Redes Complexas
Aula 10

Roteiro

- Experimento de Milgram
- Mundo pequeno
- Modelo de Watts-Strogatz
- Propriedades estruturais

Aula passada

- Aplicando o G(n,p)
- Avaliando o modelo
- Preferential attachment
- Modelo BA
- Propriedades

Mundo Pequeno

- Seria o mundo um vilarejo?
  - pessoas conectadas por enlaces de amizade
  - redes sociais (reais) como objeto de estudo
  - indagações do início do século XX
- **Experimento de Milgram**: comprovar empiricamente que o mundo é pequeno
  - caminho na rede de amizades entre duas pessoas não relacionadas é curto
- Realizadas por Milgram na década de 60
  - professor de Harvard, psicólogo social influente
  - cunhou o termo “small world”

Experimentos de Milgram

- Enviar uma carta para uma pessoa em Boston
  - nome e formação fornecidos
- Carta deve ser enviada a alguém que você conhece pessoalmente
  - e repassada até chegar ao destino
  - cartão postal enviado a Milgram a cada passo (rota das cartas)
- Origem: pessoas das cidades de Omaha e Kansas (interior dos EUA)
  - centenas de sujeitos, escolhidos aleatoriamente
  - grande distância física e social

Resultados

- Maioria das cartas nunca chegaram ao destino
  - pessoas não passavam carta adiante
- Das que chegaram ao destino
  - comprimento médio dos caminhos foi 5.5 saltos
- Surpreendente
  - população 10^8, dois indivíduos não relacionados estão muito próximos na rede social

Seis graus de separação!

- Experimento ajudou a popularizar esta noção
  - termo não foi cunhado por Milgram

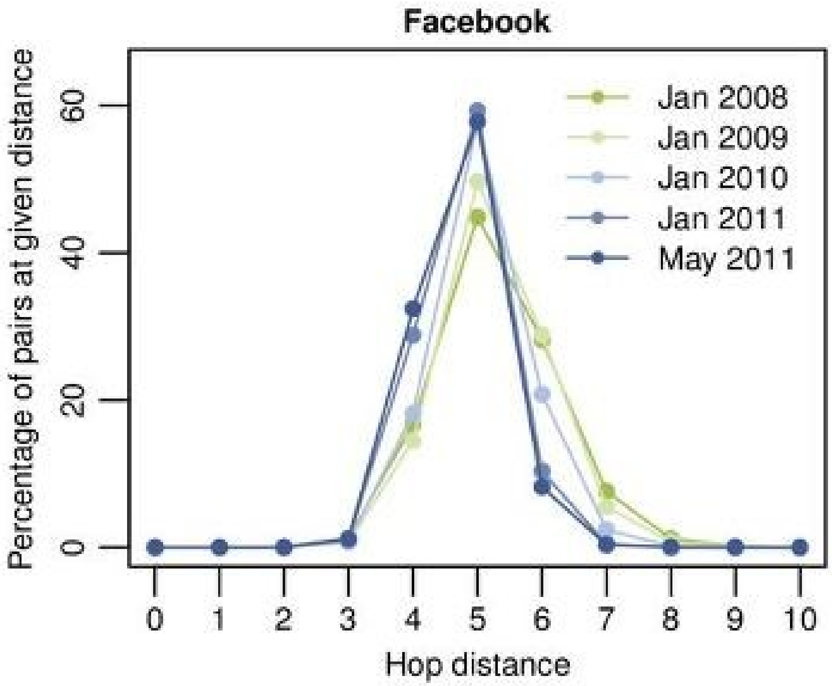
Small World Hoje

- Small World Project (2003), Columbia Univ.
  - versão moderna do experimento: *email ao invés de carta*
  - escala global: milhares de pessoas (muitos voluntários), vários países, 18 alvos diferentes
- Resultados relativamente similares
  - maioria das correntes não chegaram ao destino
  - comprimento médio dos caminhos entre 5 e 7

Peter S. Dodds, Roby Muhamad, Duncan J. Watts,
"An Experimental Study of Search in Global Social Networks", Science (2003)

Small World Hoje

- **Facebook anatomy**
  - menor caminho possível entre duas pessoas na rede de amizades do Facebook
  - diferente do caminho que pessoas encontram



  - n=721 Mi, m=68.7Bi
  - p=0.00000026 (muito esparso)

- Distância média é menor (4.7)
- Distâncias ***diminuem*** com o crescimento da rede!
- Tecnologia reduzindo caminhos na rede social?

Modelo Small World

- Inspirado nos resultados de Milgram
- Proposto por Watts and Strogatz (Cornell, Matemática Aplicada)
  - “*Collective dynamics of* '*small-world*' *networks*” Nature, 1998 – 46K+ citações (Google Scholar)
  - 30 anos depois de Milgram
- Modelo visa construir redes esparsas com distâncias curtas e alta clusterização
  - aspectos observados empiricamente em redes reais

Clusterização e Distâncias

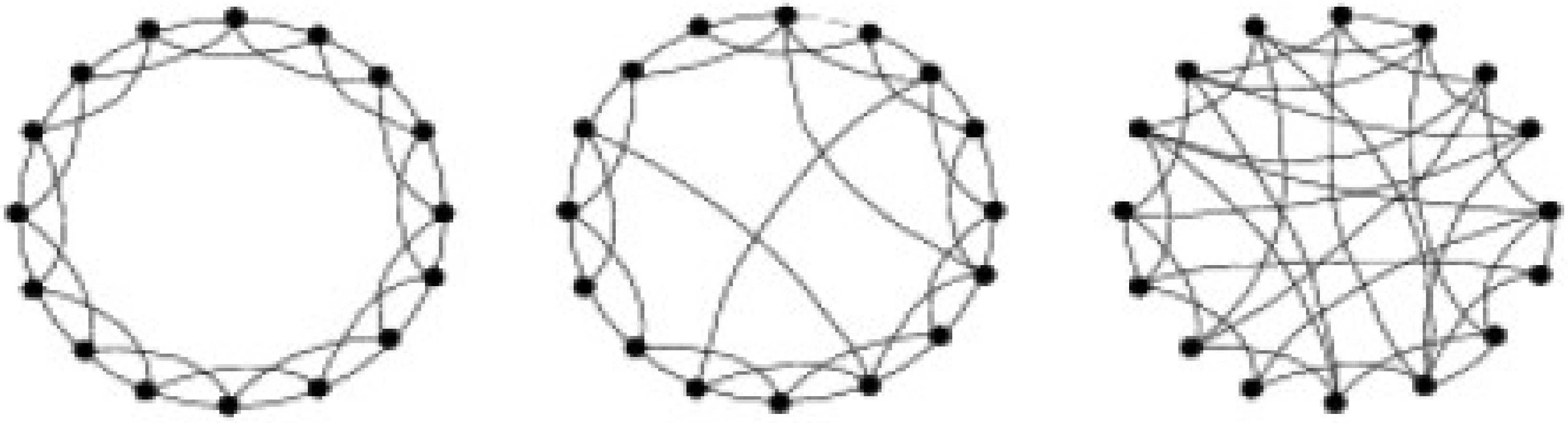
- Considere um grafo esparso
  - densidade de arestas é muito baixa
- Considere alta clusterização e caminhos curtos

Propriedades Antagônicas?

- **Alta clusterização**: muitos triângulos, arestas “desperdiçadas”, caminhos não avançam pela rede
- **Caminhos curtos**: arestas usadas eficientemente, avanço rápido pela rede

Modelo Small World

- Começar com um látice regular
  - N vértices organizados em um círculo
  - arestas para vizinhos a distância k ou menor



- Para cada aresta, reposicionar suas extremidades com probabilidade *p*
  - nova extremidade escolhida de maneira uniforme entre os pares de vértices

Modelos Small World

- Modelo possui 3 parâmetros
  - n, k, p
- k controla grau e clusterização
- p controla aleatoriedade do grafo
  - p = 0, grafo látice regular
  - p = 1, grafo aleatório (todas arestas são reposicionadas aleatoriamente)
- Propriedades topológicas do grafo
  - distância média
  - clusterização
  - distribuição dos graus, etc

Propriedades Estruturais

- Considere n grande, k constante
- C(p) : clusterização em função de p
  - C(0), alta clusterização
  - C(1), baixa clusterização
- l(p) : distância média em função de p
  - l(0), distância média alta
  - l(1), distância média baixa

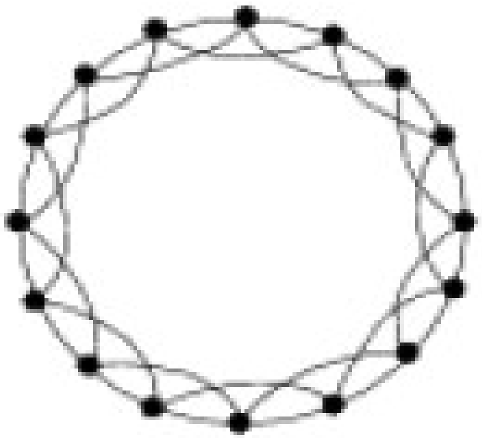
O que ocorre quando *p* é intermediário?

Calculando C(0)

- Clusterização
  - fração de aresta entre vizinhos: $C_i = \frac{E_i}{\binom{d_i}{2}}$
  - Média dos vértices
- C(0) = ?
- Não depende de n, pois é igual para todo vértice
  - mas depende de k
- Ex. k = 2

de arestas entre vizinhos de i

Grau do vértice i



C(0) = 1/2

- Para K genérico?

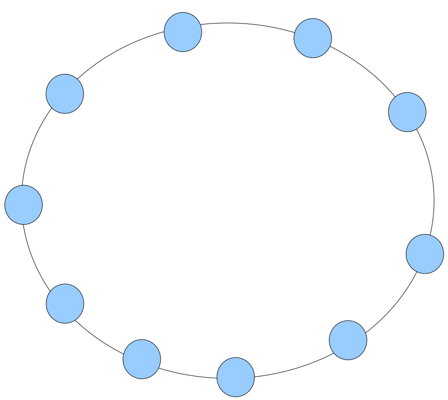
$$C(0) = \frac{3(k-1)}{2(2k-1)} \cdot \sim \frac{3}{4}$$

constante!

Calculando l(0)

$$l = \frac{\sum_{u,v \in V} d(u,v)}{\binom{n}{2}}$$

- Distância média entre vértices
  - Entre todos os pares de vértices
- l(0) = ?
- Depende de n e de k
- Ex. n ímpar, k = 1



$$l(0) = \frac{1+2+3+\ldots+n/2}{n/2} = \frac{n}{4} + \frac{1}{2}$$

- Para K genérico?

$$l(0) \sim \frac{n}{4k}$$

Crescimento
linear com n

Calculando C(1) e l(1)

- Com p=1, grafo é totalmente aleatório
- Modelo G(n,m), onde m é o número fixo de arestas
  - ponta das arestas escolhidas aleatoriamente entre todos os vértices (sem repetição)
  - primo do G(n,p)
  - propriedades estruturais equivalentes
- No G(n,p): clusterização e distância média?

$$C = p \qquad\qquad l \sim \frac{\log(n)}{\log(d)}$$

Calculando C(1) e l(1)

- Modelo SW
  - Grau médio é 2*k*
  - Número de arestas é *nk*
- Fração de arestas

$$\rho = \frac{nk}{\binom{n}{2}} = \frac{2k}{n-1} \sim \frac{2k}{n}$$

- Logo, com p=1

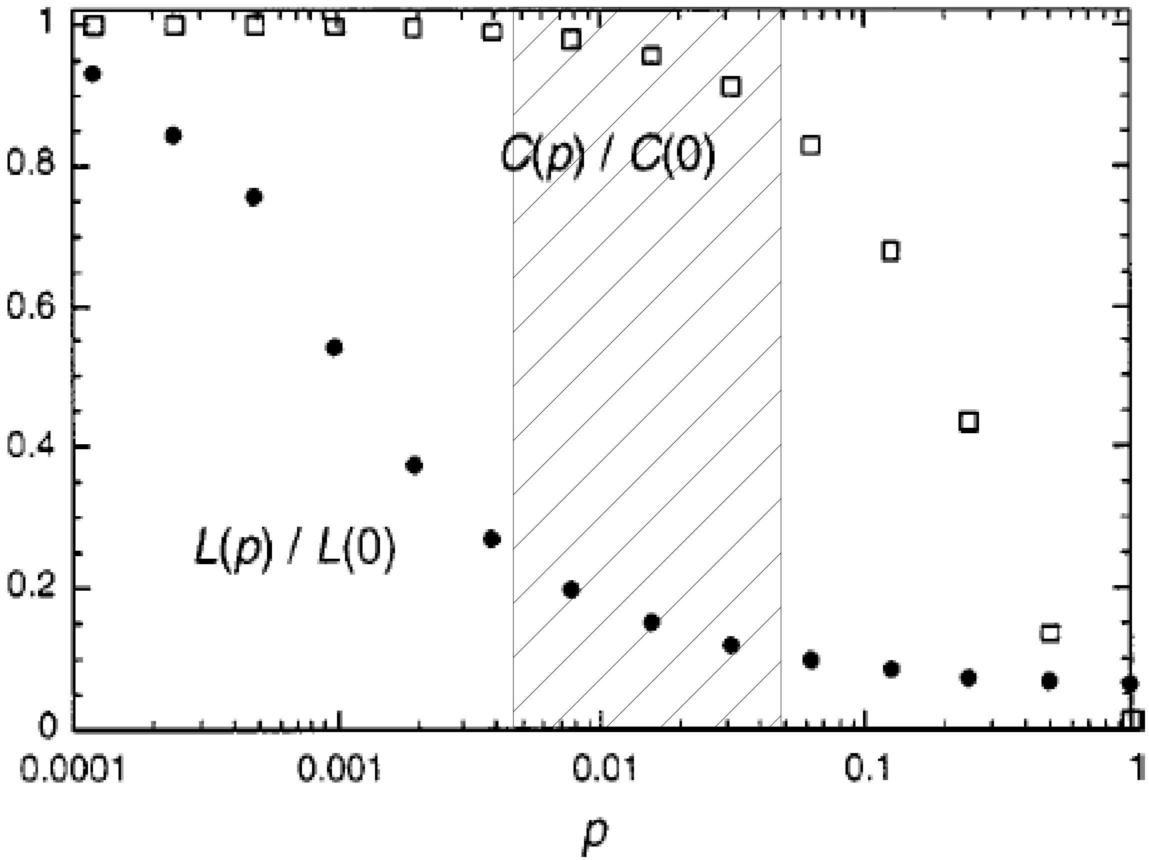
$$C(1) = \frac{2k}{n} \qquad l(1) = \frac{\log(n)}{\log(2k)}$$

decrescе
linearmente com n

Crescimento
logaritmico com n

Análise Numérica

- Relação entre C(p)/C(0) e l(p)/l(0)
  - C(0) maior clusterização, l(0) maior distância média



- Distância média decresce rapidamente
- Clusterização decresce mais devagar
- Alta clusterização, ditâncias curtas ocorrem (para alguns valores de *p*)
- Small World Networks

Distância nos Extremos

- p = 0

$$l(0) \sim \frac{n}{2k}$$

Distância cresce linearmente com n

- p = 1

$$l(1) = \frac{\log n}{\log(2k)}$$

Distância cresce logaritmicamente com n

- Mudança fundamental de comportamento em função do valor de *p*

Onde ocorre transição?

Análise da Distância

- Distância decresce rapidamente em p
- Intuição: poucas arestas aleatórias são suficientes para diminuir distâncias significativamente
  - arestas criam atalhos
- Número médio de atalhos é dado pelo número médio de arestas reposicionadas
  - $a = p\,k\,n$
- Transição de fase em função de *a*

$$l(n,k,p) \sim n/k$$ quando $a \ll 1$, $a = o(1)$

$$l(n,k,p) \sim \log(p\,k\,n)/(p\,k^2)$$ quando a >> 1, a = $\Omega(1)$

Críticas ao Modelo SW

- Distribuição de grau não possui cauda pesada
  - próximo da binomial, baixa diversidade de graus
- Modelo não permite crescimento da rede
- Modelo não é inspirado em um processo de formação
- Análise teórica rigorosa de suas propriedades é complicado

Todas as limitações acima superadas em modelos subsequentes!