Redes Complexas
Aula 12

Aula passada

- Preferential Attachment
- Modelo BA
- Lei de potência

Aula de hoje

- Experimento de Milgram
- Modelo “Small World”
- Propriedades estruturais

Mundo Pequeno

- Seria o mundo um vilarejo?
  - início do século XX
  - redes sociais (reais) como objeto de estudo
- **Experimento de Milgram**: “comprovar” que o mundo era pequeno
  - caminho entre duas pessoas quaisquer pela rede social (real) deve ser 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 fornecida
- Carta deve ser enviada a alguém que você conhece pessoalmente
  - e repassada até chegar ao destino
  - postal enviado a Milgram a cada passo (rota das cartas)
- Sujeitos (origem): pessoas em 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
  - média do comprimento do caminho: 5.5
- Surpreendente!
  - População 10^8, dois indivíduos aleatórios estão próximos na rede social

Seis graus de separação!

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

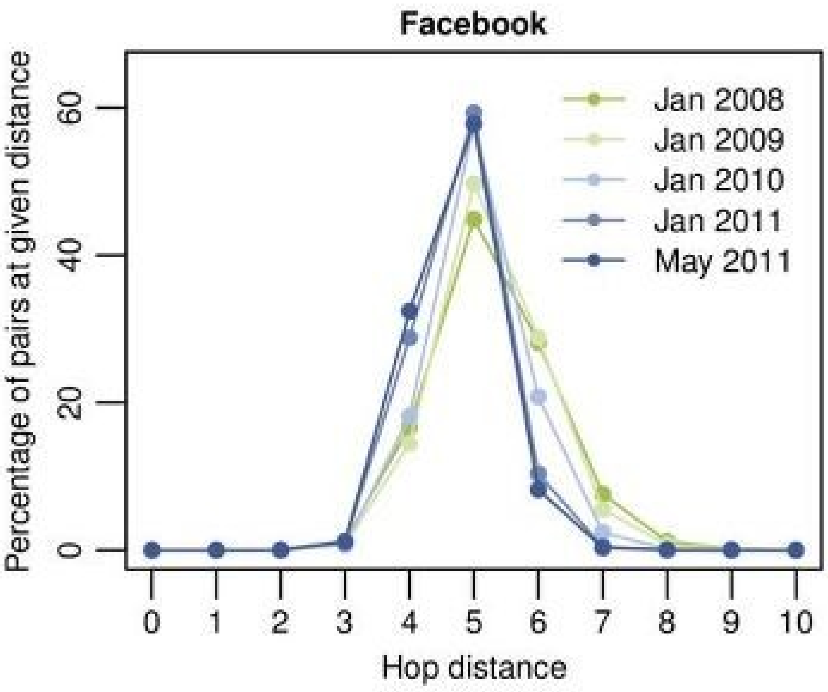
Small World Hoje

- Small World Project (2003), Columbia Univ.
  - versão moderna do experimento: *email ao invés de carta*
  - escala global (milhares de pessoas, em vários países, 18 alvos diferentes)
- Resultados similares
  - maioria das correntes não chegam ao destino
  - média do caminho 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**: Cuidado, experimento é diferente!
  - menor caminho possível na rede 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
  - e alta clusterização observada empiricamente
- Proposto por Watts and Strogatz (Cornell, Matemática Aplicada)
  - “*Collective dynamics of* '*small-world*' *networks*” Nature, 1998 – 32000 citações (Google Scholar)
  - 30 anos depois de Milgram!
- Modelo captura as duas propriedades
- Mostra que ambas podem coexistir em um mesmo grafo esparso

Clusterização e Distâncias

- Considere um grafo esparso
  - número de arestas relativamente baixo
- Considere alta clusterização e caminhos curtos

Propriedades Antagônicas?

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

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 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 (idêntico ao G(n, p)?)
- Propriedades topológicas do grafo
  - distância média
  - clusterização
  - distribuição dos graus, etc

Propriedades Estruturais

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

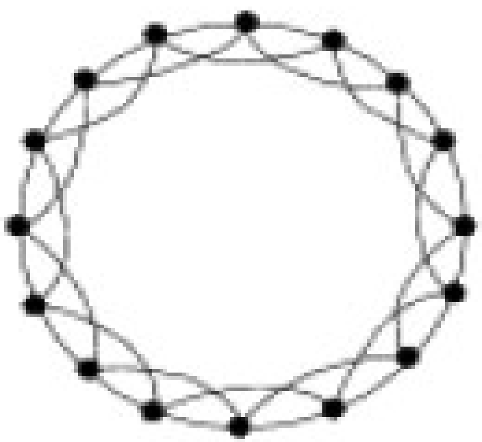
O que ocorre quando *p* é intermediário?

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 n
  - mas depende de K
- Ex. K = 2

de vizinhos de i

Grau do vértice i



C(0) = 1/2

- Para K genérico?

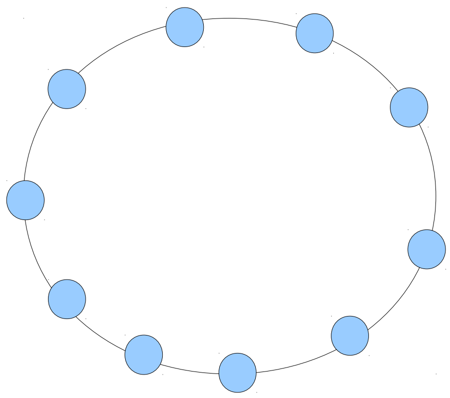
$$C(0) = \frac{3(K-1)}{2(2K-1)} \sim \frac{3}{4}$$

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
  - 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, com p = 1
- Dado K, grau médio é 2K
- Dado N e K, $m = N K$
- Fração de arestas: $p = \frac{2K}{N-1} \sim \frac{2K}{N}$
- Logo

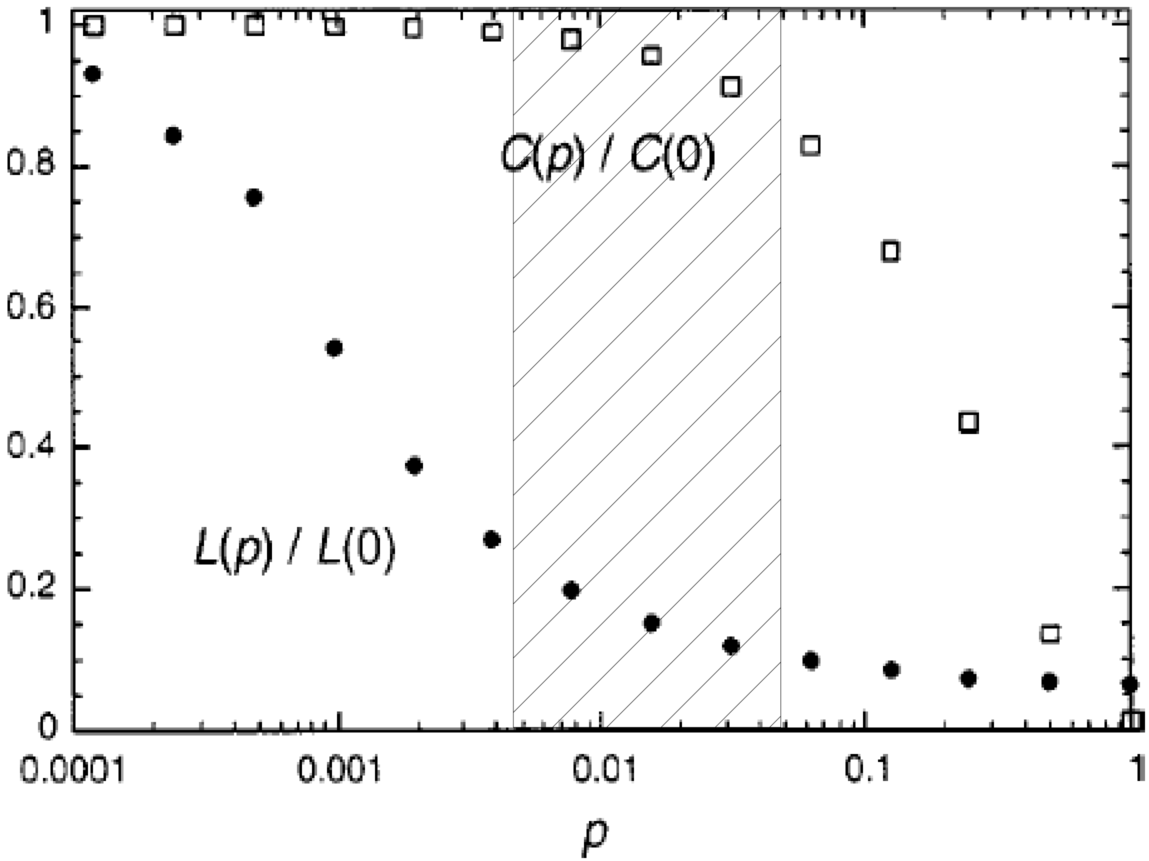
$$C(1) = \frac{2K}{N} \qquad l(1) = \frac{\log(N)}{\log(2K)}$$

decresce
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 é possível!
- 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

Onde ocorre transição?

Análise da Distância

- Distância decresce rapidamente em p
- Intuição: poucas arestas aleatórias são suficientes para significativamente diminuir distâncias
  - arestas criam atalhos
- Número médio de atalhos?
- a = p K N
- Transição de fase em a

$$l(N, K, p) \sim N/K$$ quando a << 1

$$l(N, K, p) \sim \log(p\,K\,N)/(p\,K^2)$$ quando a >> 1

Críticas ao Modelo SW

- Modelo não é inspirado em nenhum processo de formação de redes
  - nem é muito interessante do ponto de vista teórico
  - variação do modelo possui melhor tratabilidade
- Distribuição de grau não possui cauda pesada
  - Baixa variância, próximo binomial
- Modelo não generativo e sem crescimento