Pontifícia Universidade Católica de Minas Gerais
Departamento de Ciência da Computação
Redes Complexas

Redes Complexas II

Prof.: Felipe Domingos
felipe@pucminas.br

Centralidade

Como medir a *importância* de um vértice?

- Utilizando apenas a estrutura
- Relativo a outros vértices

- Grau
- Betweeness
- Closeness
- Autovetor
- Random walks
- etc.

- Métricas locais
  - dependem apenas da vizinhança do vértice (Ex. grau, random walk)
- Métrica globais
  - Dependem do grafo inteiro (Ex. Closeness, pagerank)

Centralidade de Grau

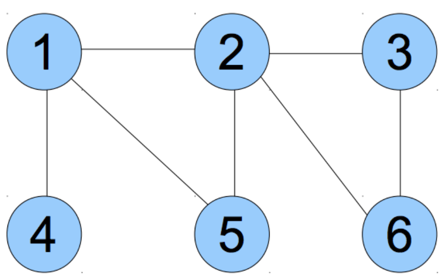
- Grau do vértice ou grau do vértice normalizado
  - valor entre 0 e 1

$$C_v = \frac{d_v}{n-1}$$

- Grafo direcionado, grau de entrada/saída
  - Duas centralidades em grau por vértice

Centralidade de Betweeness

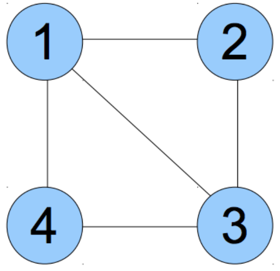
- Mede o quanto no “meio do caminho” um vértice está
- Considerar todos os caminhos mínimos do grafo
- Número de caminhos mínimos que passam pelo vértice
- Exemplo



- Grafo completo, K_n?
- Grafo estrela, com *n* folhas?

Centralidade de Betweeness

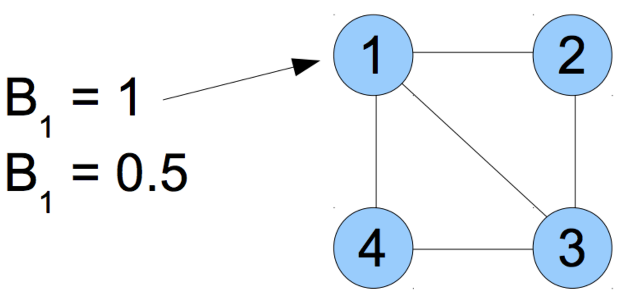
- **Problema:** Como definir métrica quando mais de um caminho mínimo existe entre um par origem/destino?
  - empate no custo do caminho mínimo
- Exemplo



- Caminho mínimo entre 2 e 4?
- 2,1,4 ou 2,3,4?
- Centralidade do vértice 1 e 3?

Centralidade de Betweeness

- Duas abordagens
  - Cada caminho mínimo conta 1 vez
  - “Carga” dividida pelos caminhos mínimos (cada caminho mínimo conta 1/k para a métrica, para k caminhos)
- Exemplo



- Para muitas redes, diferença é pequena
- Mas nem sempre!

Calculo Betweeness

- Mais precisamente

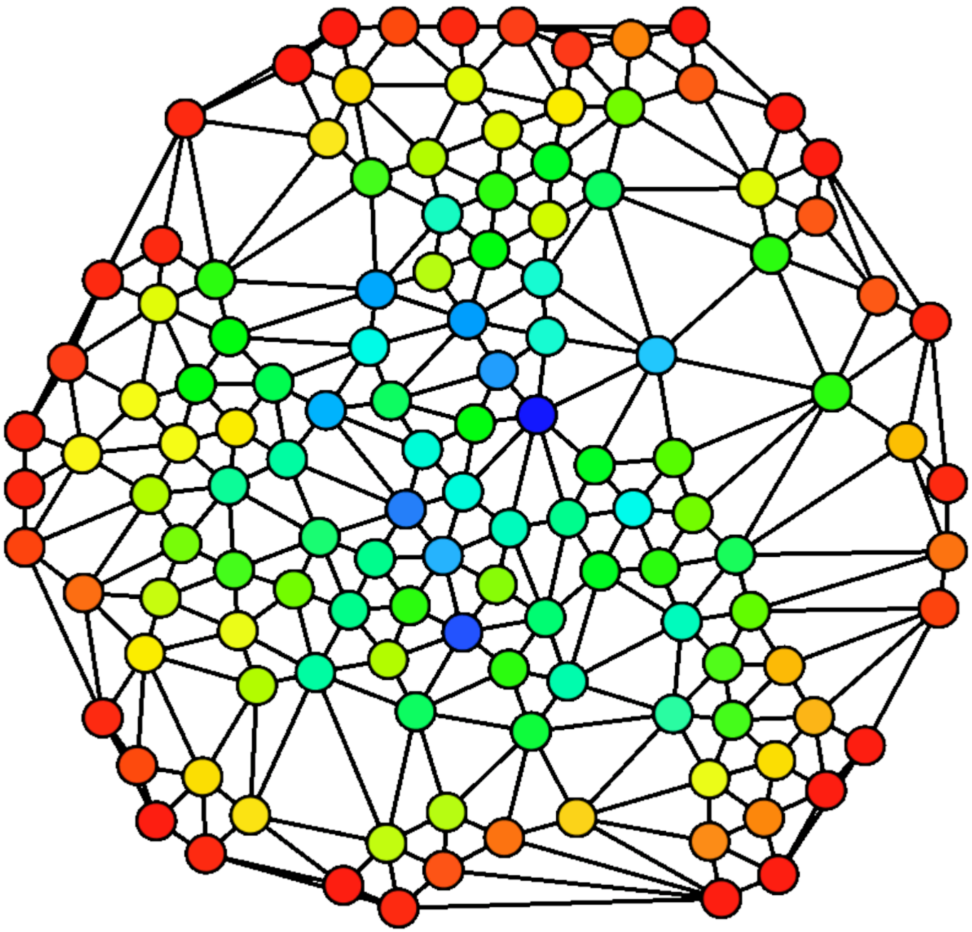
$$C_v = \sum_{s,t \in V;\, s,t \neq v} \frac{\sigma_v(s,t)}{\sigma(s,t)}$$

Número de caminhos mais curtos entre s e t que passam por v

Número de caminhos mais curtos entre s e t

- Pode ser normalizada pelo número total de pares origem/destino (sem contar v)
  - métrica entre 0 e 1

Exemplo



- Cores indicam betweeness
- Vermelho = 0, azul = máximo
- Ilustra vértices mais centrais

Centralidade de Closeness

- Utiliza conceito de distância
  - com ou sem pesos
- Distância média entre vértice e o resto do grafo
  - capturar o quão central é o vértice

$$C_v = \frac{\sum_{t \in V - \{v\}} d(v,t)}{n-1}$$

- "Velocidade" com a qual informação se propaga de um vértice para o resto da rede

Quatro importantes características

- Observada em diversas redes reais
  - a partir do final da década de 90

- Mundo pequeno

- Meus amigos são amigos

- Normalidade ausente

- Tudo Conectado

Mundo Pequeno

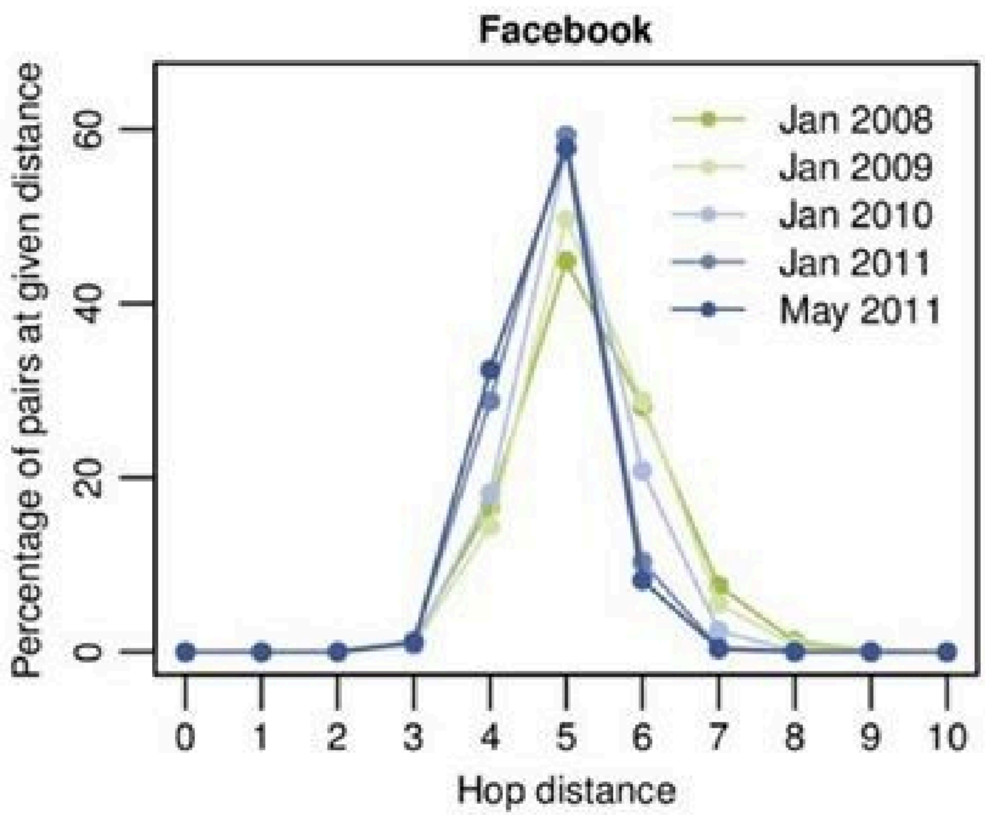
- “It's a small world, after all”
- Distância média muito pequena, diâmetro também
- Mesmo para redes muito grandes
  - Web (parte) – 10^8 vértices
  - Rede de colaboração – 10^6 nodes
  - Facebook - 10^9 nodes
  - “seis graus de separação”
  - e muitas outras!

ORDENS DE GRANDEZA MENOR!

APARENTEMENTE DA ORDEM DE LOG N

Exemplo: Facebook

- Distribuição da distância ao longo do tempo (truncada em 10)



vértices (arestas)

	fb
2007	13.0 M (644.6 M)
2008	56.0 M (2.1 G)
2009	139.1 M (6.2 G)
2010	332.3 M (18.8 G)
2011	562.4 M (47.5 G)
current	721.1 M (68.7 G)

- Rede cresce (50 x), mas distância média diminui!
- current = 12/2011

Exemplo: Facebook

- Mudanças na rede

distância média (std)

	fb
2007	4.46 (±0.04)
2008	5.28 (±0.03)
2009	5.26 (±0.03)
2010	5.06 (±0.01)
2011	4.81 (±0.04)
current	4.74 (±0.02)

grau médio

	fb
2007	99.50
2008	76.15
2009	88.68
2010	113.00
2011	169.03
current	190.44

densidade

	fb
2007	7.679E-06
2008	1.359E-06
2009	6.377E-07
2010	3.400E-07
2011	3.006E-07
current	2.641E-07

- Grau médio cresce
- Rede muito, muito esparsa
- Grau médio aumenta e densidade diminui?

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”

Mundo Pequeno

- 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

Mundo Pequeno

- 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!

Mundo Pequeno (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)

Mundo Pequeno

- 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

Meus amigos também são amigos

- A relacionado com B e C faz com que B e C se relacionem mais provavelmente
- Rede possui transitividade – caminhos de comprimento dois viram triângulos
  - métrica: coeficiente de clusterização
  - densidade: chance de dois vértices ao acaso estarem relacionados

	clusterização	densidade
AS graph – 10^4 nodes	**0.39**	**0.00056**
Facebook – 10^9 nodes	**0.14**	**0.00000026**
Biology coauthorship	**0.67**	**0.00001**

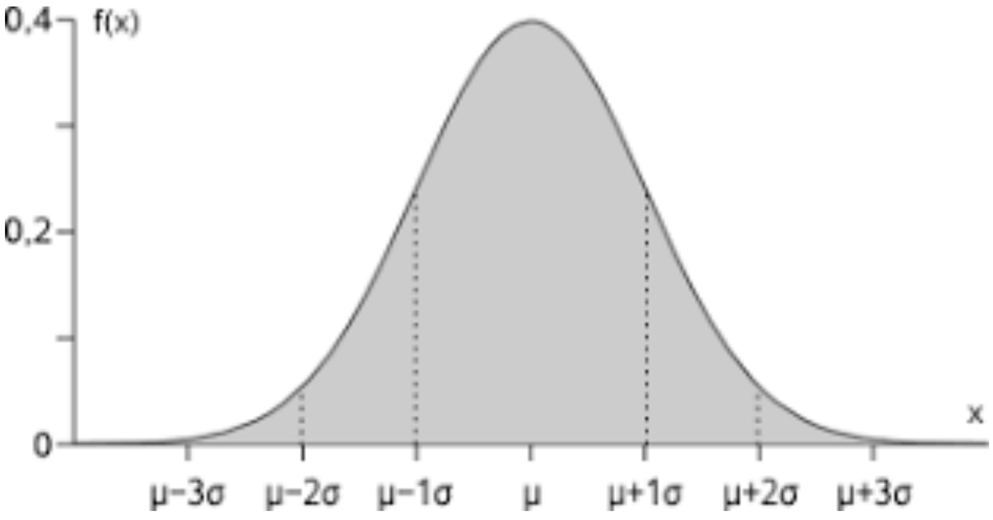
Normalidade Ausente

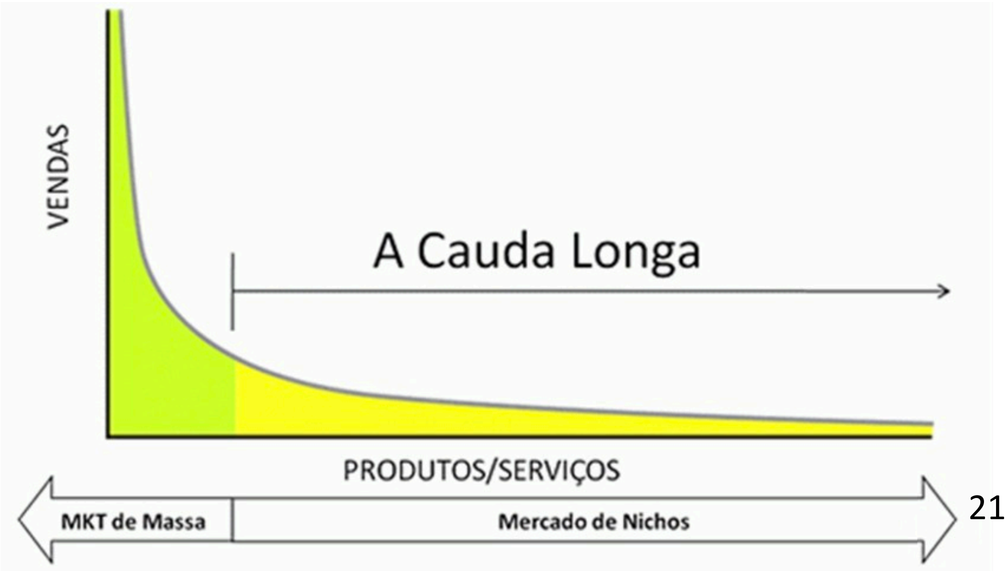
- Grau dos vértices é muito desigual
- Muitos com grau pequeno, poucos com grau muito grande

PARECIDO COM DISTRIBUIÇÃO DA RENDA NO BRASIL!

  - AS Graph – 10^4 nodes Grau médio: 5.9 Grau máx: ~2100
  - Citações – 10^6 nodes Grau médio: 8.6 Grau máx: ~9000
- Distribuição de grau possui cauda pesada
- Abrange diversas ordens de grandeza
- Muito diferente de distribuição normal

Distribuição Normal X Cauda Pesada





Tudo Conectado

- Maior componente conexa possui **quase todos** os vértices
  - CC gigante
- Outras componentes muito pequenas
  - Muitas outras componentes
- Rede de sinônimos – 23K vértices
  - Maior componente: ~**22K**
- Rede social, rede neural, etc

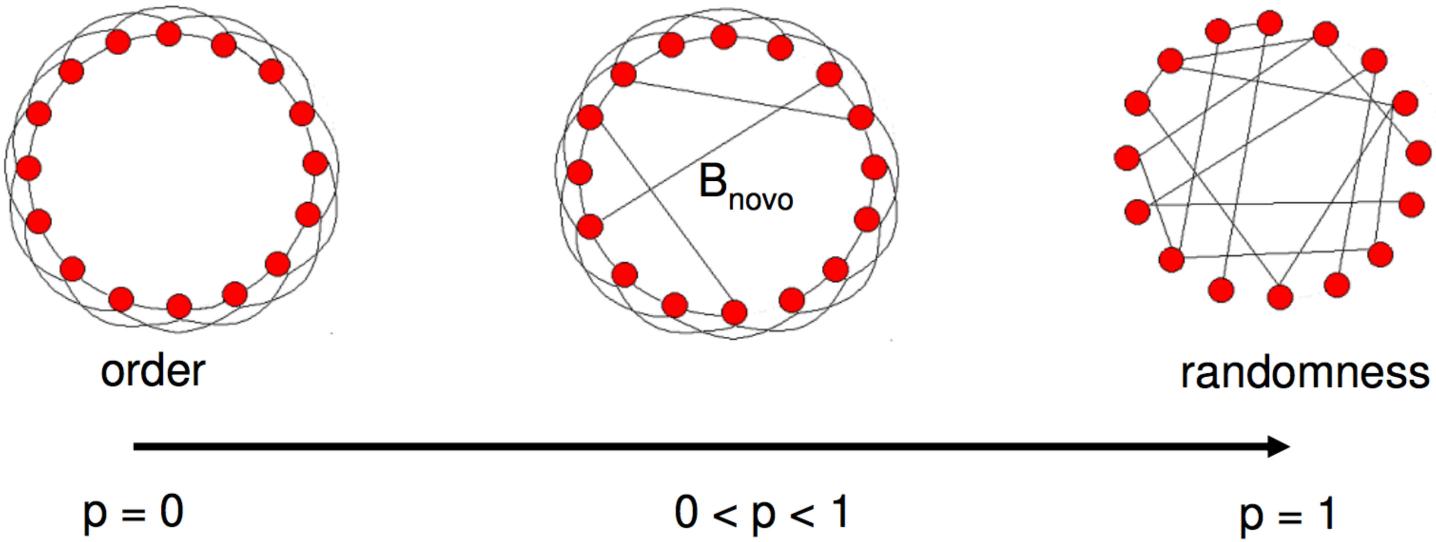
Modelo de Rede Small World

- 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

Modelo de Rede Small World

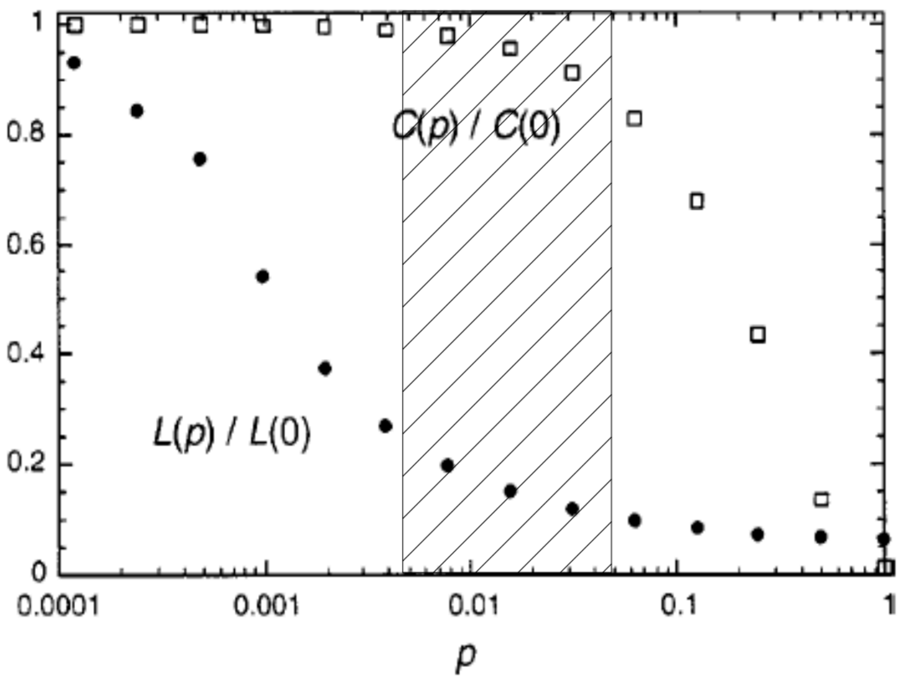
- Modelo possui 3 parâmetros N, k, p
- k controla grau e clusterização
- p controla aleatoriedade do grafo
  - p = 0, grafo látice regular
  - p = 1, grafo aleatório (idêntico ao G(n, p)?)
- Propriedades topológicas do grafo
  - distância média clusterização distribuição dos graus, etc

Modelo de Rede Small World



Modelo de Rede Small World

- 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

Modelo G(n,p)

- Modelo clássico para grafos aleatórios
- Proposto por Erdös e Rényi (1959)

Como funciona o modelo?

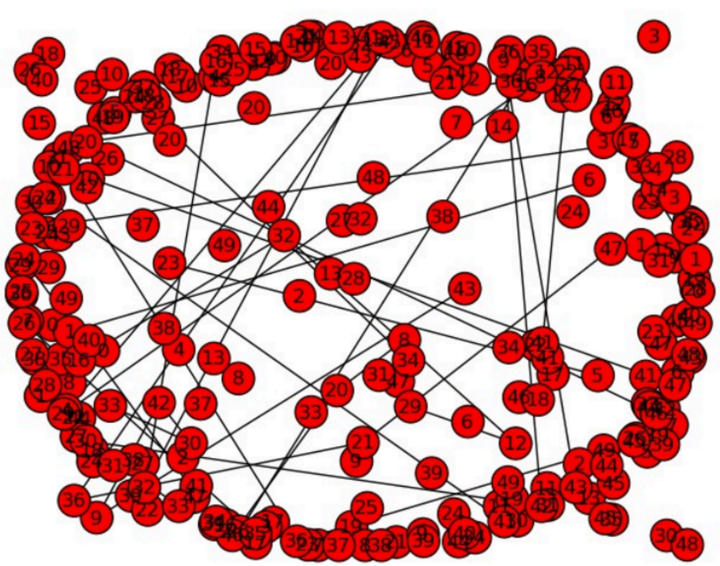
- Grafo possui ***n*** nós
- Cada aresta do grafo existe, independentemente, com probabilidade ***p***

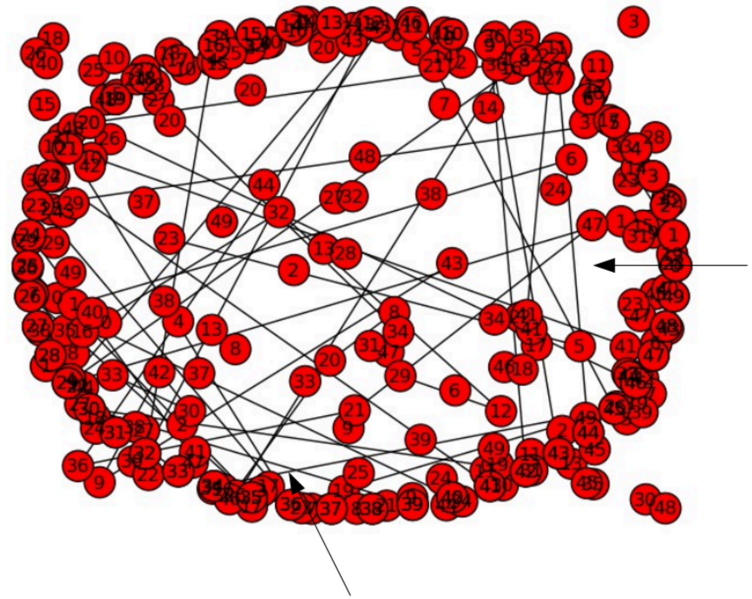
Modelo G(n,p)

- Modelo possui dois parâmetros determinísticos
  - n: número de vértices
  - p: prob. de existência de cada aresta

- Para um par de valores de (n,p), diversos grafos podem ser gerados

Modelo G(n,p)





n = 50; p = 0.01

Modelo G(n,p)

- A estrutura esperada de um grafo aleatório G(n,p) varia com o valor de **p**
- Arestas conectam nós formando componentes, i.e. subconjuntos de nós alcançáveis por caminhos pela rede

- **TED Talk:** The hidden influence of social networks

  https://www.ted.com/talks/nicholas_christakis_the_hidden_influence_of_social_networks