Redes Complexas
Aula 11

Aula passada

- Modelo G(n, p)
- Threshold functions
- Evolução da estrutura
- Mais propriedades

Aula de hoje

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

Aplicando o G(n,p)

- Como aplicar modelo G(n,p) a uma rede real?

- Determinar parâmetros do modelo de acordo com rede real
  - parâmetros: *n* e *p*
- Ex. AS Graph, 10697 vértices e 31992 arestas

Quanto vale *n* e *p* neste caso?

Aplicando o G(n,p)

- Determinar *p* de forma a preservar grau médio da rede real
- Grau médio do grafo empírico com n vértices e m arestas $\longrightarrow$ $$\bar{d} = \frac{2m}{n}$$
- Grau médio do G(n,p) $\longrightarrow$ $$\bar{d}_G = (n-1)p$$
- Temos então que

$$p = \frac{2m}{n(n-1)}$$

Exemplos

- AS Graph, 10697 vértices e 31992 arestas
- Rede de atores, 449913 vértices e 25516482 arestas
- Rede metabólica, 765 vértices 3686 arestas

- d = 5.98
- p = 0.00056

- d = 113.4
- p = 0.00025

- d = 9.67
- p = 0.0126

Observação?

- p geralmente muito pequeno (grafos esparsos)

Adequação do Modelo G(n,p)

- O quão adequado é o modelo G(n,p) para representar redes reais?

Captura número de vértices e grau médio!

- Suficiente?
- Depende dos objetivos...

Captura outros aspectos estruturais de redes reais?

Aspectos Estruturais

- Presente em muitas redes reais
  - Distâncias pequenas
  - Alta clusterização
  - Distribuição do grau em cauda pesada
- Modelo G(n,p)
  - Distâncias: O(log n / log z)
  - Clusterização: p
  - Distribuição do grau: Binomial(n-1, p)

Muito diferentes!

Aspectos fundamentais para muitas aplicações

Exemplo

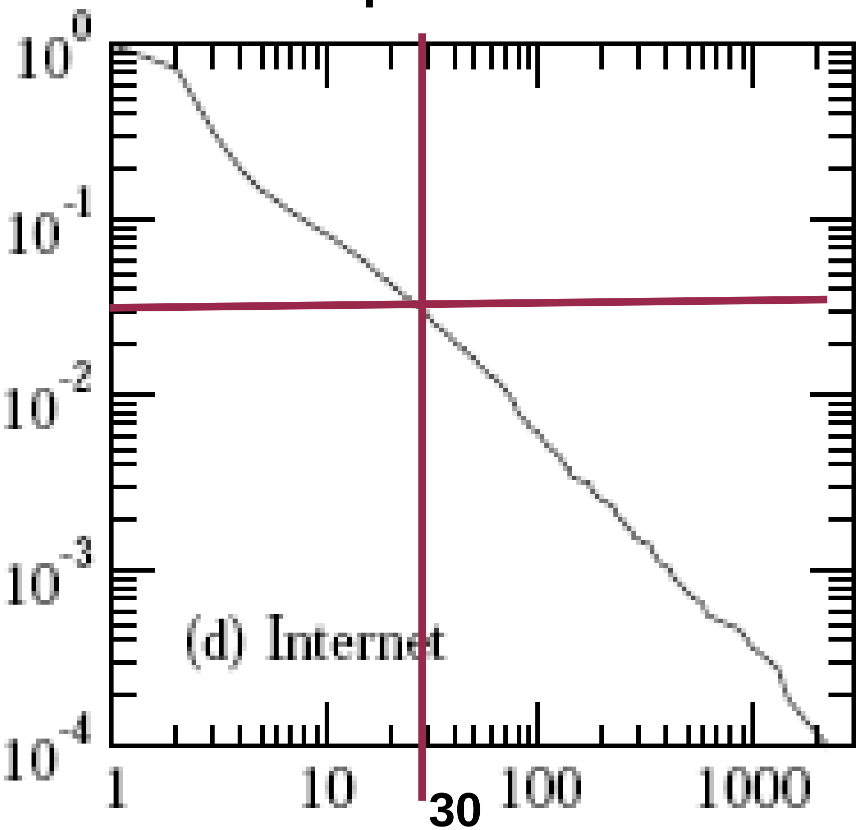
- AS Graph, 10697 vértices e 31992 arestas
- Distância média
  - empírica: 3.31
  - modelo: O(log n/log z) = 5.18
- Clusterização
  - empírica: 0.39
  - modelo: p = 0.00056

~1000 vezes menor!

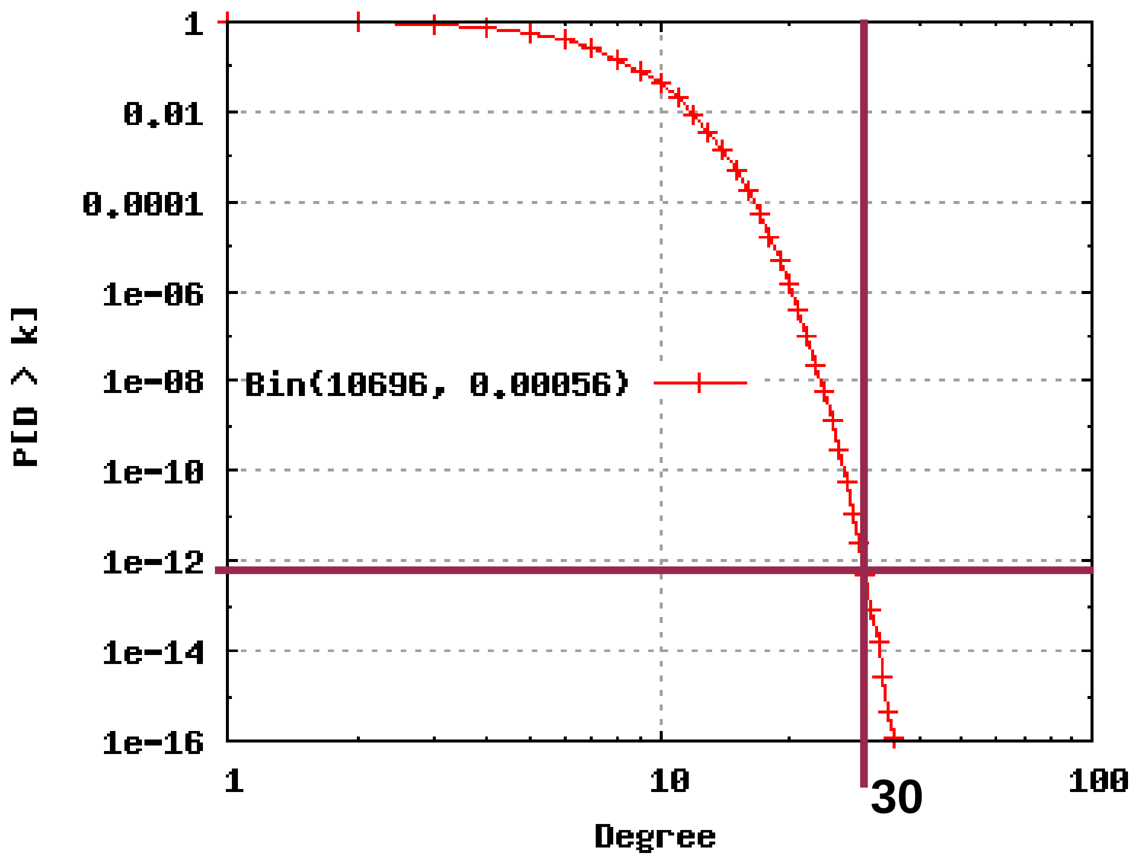
Exemplo

- AS Graph, 10697 vértices e 31992 arestas
- Distribuição do grau: P[D > k]

Empírica



Modelo G(n,p)



Fudamentalmente diferente!

Figueiredo – 2015

Modelos de Redes

- **Objetivo**: modelos que capturem aspectos importantes de redes reais

- Duas abordagens
- Modelos diretos
  - Modelo define diretamente estrutura do grafo
  - Ex. sequência de graus é dada
- Modelos de crescimento (*growth*)
  - Baseados em algum processo incremental ou evolucionário
  - Processo iterativo *leva* à rede

Modelos de Crescimento

- Explicam o surgimento das propriedades estruturais
  - estrutura é função do processo gerador
- Vértices e arestas são adicionados incrementalmente
  - refletem o "crescimento" da rede
- Processo de crescimento tenta capturar a realidade
  - estrutura observada é consequência da dinâmica

Muitos modelos propostos são desta classe

Preferential Attachment (PA)

- Regra fundamental (e antiga) de formação
  - aka. cumulative advantage, rich-gets-richer, Matthew effect
- **Idéia:** objetos têm preferência em se relacionar com objetos mais populares

$$P[u \sim v] \approx pop_v$$

prob. do objeto u
se relacionar com objeto v

- Objetos populares "atraem" novos objetos
- Idéia que possui muitas aplicações

Exemplos ?

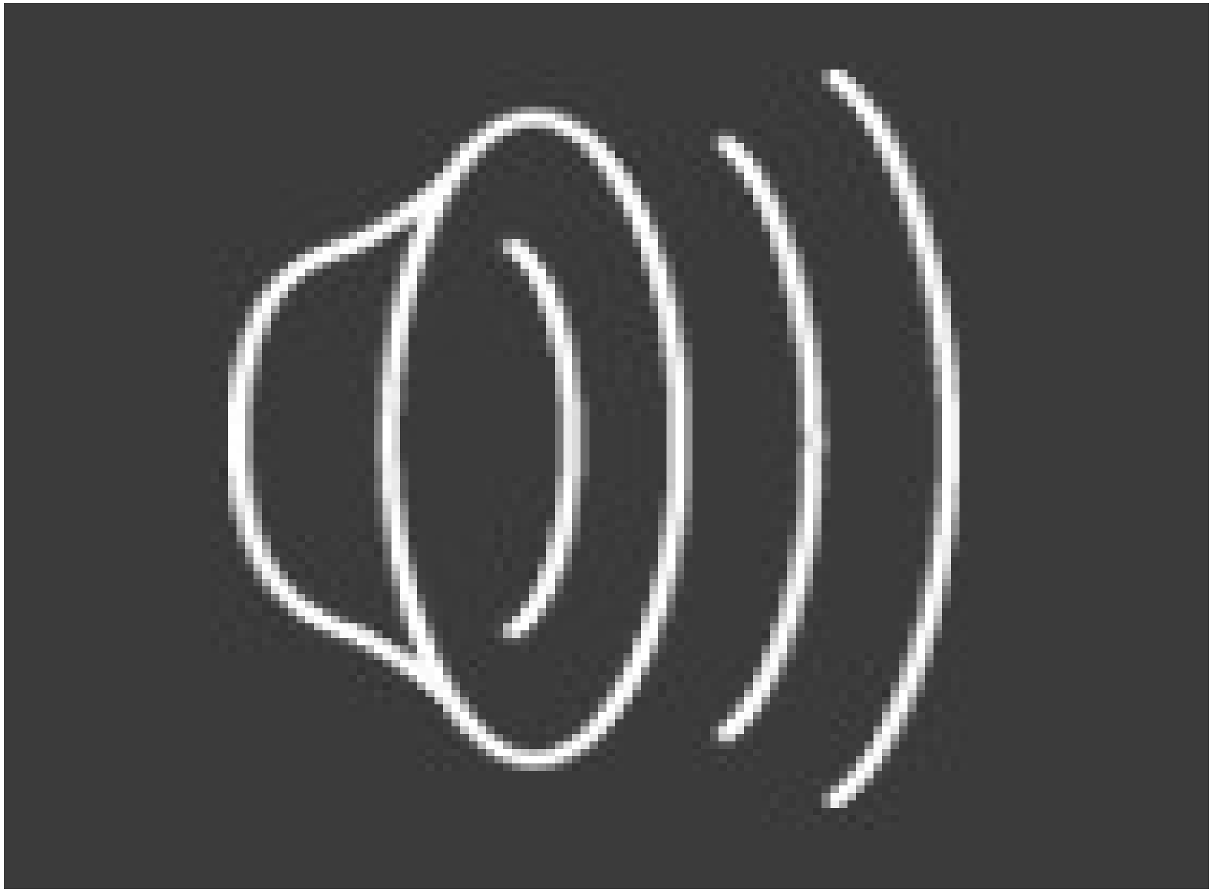
Exemplos (PA)

- Criação de links na web (Barabasi et al 99)
  - Páginas novas tendem a criar links para páginas mais populares
- Algomeração de pessoas em cidades (Zipf 32)
  - Cidades maiores tendem a atrair mais pessoas
- Espécies de plantas de um mesmo gênero (Yule 25)
  - Mutações são mais prováveis em gênero com mais espécies, levando a outras espécies
- Número de artigos publicados (Simon 55)
  - artigos novos tendem a ter co-autores mais populares

Modelo BA (Barabási-Albert)

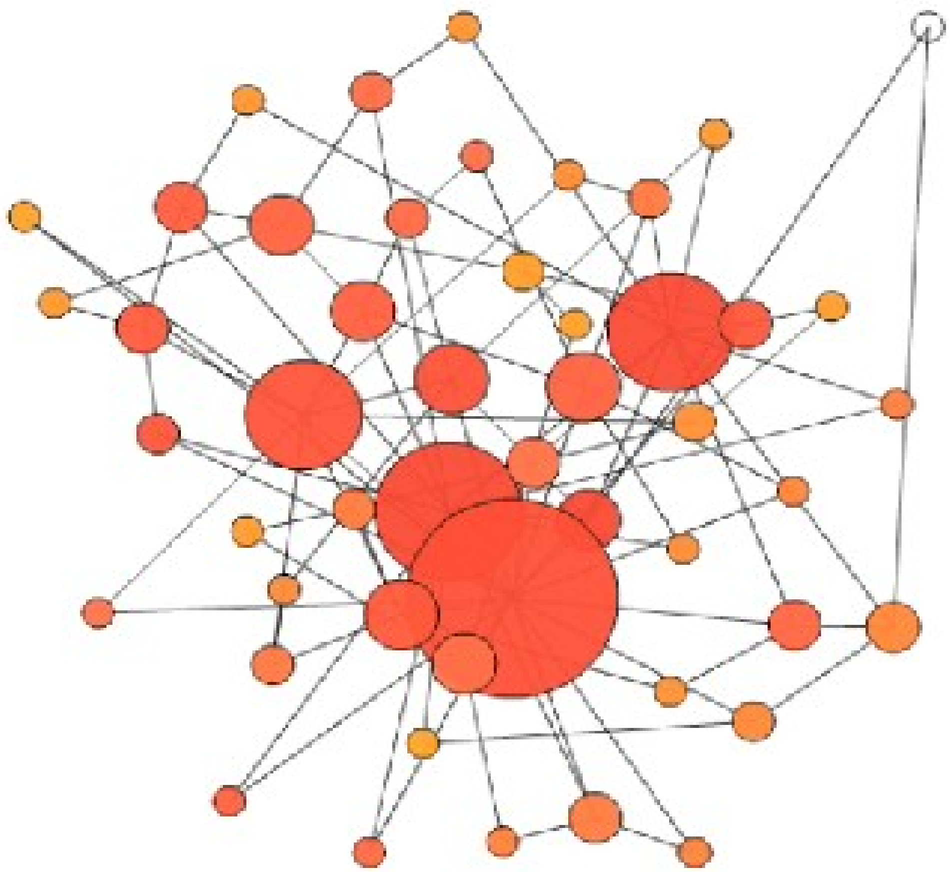
- Modelo para redes aleatórias baseado em PA (Barabasi, Albert 99, +32000 citações)
  - objetos são vértices e popularidade é dada pelo grau do vértice
- Processo de formação
  - inicialmente temos pequeno clique
  - a cada passo, adicionar 1 vértice com grau m
  - vértice adjacente é escolhido aleatoriamente, com prob. proporcional ao seu grau

PA Exemplo



- Clique inicial com 3 vértices, m = 2
- Tamanho do vértice proporcional ao grau

PA Exemplo



- O que está acontecendo?
- "*Rich gets richer*" → Lei de potência?

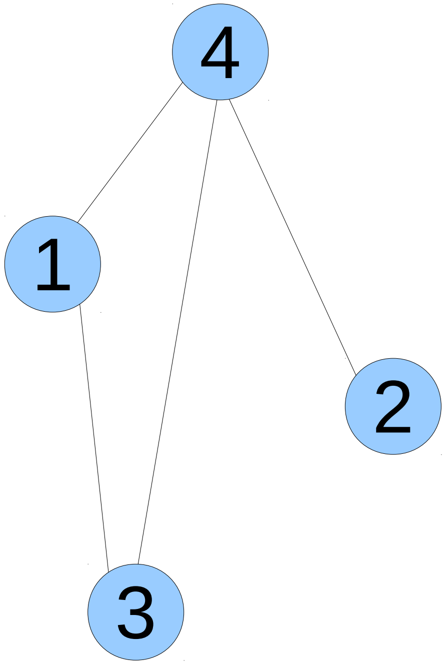
Definindo Preferências

- Dado grafo G=(V, E) no instante t = 1, 2, ...
- $d_u(t)$: grau do vértice u no instante t
- Considere a chegada de um novo vértice no instante t
  - vértice traz m arestas
- $p_u(t)$: prob. do vértice u ser incidente a uma nova aresta no instante t

$$p_u(t) = \frac{d_u(t)}{\sum_{v \in V} d_v(t)} = \frac{d_u(t)}{2mt}$$

número de arestas do grafo em t

PA Exemplo



$$p_u(t) = \frac{d_u(t)}{\sum_{v \in V} d_v(t)}$$

$$\sum_{v \in V} d_v = 2+1+2+3 = 8$$

$$p_1 = 2/8 = 1/4$$

$$p_2 = 1/8$$

$$p_3 = 2/8 = 1/4$$

$$p_4 = 3/8$$

Evolução do Grau do Vértice

- Quanto vale $d_u(t)$?
  - $d_u(t)$ é uma v.a. discreta
- Iremos fazer uma aproximação
  - Assumir $d_u(t)$ é determinístico (valor esperado)
  - Logo $d_u(t)$ é contínuo
- Quanto vale variação de $d_u(t)$ no instante t ?

$$\frac{dd_u}{d_t} = m\, p_u(t) \longrightarrow \frac{dd_u}{d_t} = \frac{d_u}{2t}$$

Evolução do Grau do Vértice

- Condições inicias
  - Vértice u chega no instante t_u
  - Ao chegar, vértice tem m arestas: $d_u(t_u) = m$
- Resolver diferencial

$$\frac{dd_u}{d_t} = \frac{d_u}{2t} \longrightarrow d_u(t) = m\left(\frac{t}{t_u}\right)^{1/2}$$

- Dependência temporal no grau
  - mais antigos tem grau maiores

Distribuição do Grau

- Distribuição do grau no tempo t
  - distribuição depende to tempo?

$$P[d_u(t) < k]$$

- Assumir que instante de chegada é uniforme
  - Vértices u entra entre [1, t]
  - Substituindo $d_u(t)$

$$P[d_u(t) < k] = P[t_u > m^2 t / k^2] = 1 - \frac{m^2}{k^2}$$

Distribuição do Grau

- Resolvendo e depois diferenciando temos

$$P[d_u(t)=k] \approx \frac{2m^2}{k^3}$$

Lei de potência na distribuição do grau

Observações sobre PA

- Fenômenos observados podem ser modelados por PA
  - modelo generativo baseado em PA
- Modelo pode levar ao surgimento de *lei de potência* na popularidade dos objetos
- PA como *explicação* para leis de potência observadas empiricamente
  - lei de potência surge, pois processo de formação obedece ao PA
- **Cuidado**: nem toda lei de potência é explicada por PA

Limitação do Modelo BA

- Lei de potência com expoente 3
  - muitas redes tem outros expoentes
- Vértices mais antigos tem grau maiores
  - vértices novos não tem chance de serem populares
- Vértices criam arestas somente no instante de chegada
  - não há evolução local da rede
- Rede gerada tem baixa clusterização