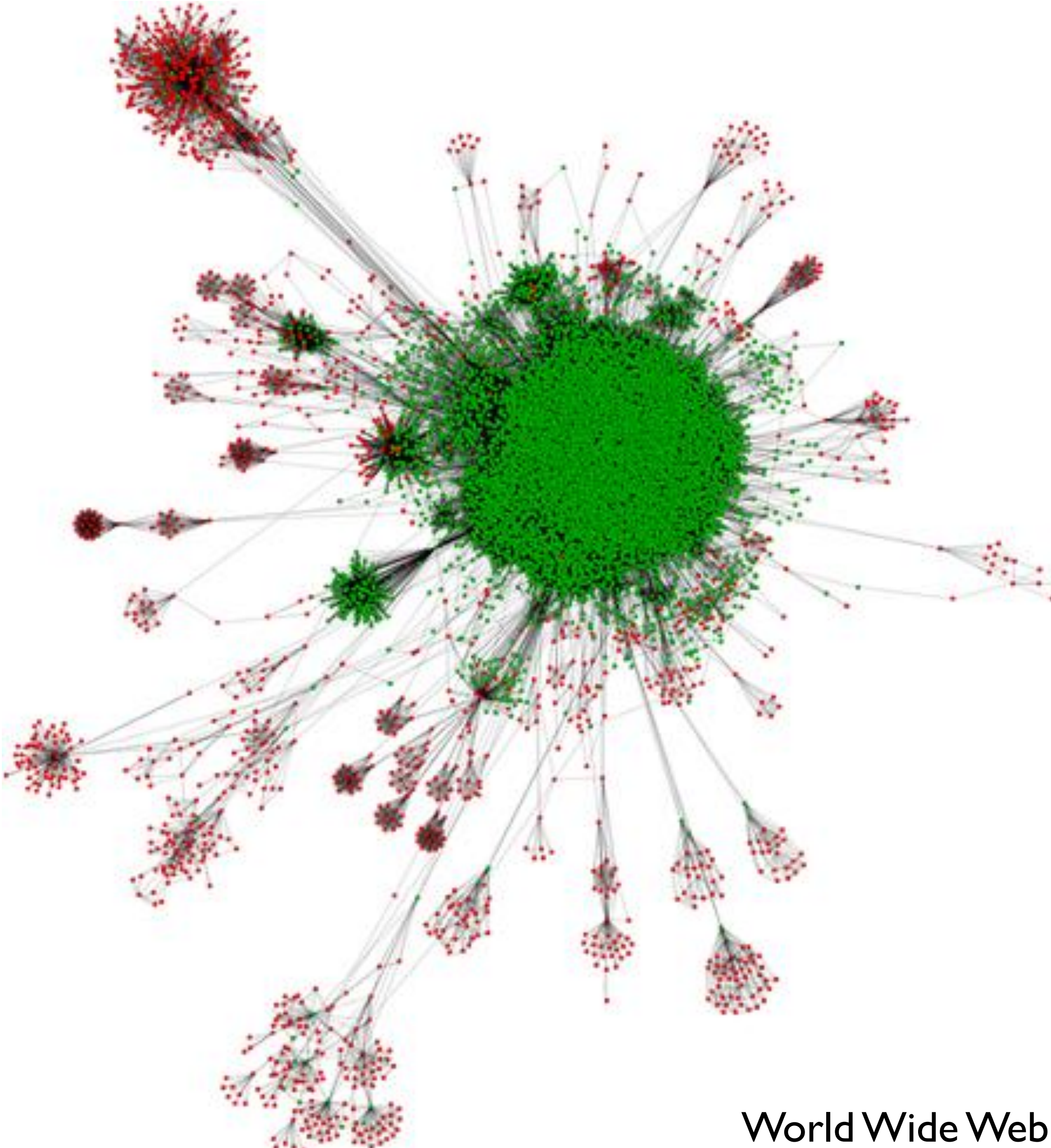
- Redes Complexas:
  Área direcionada ao entendimento e à previsão da estrutura e do comportamento de sistemas complexos modelados como grafos

- Altamente multidisciplinar
  - Computação, biologia, sociologia, física, etc

- Fundamentos:
  - Teoria dos grafos (Matemática)
  - Algoritmos e estruturas de dados (Computação)
  - Mecânica estatística (Física)

Contexto geral

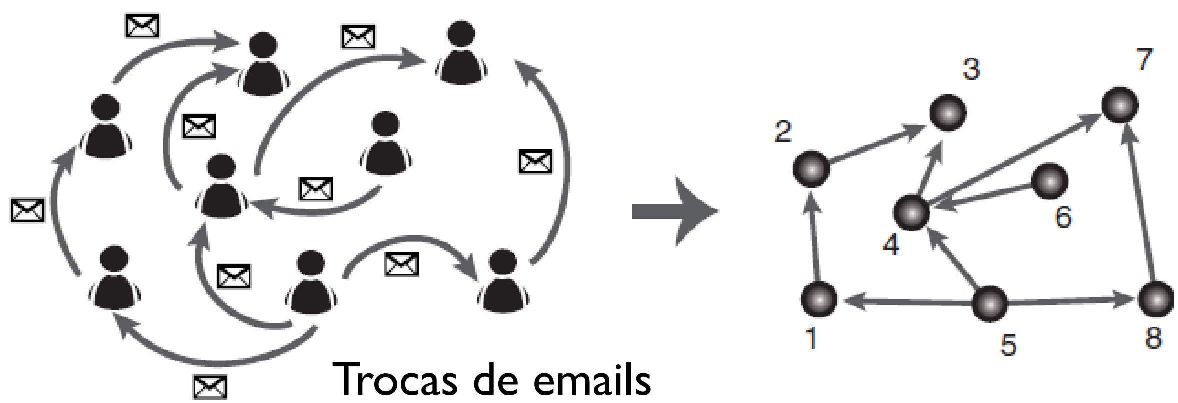
- Exemplos de redes complexas?



World Wide Web

Contexto geral

- Outros exemplos:
  - Internet
  - Sistemas de transporte (rodoviário, aéreo)
  - Infraestrutura elétrica
  - Cérebro (neurônios e sinapses)



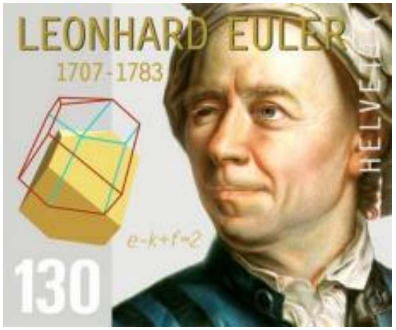
Trocas de emails

A maior parte das redes caem nas seguintes classes, de acordo com o método de construção (mapeamento em vértices e arestas):

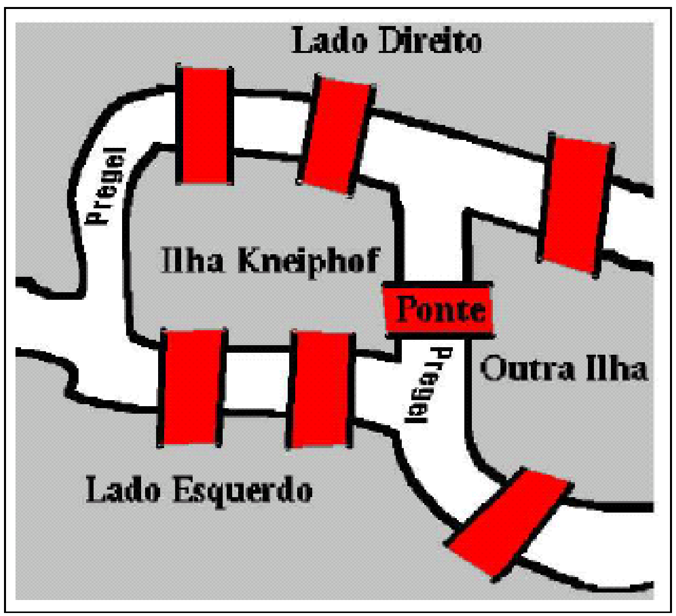
Comunicação:	e-mails, telefone, aeroportos, internet
Coexistência:	coautoria, música, filmes
Referência:	web, citações, software
Confluência:	cidades, estradas, circuitos
Correlação:	mercado financeiro, neurociência
Adjacência: *(temporal ou espacial)*	terremotos, textos

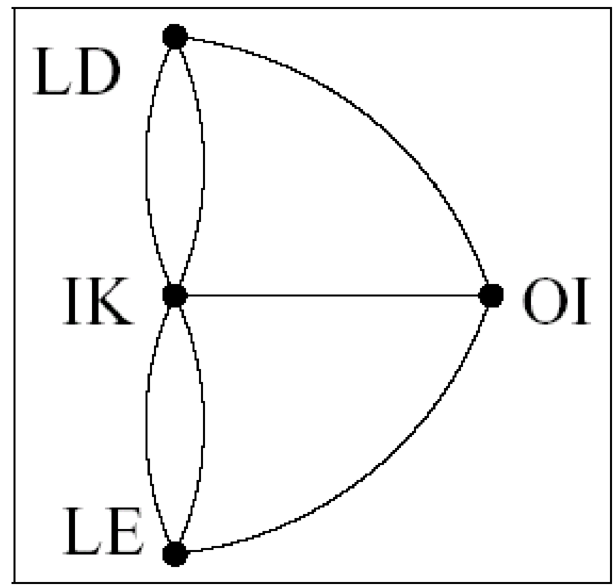
UM POUCO DE HISTÓRIA

História



Grafo:
Conceito formalmente introduzido por Leonhard Euler em 1736





Problema das Pontes de Königsberg
É possível planejar um caminho de modo que se cruze cada uma das pontes uma única vez?

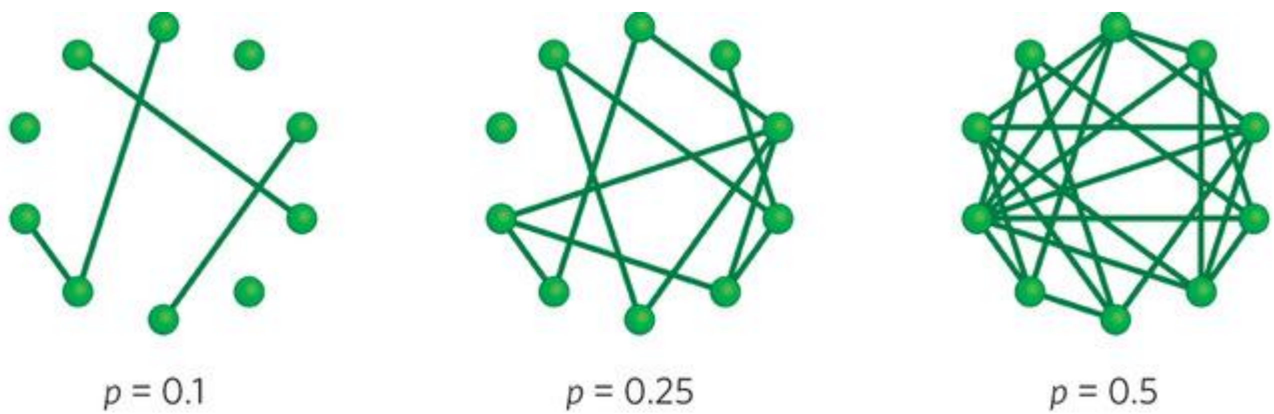
História

- Paul Erdös (1913-1996)
  - Matemático húngaro
  - Um dos grandes cientistas do século XX
  - Publicou mais de 1.500 artigos científicos

- Final da década de 1950 em diante: Criou e desenvolveu o modelo dos grafos aleatórios (em notável colaboração com Alfréd Rényi)

Modelo aleatório

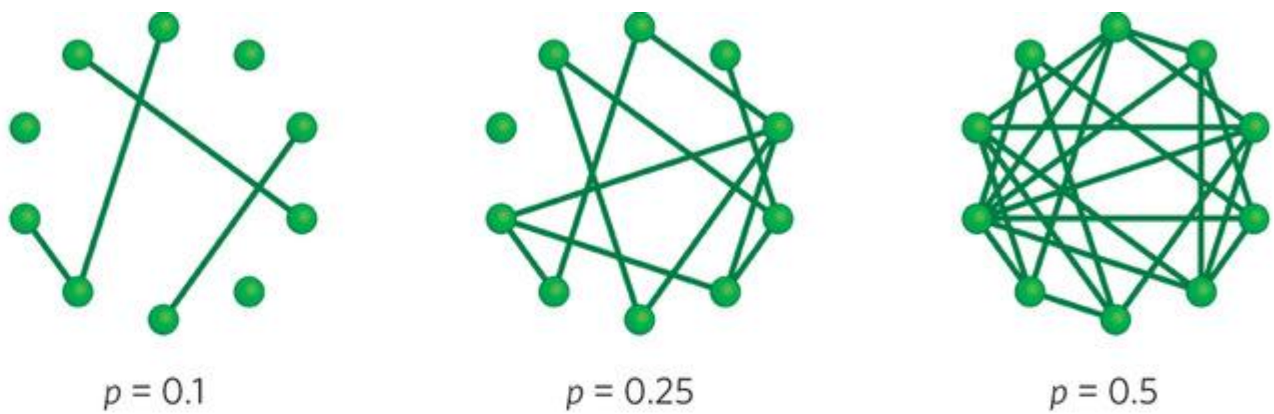
- Grafo aleatório $G(N, p)$ de Erdös e Rényi:
  Um grafo com N vértices é definido por meio de uma criação aleatória de arestas. Cada aresta é criada com probabilidade p.



Nature Physics **6**, 539–543 (2010)

Modelo aleatório

- A partir desse modelo foram desenvolvidas soluções exatas para diversas propriedades do grafo aleatório, tais como sua distribuição de graus e tamanhos de componentes conexos.



Nature Physics **6**, 539–543 (2010)

Sugestão de exercício

- Implementar em seu TAD Grafo uma função que cria um grafo aleatório de acordo com o modelo ER (Erdös-Rényi).
  - Como criar uma aresta com probabilidade p?

Algoritmo (modelo ER)

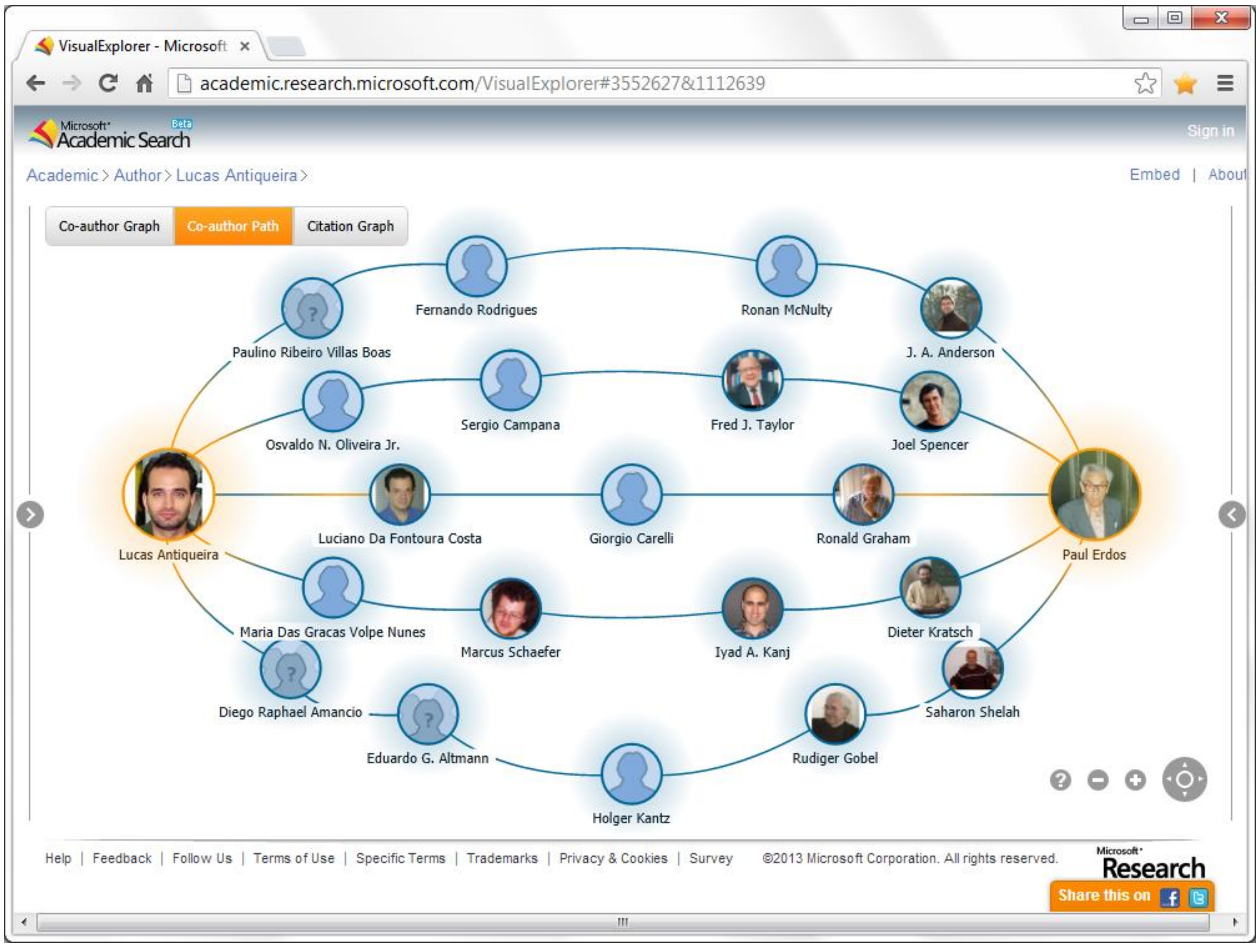
```
função random_graph(N, p) → retorna grafo G
  G ← Crie um grafo de N vértices e nenhuma aresta
  para i de 1 até N faça
    para j de i+1 até N faça
      s ← sorteie um número de 0 a 1
      se s <= p então  //cria aresta com probabilidade p
        Crie em G uma aresta não dirigida (i,j)
      fim_se
    fim_para
  fim_para
  retorne G
fim_função
```

Observação: Neste algoritmo, índices de vértices variam de 1 a N (e não de 0 a N-1)

Curiosidade: número de Erdös

- A importância de Erdös é tão grande que outros cientistas costumam calcular o chamado número de Erdös:
  - Constrói-se um grafo onde cada vértice representa um cientista
    - Cria-se uma aresta entre dois cientistas que tenham publicado ao menos um artigo científico juntos
    - É uma rede de *co-autoria* ou *colaboração* científica
  - O número de Erdös de um cientista i é dado pela distância entre i e Erdös na rede de co-autoria
    - Comprimento de caminho mínimo!

Exemplo:



Qual o comprimento do menor caminho que leva Lucas Antiqueira a Paul Erdös?

Número de Erdös é igual a 4

História

- Stanley Milgram, psicólogo (1933-1984)

- Década de 1960:
  - The Small-World Experiment
  - Milgram estudou a rede social dos EUA

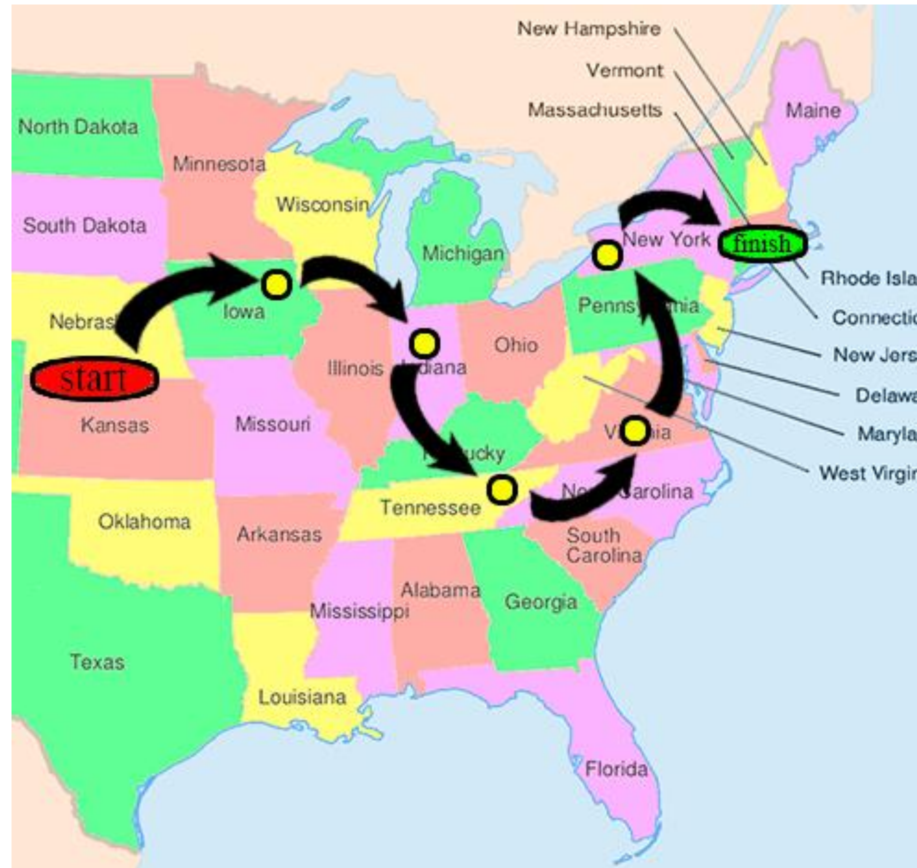
- Como, se não havia Facebook?!?

- Solução: envio de cartas!

- Pessoas foram escolhidas aleatoriamente no Nebraska e no Kansas

- Objetivo: essas pessoas deveriam enviar uma carta a uma determinada pessoa X em Boston

- Caso X não fosse conhecida, a carta deveria ser enviada a um amigo que "talvez" conhecesse X.

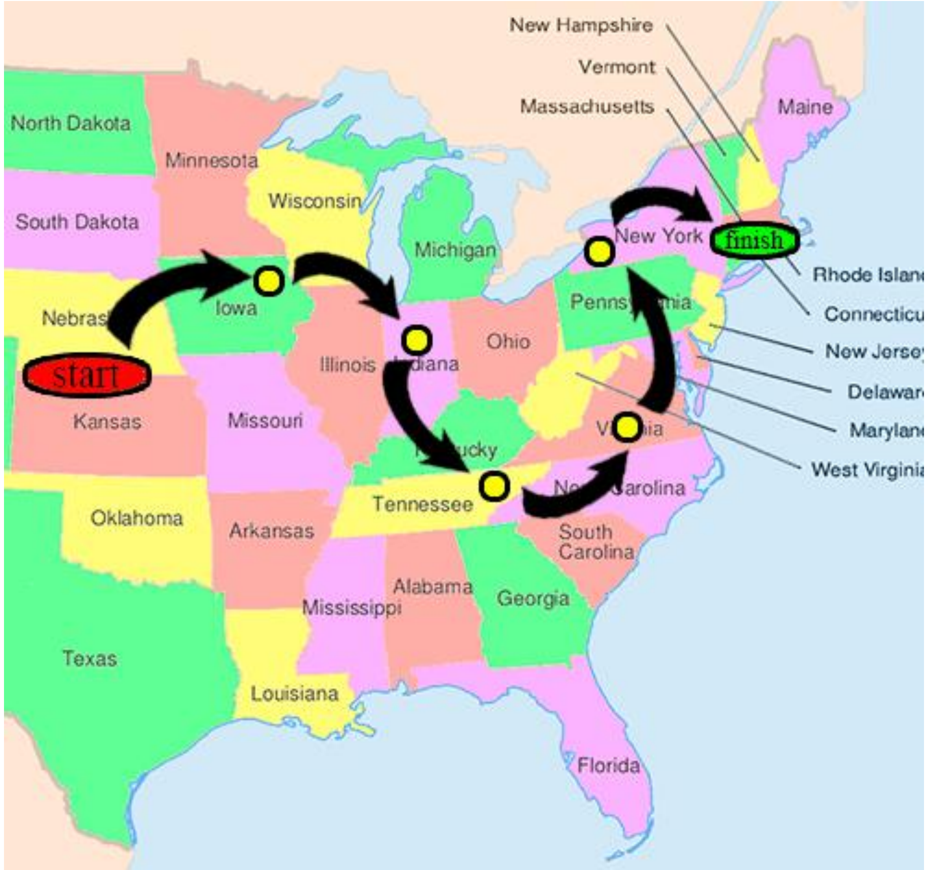


http://en.wikipedia.org/wiki/File:Experement_Small_World_(possible_option).gif

- Resultados

- Cartas que chegaram ao destino! (64 de 296)

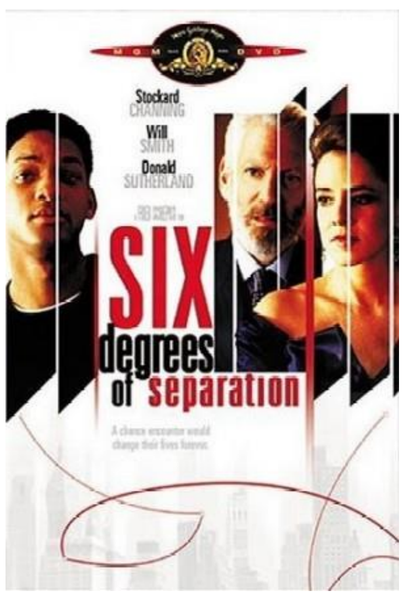
- Quantas vezes, em média, cada carta teve que ser re-endereçada até chegar ao destino?
- Resposta: **Seis**

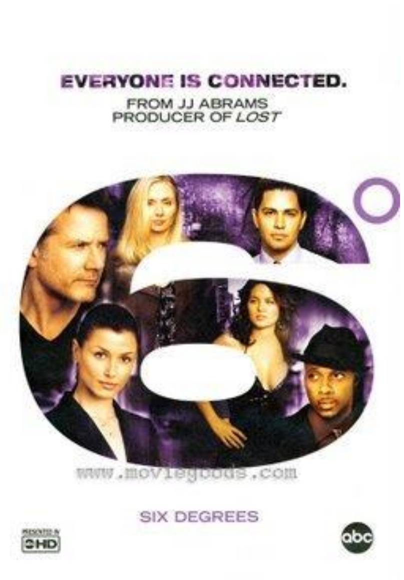


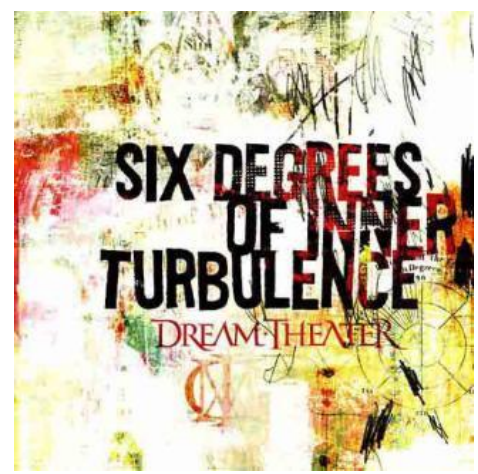
http://en.wikipedia.org/wiki/File:Experement_Small_World_(possible_option).gif

Curiosidade: *six degrees*

- O termo "seis graus" tornou-se parte da cultura popular americana:
  - Peça teatral e filme: *Six Degrees of Separation*
  - Série de TV: *Six Degrees*
  - Música: *Six Degrees of Inner Turbulence* (Dream Theater)







Contexto histórico

- Até o final da década de 1990:

  Acreditava-se que o grafo aleatório era um bom modelo para representar redes do mundo real

- *Por exemplo:*
  As curtas distâncias observadas no experimento de Milgram são reproduzidas no modelo de Erdös-Rényi

O "*BOOM*" DAS REDES COMPLEXAS

O "*boom*" das redes complexas

- Muitos pesquisadores já estudavam redes e suas representações por grafos
- *Boom*: O que motivou o surgimento do efervescente campo de pesquisa chamado Redes Complexas?

O "*boom*" das redes complexas

- *Resposta: Foi descoberto que as redes são mais complexas do que acreditava-se anteriormente*

1. Redes pequeno-mundo:
   1998: Artigo na Nature
   Watts, D. J. & Strogatz, S. H.
   Collective Dynamics of 'Small-World' Networks

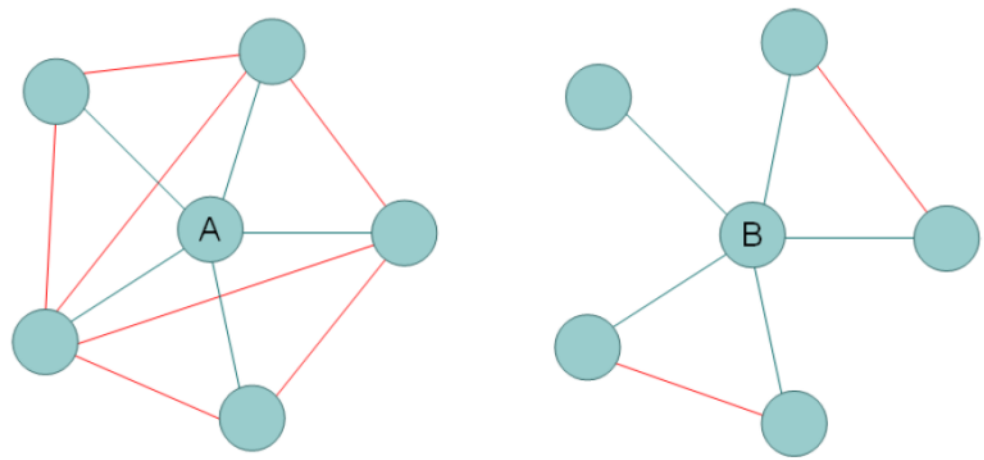
2. Redes invariantes à escala ("sem escala"):
   1999: Artigo na Science
   Barabási, A. L. & Albert, R.
   Emergence of Scaling in Random Networks

Redes pequeno-mundo

- 1998 – Duncan Watts e Steven Strogatz analisaram as seguintes redes:
  - Rede de colaboração entre atores (IMDB)
  - Rede elétrica dos EUA
  - Rede neural da *Caenorhabditis elegans* (verme)

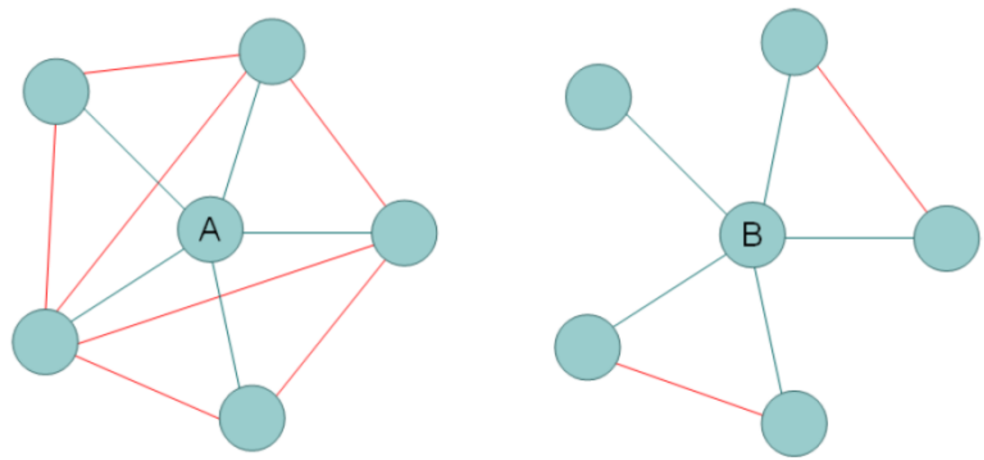

- Resultados:
  - Distâncias entre vértices são curtas
    - OK, já vimos isso em redes aleatórias
  - Agrupamento local é alto
    - Isso é novo!

Redes pequeno-mundo



Intuitivamente, qual vértice (A ou B) apresenta maior agrupamento em sua vizinhança?

Redes pequeno-mundo

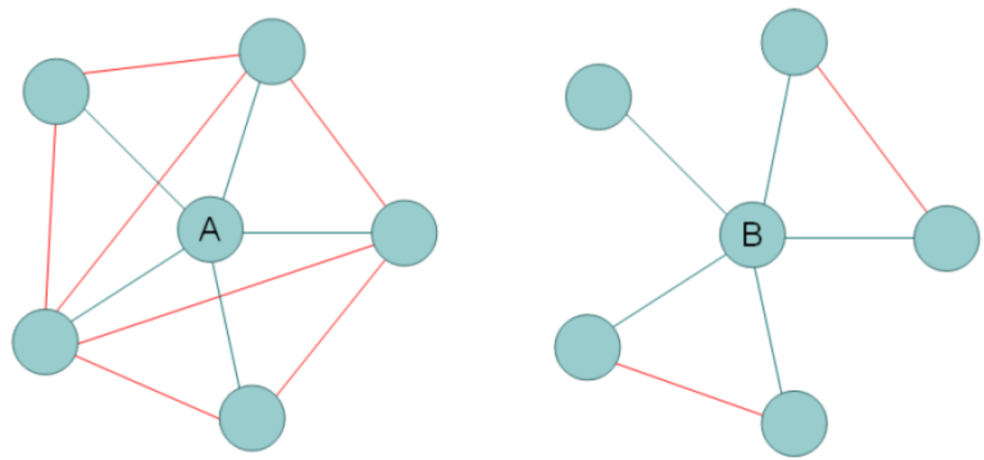


Número de arestas entre os vizinhos de A = 7
Número de arestas entre os vizinhos de B = 2

Número de vizinhos de A = 5
Número de vizinhos de B = 5

Número máximo possível de arestas entre os vizinhos de A ou B:
5 (5-1)/2 = 10

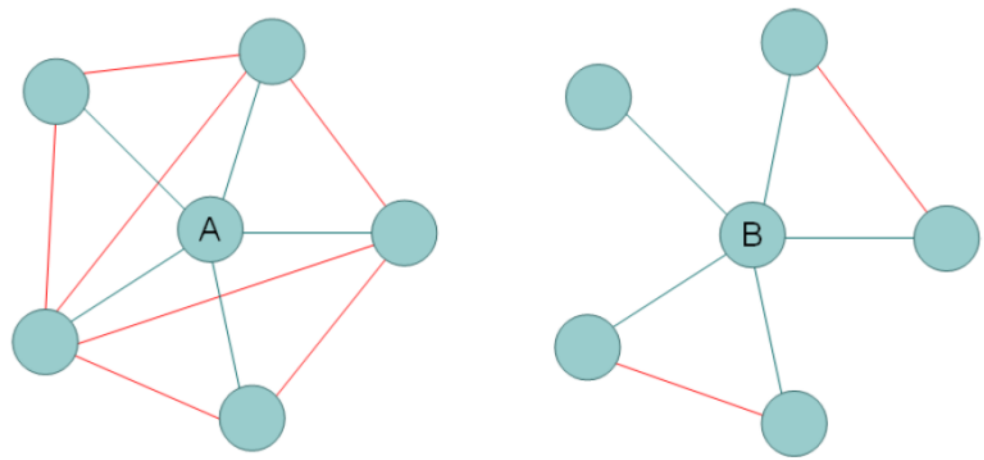
Redes pequeno-mundo



Coeficiente de agrupamento de X =

$$\frac{\text{Número de arestas entre os vizinhos de X}}{\text{Número máximo possível de arestas entre os vizinhos de X}}$$

Redes pequeno-mundo



Coeficiente de agrupamento de A = 7/10 = 0,7

Coeficiente de agrupamento de B = 2/10 = 0,2

Redes pequeno-mundo

- *As redes pequeno-mundo apresentam alto agrupamento local médio quando comparadas às redes aleatórias*

  - Coeficiente de agrupamento alto → redundância de conexões

  - Remete ao conceito de transitividade:
    Se A e B são amigos de C, é provável que A e B sejam amigos entre si

  - Característica comum de redes sociais, e de diversas outras redes, não contemplada anteriormente (p.ex. no modelo aleatório e no experimento de Milgram)

Sugestão de exercício

- Inclua em seu TAD grafo uma função que calcule o coeficiente de agrupamento de um dado vértice i. Faça também uma função que calcule a média de todos os coeficientes de agrupamento em um dado grafo.
  - O que fazer quando i tem menos de 2 vizinhos?

Sugestão de exercício

- É importante também ter implementada uma função que retorne o comprimento do caminho mínimo médio em um grafo.
  - Ou seja, considere a média das distâncias entre todos os diferentes pares de vértices.
  - O que fazer quando não há caminho entre 2 vértices?

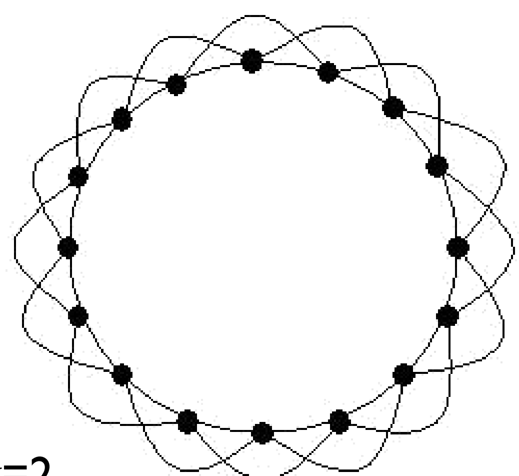
Modelo pequeno-mundo

- Por que as redes pequeno-mundo têm esse comportamento?
  - *Temos, no efeito pequeno-mundo, alto agrupamento local (quando comparado ao das redes aleatórias) e curtas distâncias (próximas do observado em redes aleatórias)*

- Watts e Strogatz propuseram um modelo de grafo intermediário entre os grafos aleatórios e os grafos regulares (uniformes)

Modelo pequeno-mundo

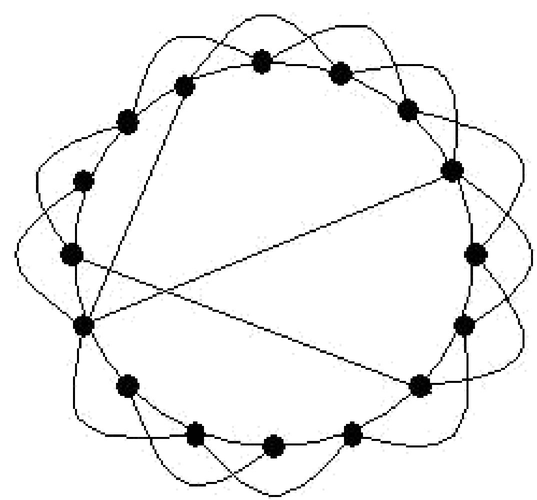
- Modelo WS (Watts-Strogatz):
  - Crie um grafo regular em forma de anel onde os k vizinhos mais próximos (de cada lado) são conectados



k=2

Modelo pequeno-mundo

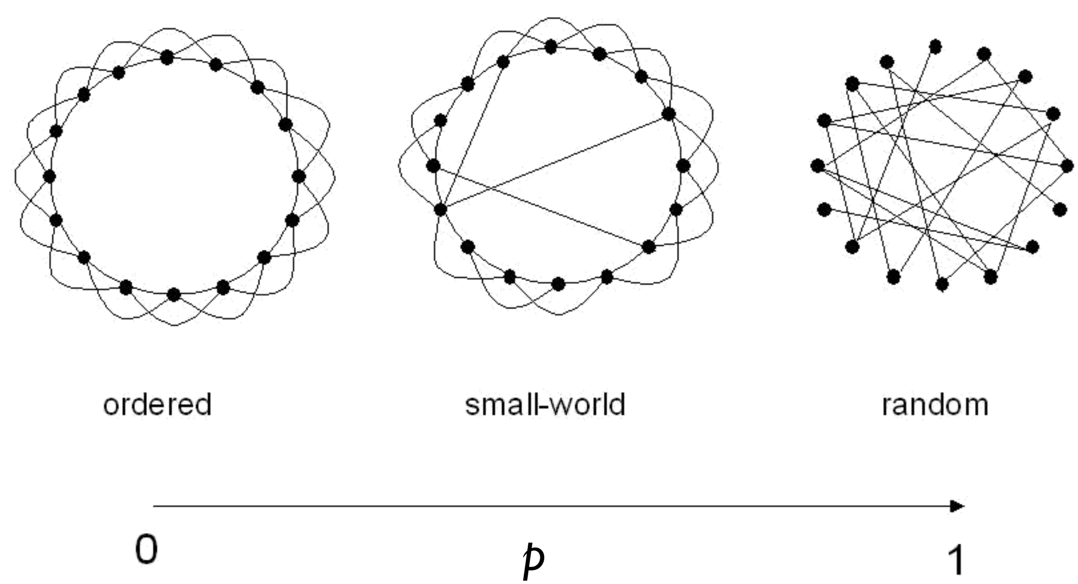
- Modelo WS (Watts-Strogatz):
  - Processo de *religação*:
    A seguir, cada aresta é movida aleatoriamente com probabilidade *p*.
    Mover uma aresta, neste caso, é alterar o vértice conectado a uma de suas pontas.



Para valores específicos de *p* o grafo apresenta o efeito pequeno-mundo.
Por exemplo, tome p=0,1

Modelo pequeno-mundo

- Modelo WS (Watts-Strogatz):
  - O parâmetro *p* controla a variação entre "ordem" e "aleatoriedade"
  - Com *p*=1 temos o grafo aleatório de Erdös e Rényi



http://home.kpn.nl/stam7883/graph_introduction.html

Sugestão de exercício

- Implemente em seu TAD Grafo uma função que cria um grafo de acordo com o modelo WS
- A seguir, crie e compare grafos ER e WS de acordo com:
  - Comprimento médio de caminho mínimo
  - Média do coeficiente de agrupamento
  - Responda: Para quais valores do parâmetro de religação do modelo WS o efeito pequeno-mundo aparece?
  - ***Obs.:*** *tome o cuidado de comparar grafos com mesmo número de vértices e aproximadamente mesmo número de arestas (ajuste os parâmetros dos modelos para tanto)*

Algoritmo (modelo WS)

```
função small_world_graph(N, k, p) → retorna grafo G
  G ← Crie um grafo de N vértices e nenhuma aresta
  //Agora cria grafo circular
  para i de 1 até N faça
    para j de 1 até k faça
      m = i + j
      se m > N então  //gera efeito circular
        m = m - N
      fim_se
      Crie em G uma aresta não dirigida (i,m)
      m = i - j
      se m < 1 então  //gera efeito circular
        m = N - m
      fim_se
      Crie em G uma aresta não dirigida (i,m)
    fim_para
  fim_para
  //Continua →
```

Observação: Neste algoritmo, índices de vértices variam de 1 a N (e não de 0 a N-1)

Algoritmo (modelo WS)

```
  //Agora faz a religação das arestas
  para todo i e j tal que a aresta (i,j) pertença ao grafo G faça
    s ← sorteie um número entre 0 e 1
    se s <= p então  //religa com probabilidade p
      Exclua a aresta (i,j) de G
      repita
        m ← sorteie um número entre 1 e N
      até_que m<>i e m<>j
      Crie em G uma aresta não dirigida (i,m)
    fim_se
  fim_para
  retorne G
fim_função
```

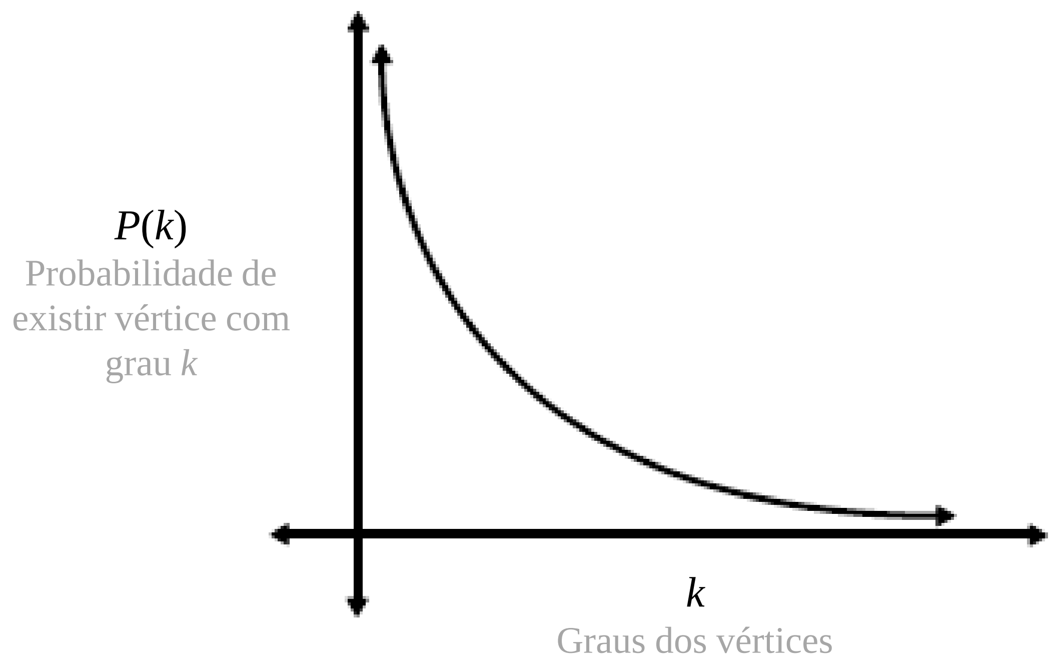
Observação: Neste algoritmo, índices de vértices variam de 1 a N (e não de 0 a N-1)

Redes "sem escala"

- 1999 – Albert-László Barabási e Réka Albert analisaram as seguintes redes:
  - Rede de colaboração entre atores (IMDB)
  - Mapa da World Wide Web
  - Rede elétrica dos EUA

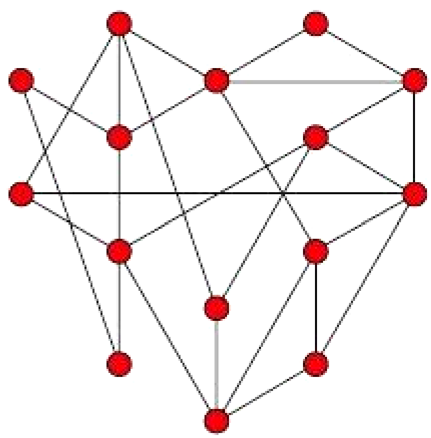
- Resultado:
  - Lei de potência na distribuição de graus (*scale-free*)
    - Isso é novo!
    - O que significa?

- O que indica a distribuição de graus de um grafo?

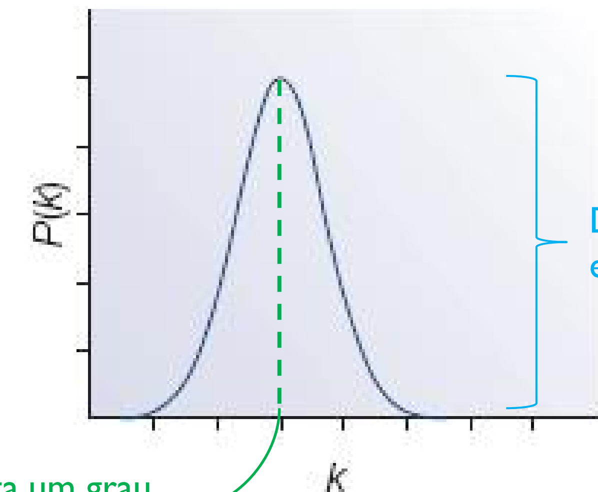


http://www.learner.org/courses/mathilluminated/units/11/textbook/05.php

- Distribuição de graus em redes aleatórias ER:



É uma distribuição de Poisson

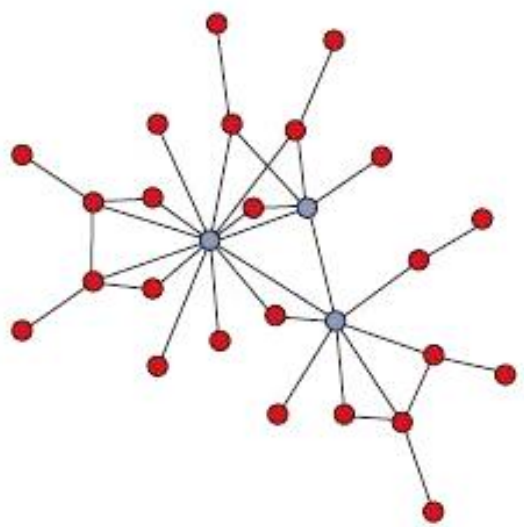


Redes pequeno-mundo WS apresentam outra distribuição, mas o comportamento é semelhante

http://cns-classes.bu.edu/cn710/Fall2007/index.php?n=Main.DiscussionPage-Class2Linked

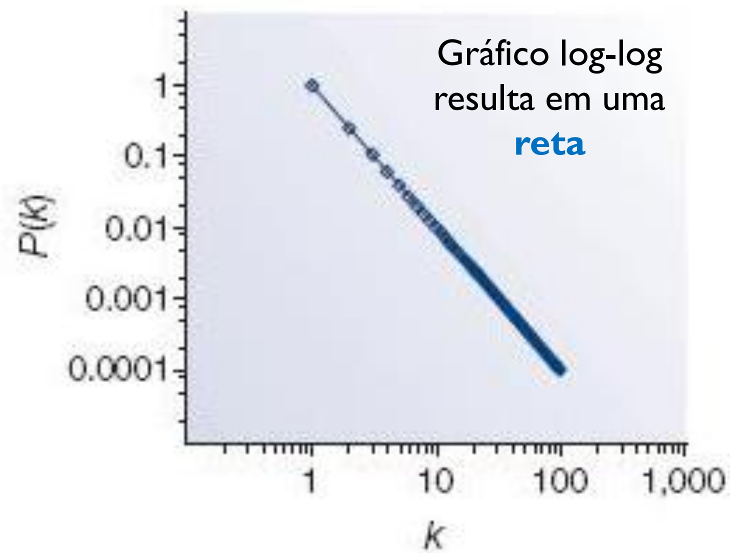
Redes “sem escala”

- Distribuição de graus em redes *scale-free*:



Segue uma lei de potência

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



http://cns-classes.bu.edu/cn710/Fall2007/index.php?n=Main.DiscussionPage-Class2Linked

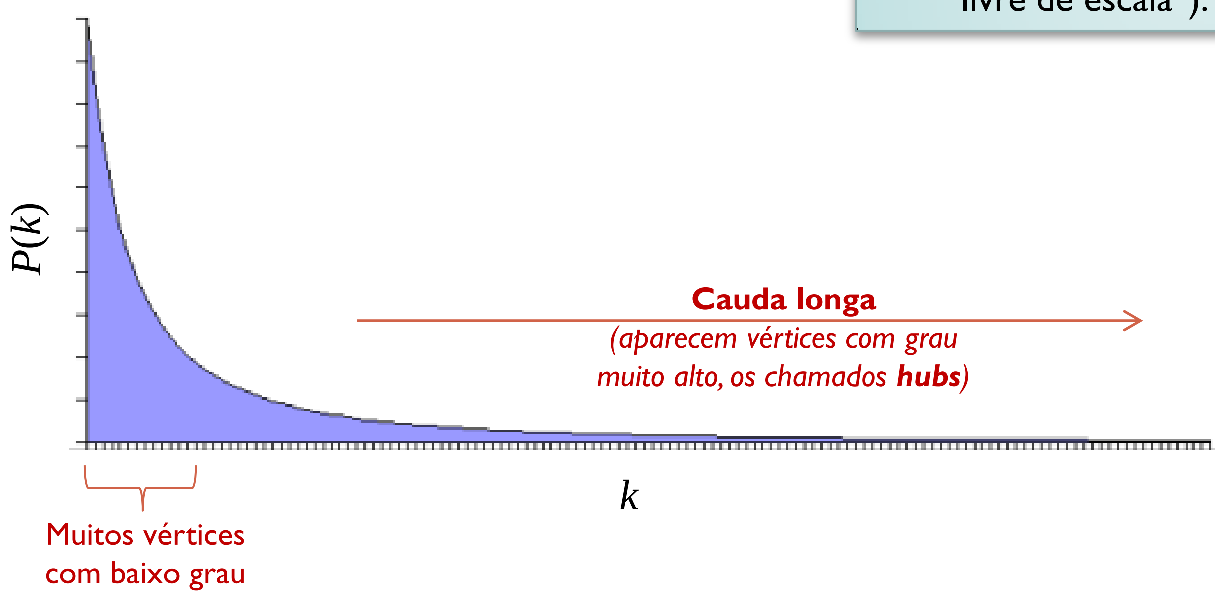
Redes "sem escala"

- Distribuição de graus em redes *scale-free*:

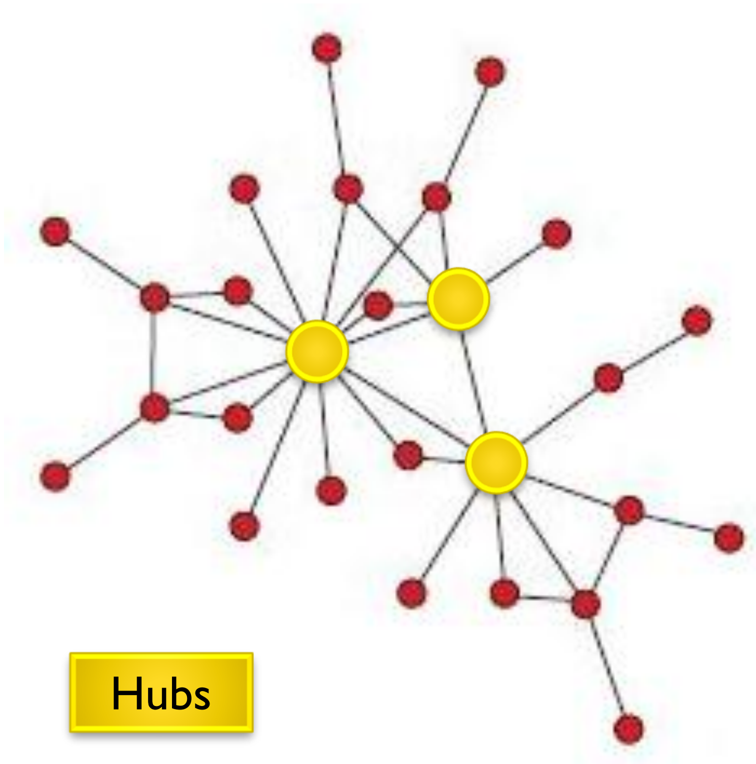
Lei de potência: $P(k) \sim k^{-\gamma}$
(onde γ depende da rede)

Distribuição invariante à escala.
Daí o nome *scale-free* ("sem escala" ou "livre de escala"). Vejam explicação!

Sem log-log:



Redes "sem escala"



Rede altamente tolerante a falhas

Ex.: Internet não deixa de funcionar totalmente caso alguns roteadores falhem

Rede altamente vulnerável a ataques

Ex.: Internet pode "parar" se os hubs deixarem de funcionar

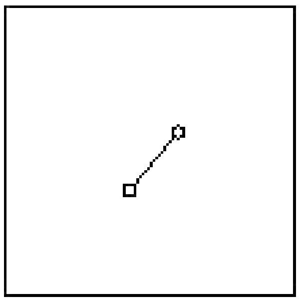
Albert, R.; Jeong, H. & Barabási, A.-L. Nature, 2000, 406, 378-382

Modelo *rich-get-richer*

- Mas como explicar o mecanismo por trás das redes sem escala?
- Barabási e Albert propuseram um modelo de grafo baseado em *crescimento* e *ligação preferencial*

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):

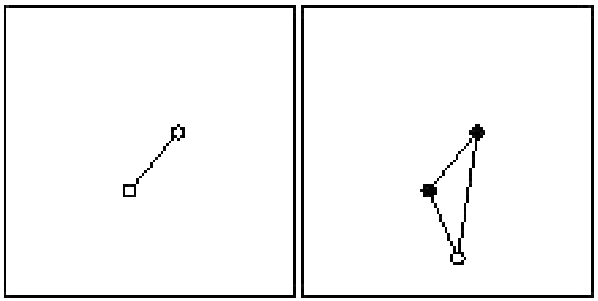


Crie um grafo com alguns vértices conectados

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):

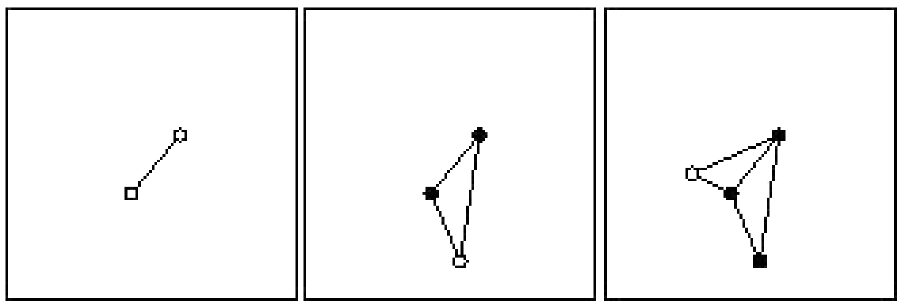


Adicione outro vértice (exibido em branco) e crie outras *m* arestas, conectando o novo vértice aos vértices previamente criados
Obs.: Nesse caso (m=2)

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):



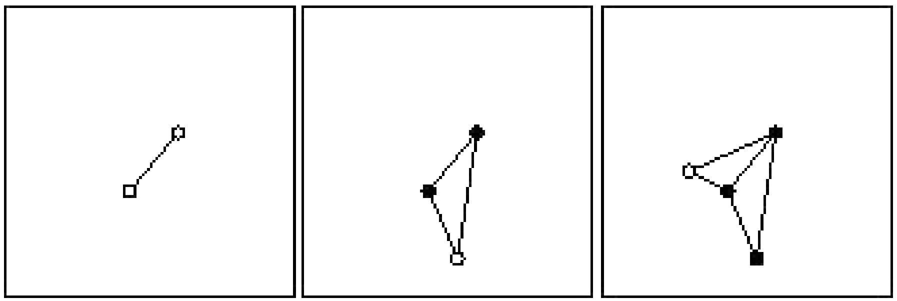
Adicione outro vértice e crie outras *m*=2 arestas.

Como escolher com quem conectar o novo vértice?

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):



Adicione outro vértice e crie outras m=2 arestas.

Como escolher com quem conectar o novo vértice?

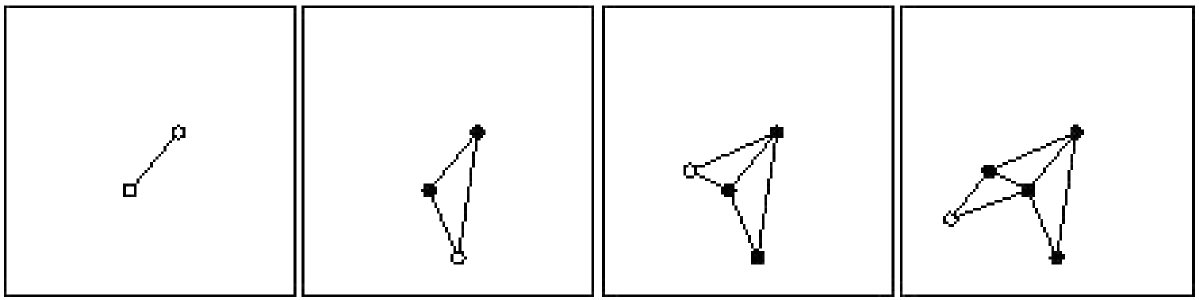
Ligação preferencial → Quem tem maior grau tem maior probabilidade de receber novas conexões

A probabilidade de um vértice receber conexões é linearmente proporcional ao seu grau

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):

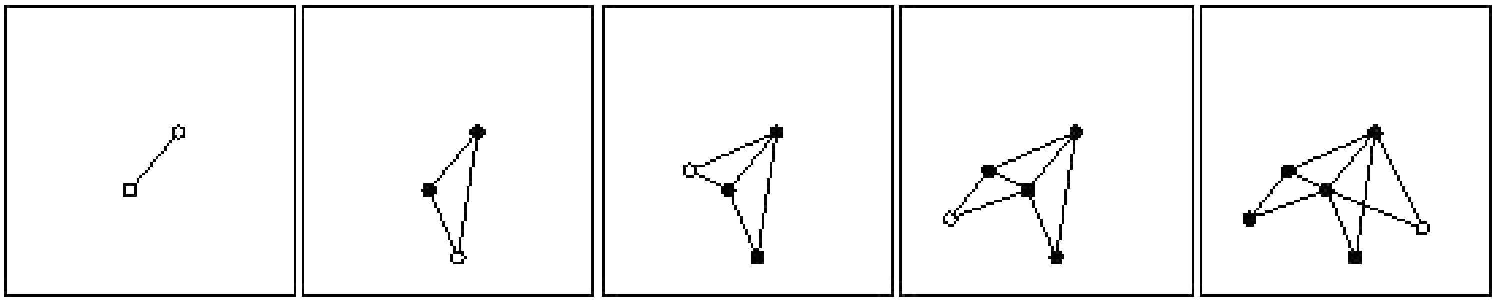


Adicione outro vértice e crie outras *m*=2 arestas por ligação preferencial

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):

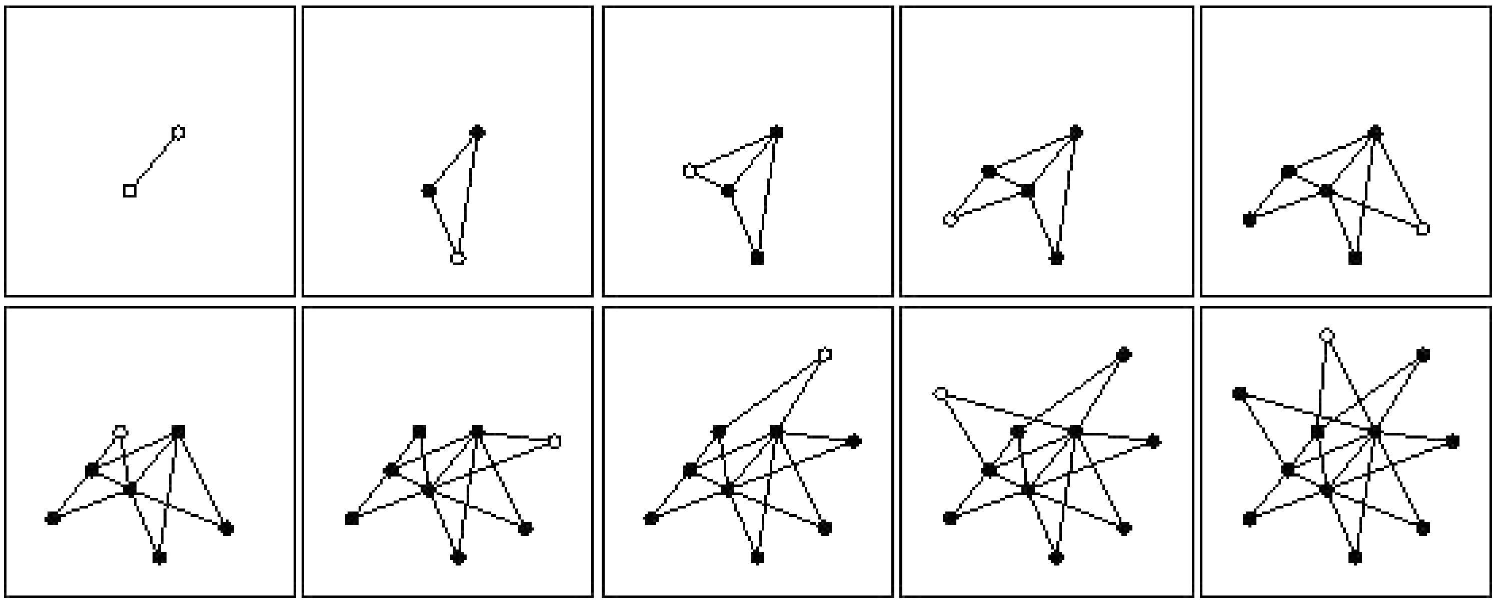


Adicione outro vértice e crie outras m=2 arestas por ligação preferencial

http://www3.nd.edu/~networks/Linked/newfile11.htm

Modelo *rich-get-richer*

- Modelo BA (Barabási-Albert):



E assim sucessivamente, até o grafo ter o número *N* de vértices desejado

http://www3.nd.edu/~networks/Linked/newfile11.htm

Sugestão de exercício

- Implemente em seu TAD Grafo uma função que cria um grafo de acordo com o modelo BA
  - Como vocês implementariam a ligação preferencial?
- A seguir, crie e compare grafos ER, WS e BA de acordo com suas distribuições de graus
  - Estime as distribuições usando histogramas!

Algoritmo (modelo BA)

```
função scale_free_graph(N, m) → retorna grafo G
  G ← Crie um grafo pequeno com T vértices (T>=m) e algumas arestas
       //para criar o grafo, pode usar a função random_graph
  node_list ← Inicialize uma lista vazia e dinâmica de número inteiros
  para todo i e j tal que a aresta (i,j) pertença ao grafo G faça
    Adicione o inteiro i a node_list
    Adicione o inteiro j a node_list
  fim_para
  //Segue agora o crescimento do grafo com ligação preferencial
  para i de T+1 até N faça
    Adicione a G o vértice i
    para r de 1 até m faça  //adiciona m novas conexões ao vértice i
      repita
        j ← sorteie um elemento de node_list  //isso guia a ligação preferencial
      até_que i<>j e a aresta (i,j) não pertença a G
      Crie em G uma aresta não dirigida (i,j)
      Adicione o inteiro i a node_list
      Adicione o inteiro j a node_list
    fim_para
  fim_para
  retorne G
fim_função
```

Observação: Neste algoritmo, índices de vértices variam de 1 a N (e não de 0 a N-1)

Resumindo...

- Redes complexas são...

Resumindo...

- Redes complexas são sistemas representados por grafos que apresentam propriedades não triviais

- Quais propriedades?
  - É exatamente isso que os pesquisadores de redes complexas tentam descobrir, compreender e prever!
  - *Exemplos: pequeno-mundo e distribuição "sem escala"*

OUTROS CONCEITOS

- Conceitos clássicos usados em redes complexas:
  - Ciclos
  - Caminhos mínimos
  - Árvores (hierarquias)
  - Árvores geradoras mínimas
  - Componentes conexos (conectados)
  - Percurso (travessia) em largura
  - Etc...

- Além dos algoritmos e estruturas de dados associados

- Além destes, outros conceitos costumam ser aplicados no estudo de redes complexas
  - *Veremos alguns adiante...*
- Note que muitos já existiam antes de se falar em "redes complexas"
  - *Pode-se dizer que na última década houve um renascimento dos estudos em grafos*