Introdução a Redes Complexas

Jornadas de Atualização em Informática (JAI)
CSBC 2011

Encontro 3/3

Daniel R. Figueiredo
LAND – PESC/COPPE/UFRJ

Organização do Mini-curso

- Três encontros (Qui 17h, Sex 11h, Sex 17h)
  - duas horas cada, brinde na Sex 17h
- **Encontro 1**
  - Redes por todos os lados, formalismo e propriedades, descobrimentos empíricos
- **Encontro 2**
  - Lei de potência, modelos de redes: G(n,p), Barabási-Albert, Watts-Strogatz, propriedade dos modelos
- **Encontro 3**
  - Funcionalidade em redes: robustez, busca e navigabilidade

Interatividade com participação do público!

Funcionalidade em Redes

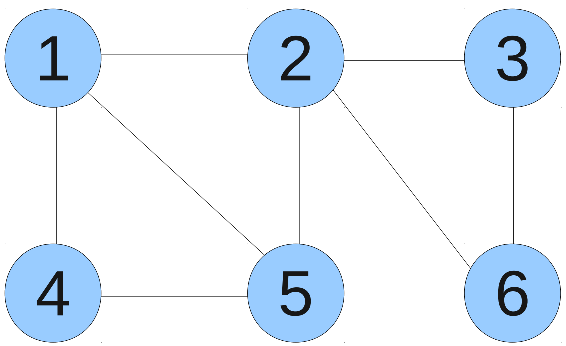
- Como estudar a estrutura de como as coisas se conectam?
  - ex. Internet, Facebook

Funcionalidade!

- Funcionalidade são os processos que operam sobre a rede
  - a razão para a rede (ou outros processos)
- Funcionalidade em função da estrutura

Redes e Falhas

- Rede: abstração de um sistema real
  - ex. Internet
- Componentes da rede podem falhar
  - vértices ou arestas



O que acontece quando falhas ocorrem?

Resilience ("robustez")

- Capacidade da rede de *operar* na presença de falhas
  - vértices ou/e arestas
- Medida da redundância existente na rede
  - muitas métricas
- Robustez em função da topologia
  - impacto da topologia na robustez
- Aplicação depende do contexto
  - ex. robustez da Internet a falhas nos AS

Métricas

- Métricas locais
  - impacto em um ou poucos vértices
  - ex. número de arestas para desconectar um vértice qualquer
- Métricas Globais
  - impacto na rede como um todo
  - tamanho da componente gigante
  - distância média entre os vértices

Interesse em métricas globais

Tipo de Falhas

- Como vértices ou arestas falham?
- Aleatoriamente
  - uniforme: todos tem mesma prob de falhar
  - não-uniforme: algum fenômeno
- Deterministicamente
  - Definir ordenação: quem falha primeiro, segundo, etc.
  - ex. ordenação decrescente por grau
- Falhas por defeito: aleatório
  - ex. roteador queimou
- Falhas propositais: determinístico
  - ex. roteador foi atacado

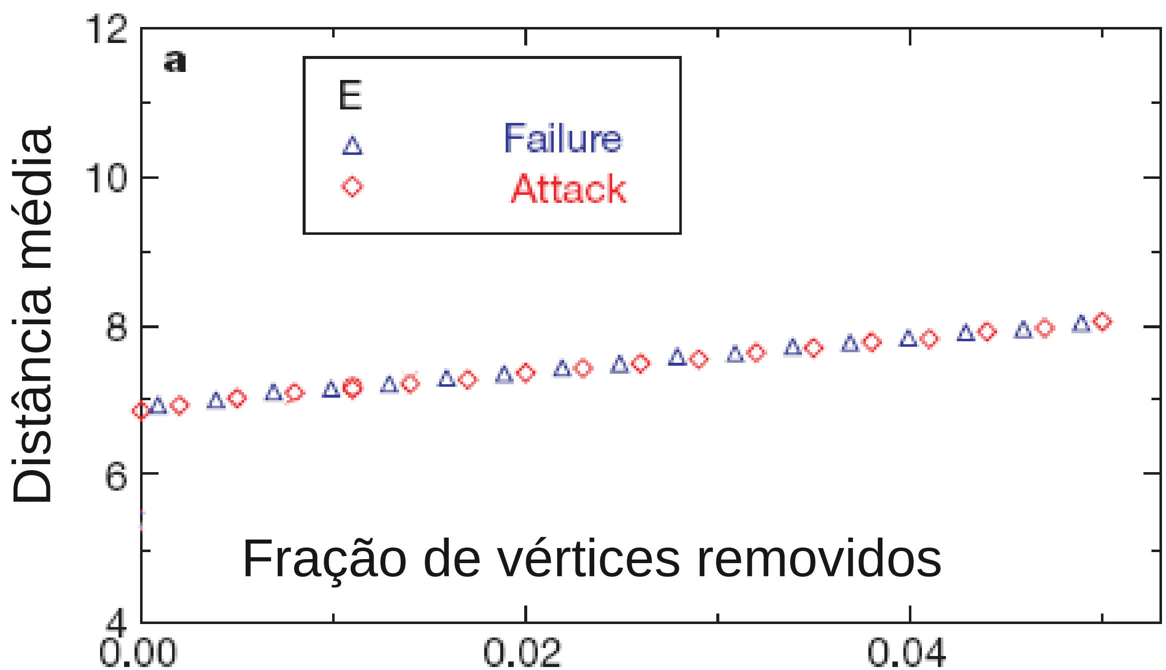
Resilience e Falhas

- Robustez da rede na presença de falhas
  - Aleatórias (uniforme), determinísitica (grau)
  - Ex. distância média entre pares
- Diferença com relação ao tipo de falha?
  - para uma mesma fração que falha
- Influência da topologia da rede?
  - para mesmo n e grau médio

Modelo G(n,p)? Modelo BA?

Falhas no G(n,p)

- **Intuição?**
- Falhas aumentam (mas não muito) a distância média entre vértices
- Falhas aleatórias se parecem com falhas determinísticas
  - grau máximo é próximo ao grau médio



- N = 10000, <D> = 4

Falhas no BA

- **Intuição?**
- Falhas aleatórias **não** se parecem com falhas determinísticas
- Falhas aleatórias: pouco impacto; falhas determinísticas: muito impacto

Falhas Aleatórias

- Rede AP, lei de potência

- Remoção aleatória (uniforme)
- Pouco impacto, rede robusta

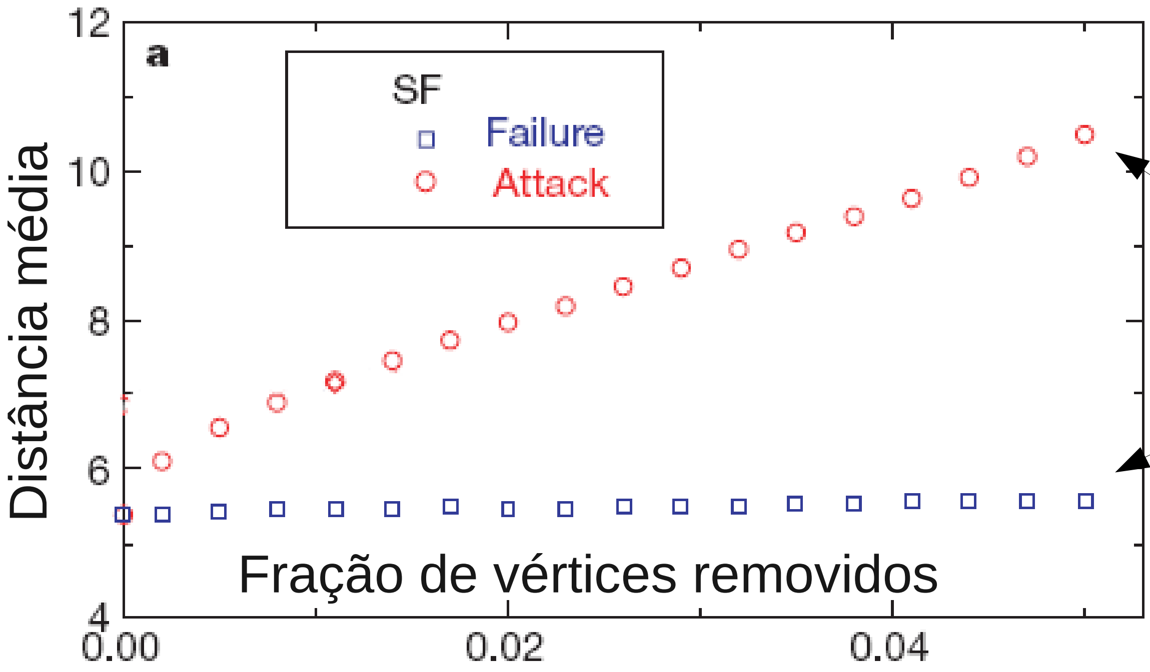
Falhas Determinísticas

- Rede AP, lei de potência
- Remoção por ordem decrescente do grau

- Impacto devastador
- Estrutura tem papel fundamental

Falhas no BA

- **Intuição?**
- Falhas aleatórias **não** se parecem com falhas determinísticas
- Falhas aleatórias: pouco impacto; falhas determinísticas: muito impacto



- N = 10000, <D> = 4

5% dos vértices removidos; distância média dobrou!

5% dos vértices removidos; distância média não variou

Modelo BA

- Distribuição do grau segue lei de potência

- Maioria dos vértices tem grau pequeno
- Pouca importância na rede

↓

- Tolerante a falhas aleatórias

- Poucos vértices com grau grande
- Interconectam a rede

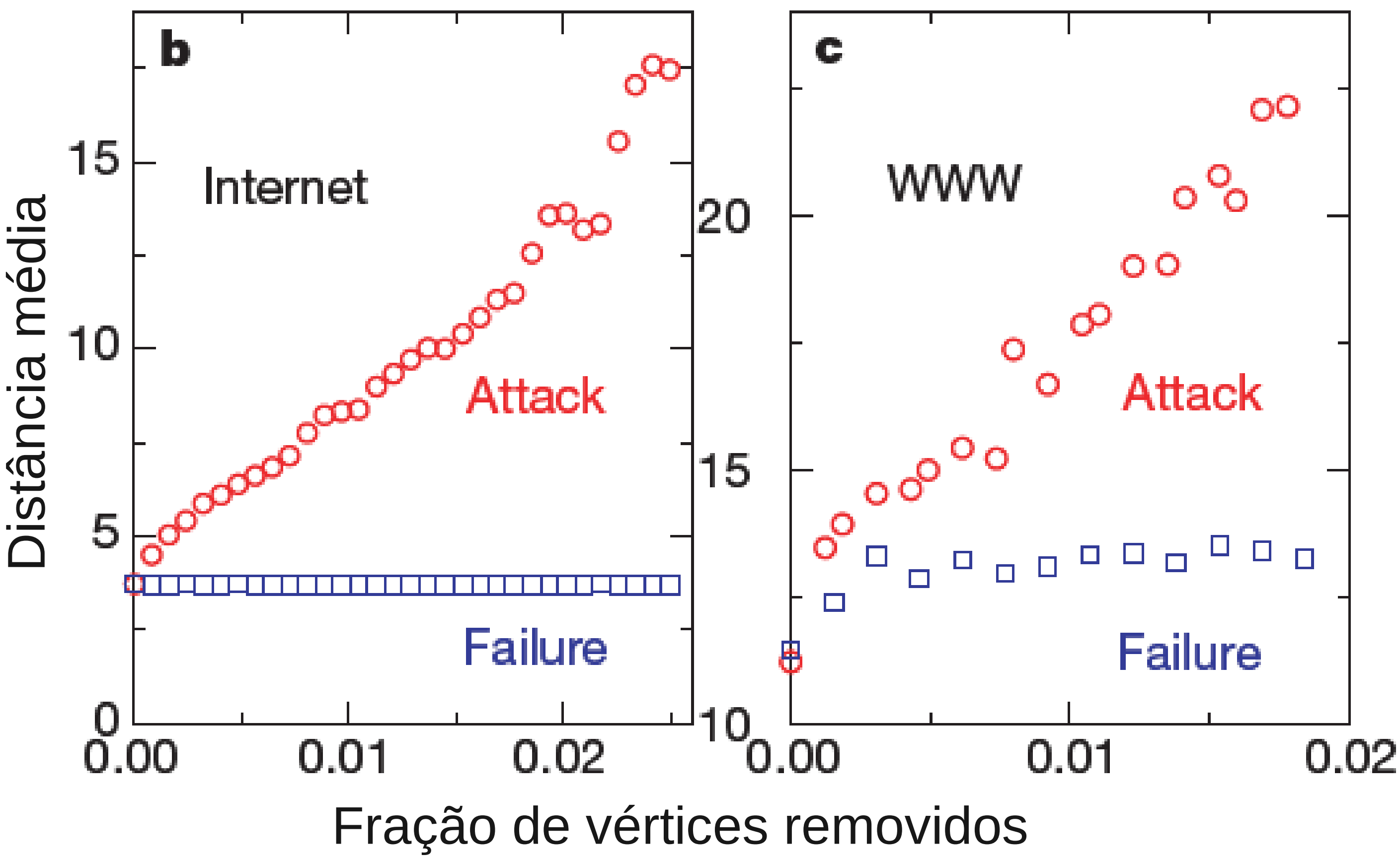
↓

- Vulnerável a ataques direcionados

Propriedade de redes livre de escala!

Robustez em Redes Reais

- WebGraph, AS Graph
  - Redes seguem lei de potência no graus



- Comportamento parecido com modelo BA
- Tolerante a falhas, vulnerável a ataques
- AS não é roteador!

Problema de Busca

- Como encontrar informação em uma rede?
  - grande e desconhecida (globalmente)
- Ex. Encontrar uma pessoa que já orbitou a terra?
- Ex. Ser apresentado a um prêmio Nobel?

Busca apenas com informação local

- Algoritmo distribuído, navegação pela estrutura
- Como explorar a estrutura?

Experimento de Milgram

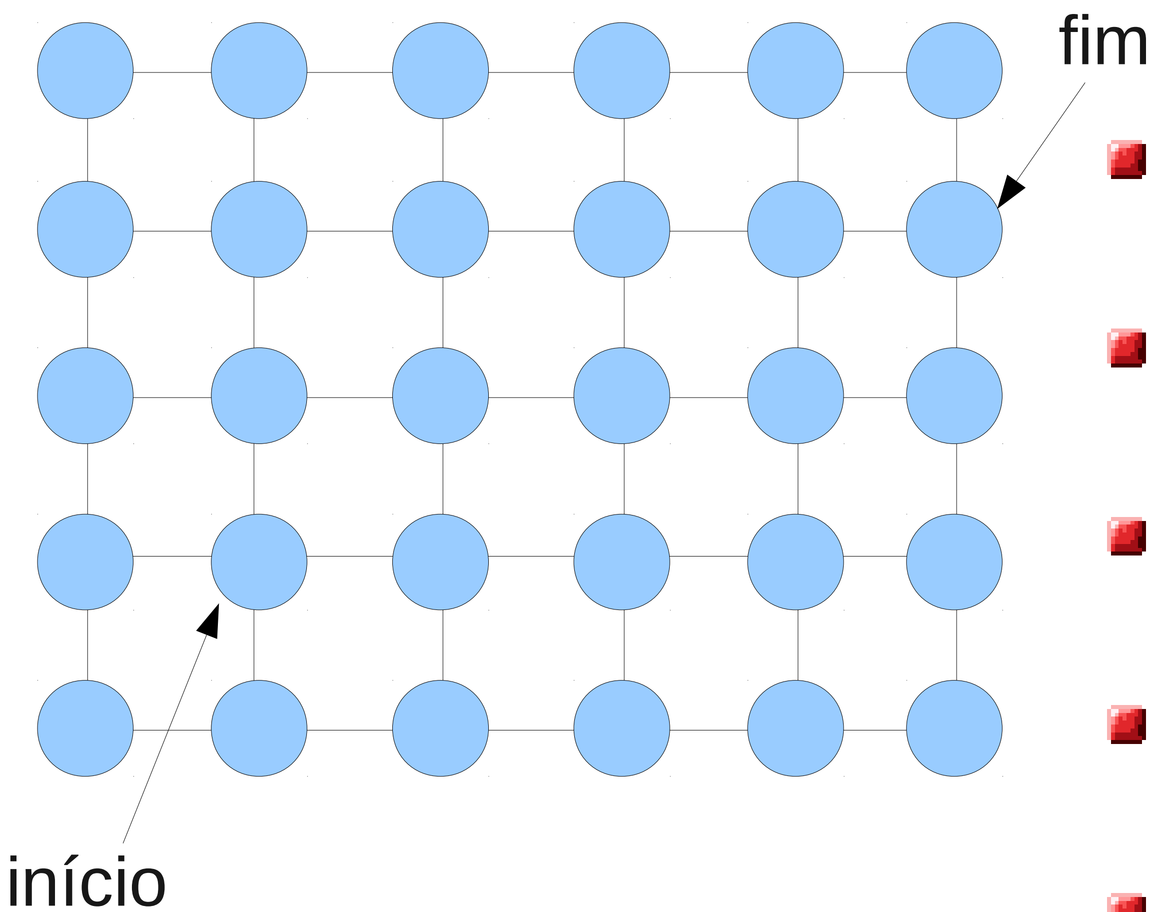
- Enviar carta a um desconhecido através apenas da rede social
- **Surpreendente**: Caminhos curtos existem
  - ~ 6 passos na rede social
- **Igualmente surpreendente**: Pessoas comuns encontraram caminhos curtos
  - rede social é vasta, busca aleatória levaria a caminhos muito mais longos
  - pessoas não têm conhecimento da rede social
- Pessoas navegam bem na rede social
- Mas como?

Algoritmo de Busca Guloso

- Algoritmo guloso
  - a cada passo, busca transferida para vizinho ***mais próximo*** do destino
- Algoritmo é míope
  - considera apenas próximo passo
  - tenta se aproximar o máximo do destino a cada passo
- Métrica para proximidade depende do contexto
  - distância física, distância social, etc.
- Provavelmente usado por pessoas para navegar na rede social

Exemplo

- Rede: grid em 2D
- Distância: distância no grid (em saltos)



- Vértice inicial conhece coordenadas do fim
- Escolhe vizinho que está mais próximo do fim
- Passa busca para vértice vizinho
- Vizinho repete procedimento
- **Complexidade?**

Desempenho do Guloso

- Desempenho do algoritmo depende da topologia da rede
- Guloso pode **não encontrar** caminho mais curto na rede
  - apesar do caminho existir!
- Estrutura da rede é fundamental para desempenho do algoritmo
  - caminhos exponencialmente menores
- Observação e demonstração feita por Kleinberg em artigo influente

Modelo de Kleinberg

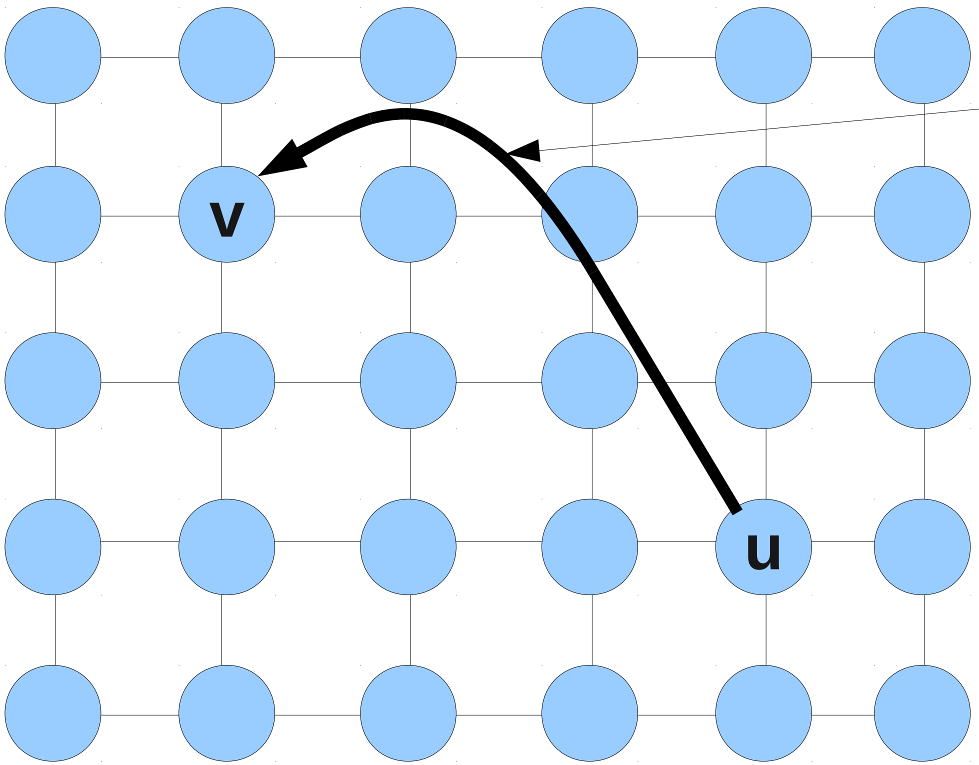
- Rede: grid em 2D
- Distância: distância no grid (em saltos)
- Adicionar “atalhos” no grid
  - criação de caminhos curtos
  - cada vértice adiciona q atalhos
- Atalhos probabilísticos
  - probabilidade inversamente proporcional a distância entre os vértices no grid

$$p_{uv} \propto \frac{1}{d(u,v)^{\alpha}}$$

, onde *d(u,v)* é a distância entre os vértices *u* e *v* no grid

Modelo de Kleinberg

- Rede: grid em 2D + atalhos probabilísticos



Atalho ocorre com probabilidade ~ $5^{-\alpha}$

- Atalhos é função de α
- α grande: atalhos longos ocorrem mais raramente
- α pequeno: atalhos muito longos podem ocorrer

- **Problema**: Calcular desempenho do algortimo guloso em função de α

Modelo de Kleinberg

- Assumir rede é um látice, n x n vértices
- Cada vértice tem exatamente um atalho
- Probabilidade de atalho de u levar ao vértice v

$$p_{uv} = \frac{d(u,v)^{-\alpha}}{\sum_{u \neq v} d(u,v)^{-\alpha}}$$

, onde *d(u,v)* é a distância entre os vértices *u* e *v* no grid

- Parâmetro α constante > 0
- Desempenho: numero médio de saltos para encaminhar mensagem do origem ao destino (T)
  - escolhidos também aleatoriamente

Avaliação Teórica

- 0 <= α < 2 ⟶ $T \geq c\, n^{(2-\alpha)/3}$
- α > 2 ⟶ $T \geq c\, n^{(\alpha-2)/(\alpha-1)}$

Tempo médio é polinomial!

- α = 2 ⟶ $T \leq c \log^2 n$

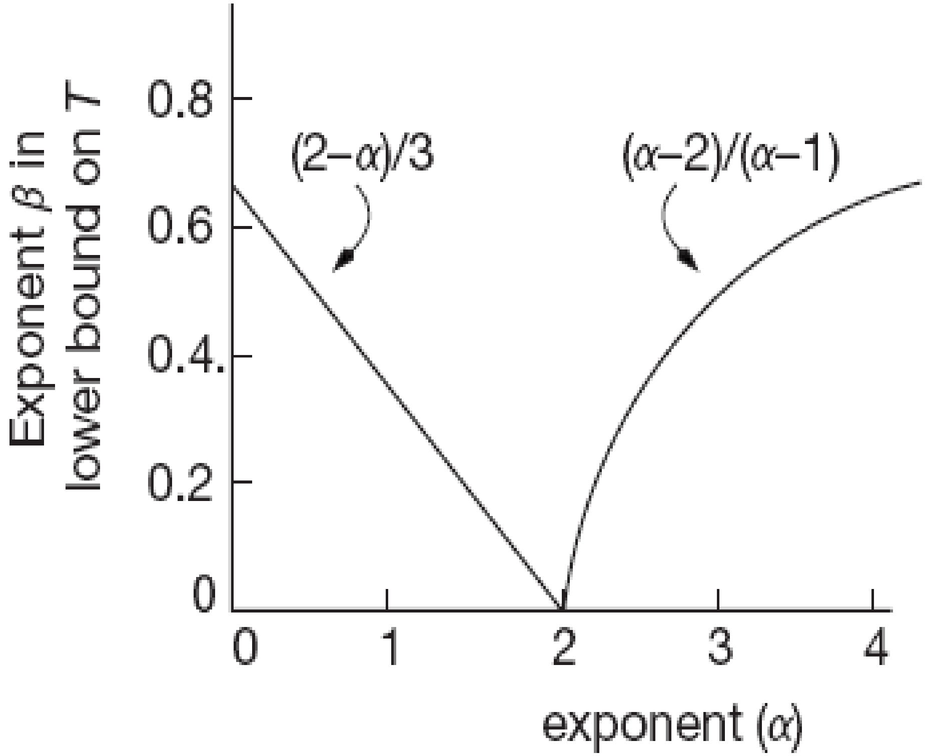
Exponencialmente menor!

- Resultado vale para qualquer algoritmo que utiliza apenas informação local (qualquer variação do guloso)

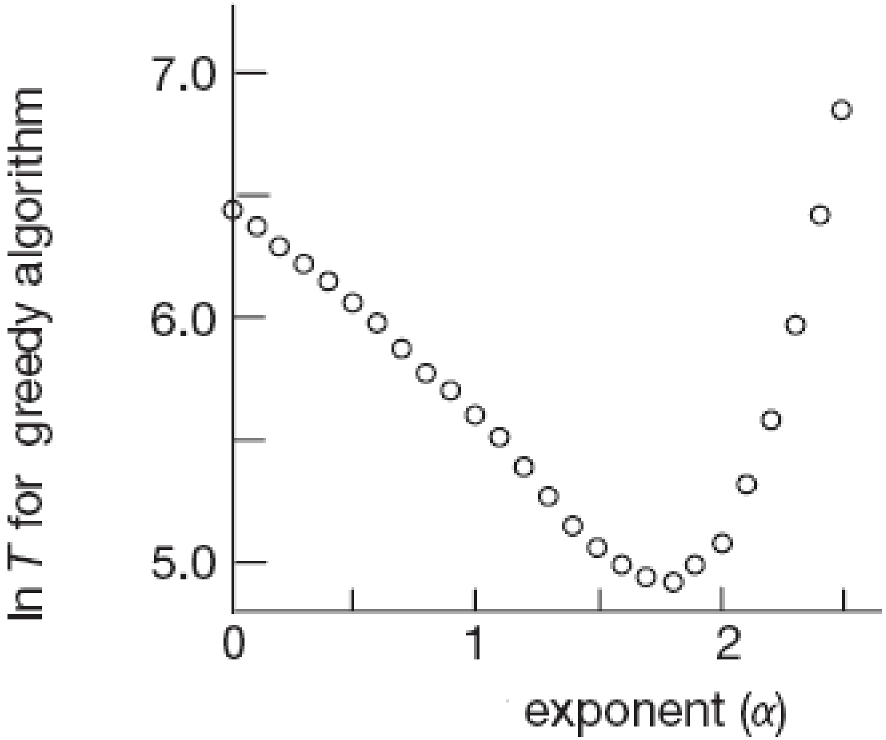
Avaliação

- Gráfico do expoente β

$$T \geq c\, n^{\beta}$$



- Avaliação empírica (simulação), n = 20K, 1K rodadas, log T



Observações

- $\alpha = 2$: vértices tem em média o mesmo número de vizinhos em todas as distâncias
  - Efeitos se cancelam
- $\alpha < 2$: distribuição de atalhos muito uniforme
  - algoritmo não consegue explorar os atalhos
- $\alpha > 2$: não há atalhos longo suficientes
  - não existem caminhos curtos

Idéia da Prova

- Para o caso $\alpha = 2$
- Argumento geométrico
- Algoritmo procede em fases
  - Fase j se distância até destino entre 2^j e 2^{j+1}
- Quantas fases, no máximo?
  - log n
- Número médio de passos até mudar de fase?
  - encontrar atalho para próxima fase
  - log n
- Logo, $T <= c (\log n)^2$

Generalizações

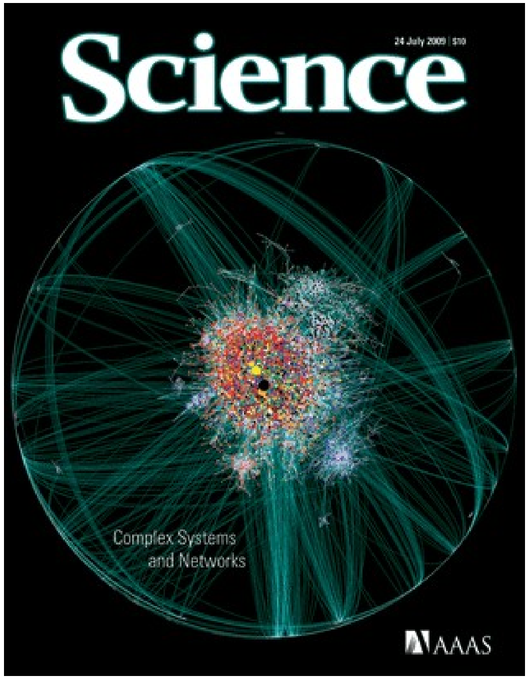
- Generalizado para qualquer número constante de atalhos
  - não muda complexidade
- Generalizado para qualquer algoritmo que utiliza apenas informação local
  - guloso não usa o passado
- Generalizado para grids com d dimensões
- Ponto crítico ocorre quando $\alpha = d$
  - guloso encontra caminhos curtos apenas nestes caso
  - mesma intuição

O Essencial

- Redes por todos os lados, estrutura no centro das atenções
  - Implicações da estrutura
- Modelos de rede (estrutura)
  - Como as coisas se conectam
- Funcionalidade e estrutura
  - Objetivo maior, por que?
- Multidisciplina intrínsica
  - física, matemática, computação, biologia, sociologia, etc

Redes Complexas

- Área de pesquisa emergente
  - completou 12 anos
- Diversos institutos em Redes Complexas sendo criados
  - multidisciplinares
- Muitas questões fundamentais ainda em aberto
  - razão destas estuturas
  - evolução e dinamismo
  - funcionalidade e estrutura



FIM

- Obrigado pela atenção
- Perguntas, dúvidas, comentários?

- Contato: daniel@land.ufrj.br
- http://www.land.ufrj.br/~daniel