

Redes Complexas: Internet, Web e outras aplicações em computação

-Bloco #2-

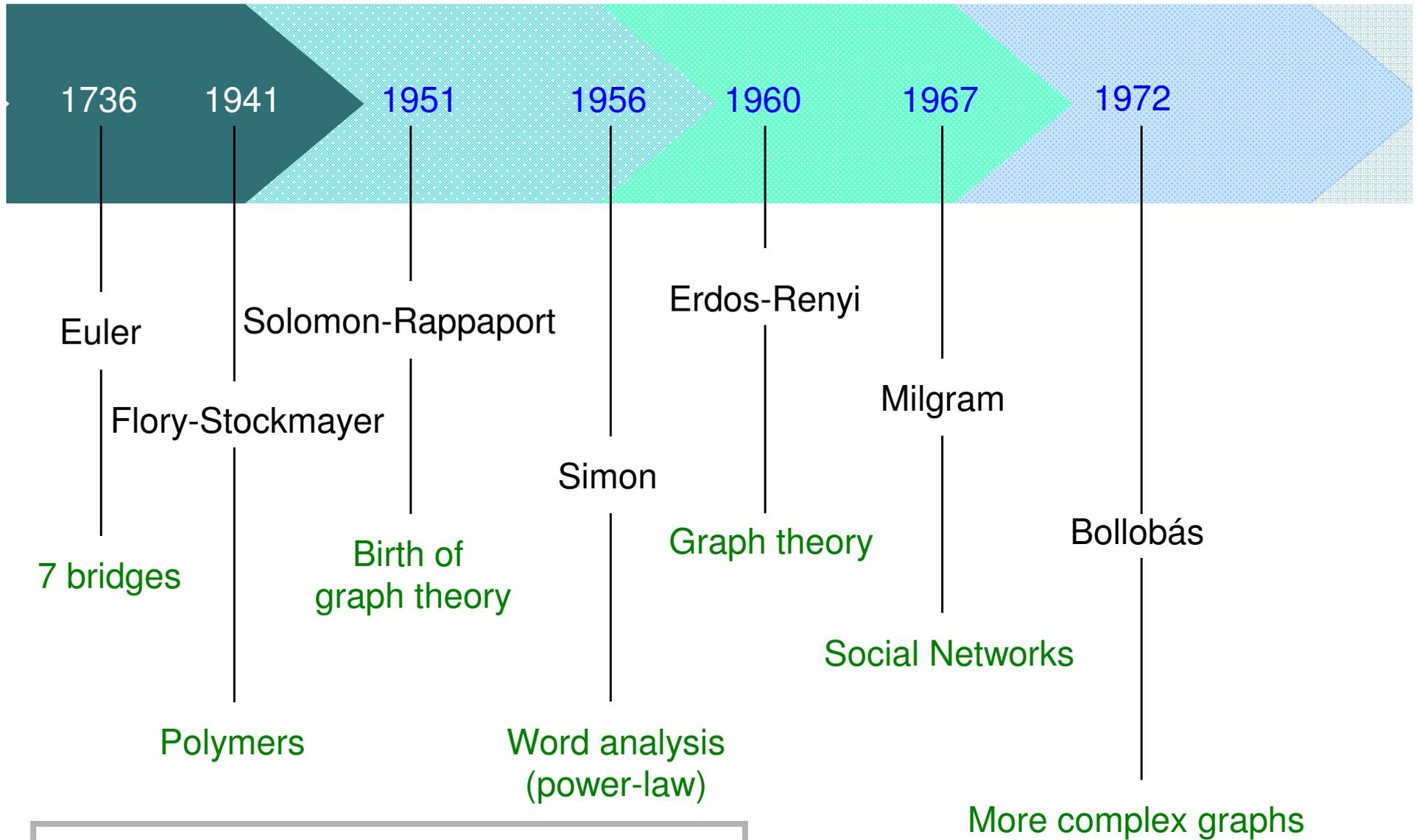
Introdução a Redes
2º semestre de 2008

Virgílio A. F. Almeida
Agosto de 2008



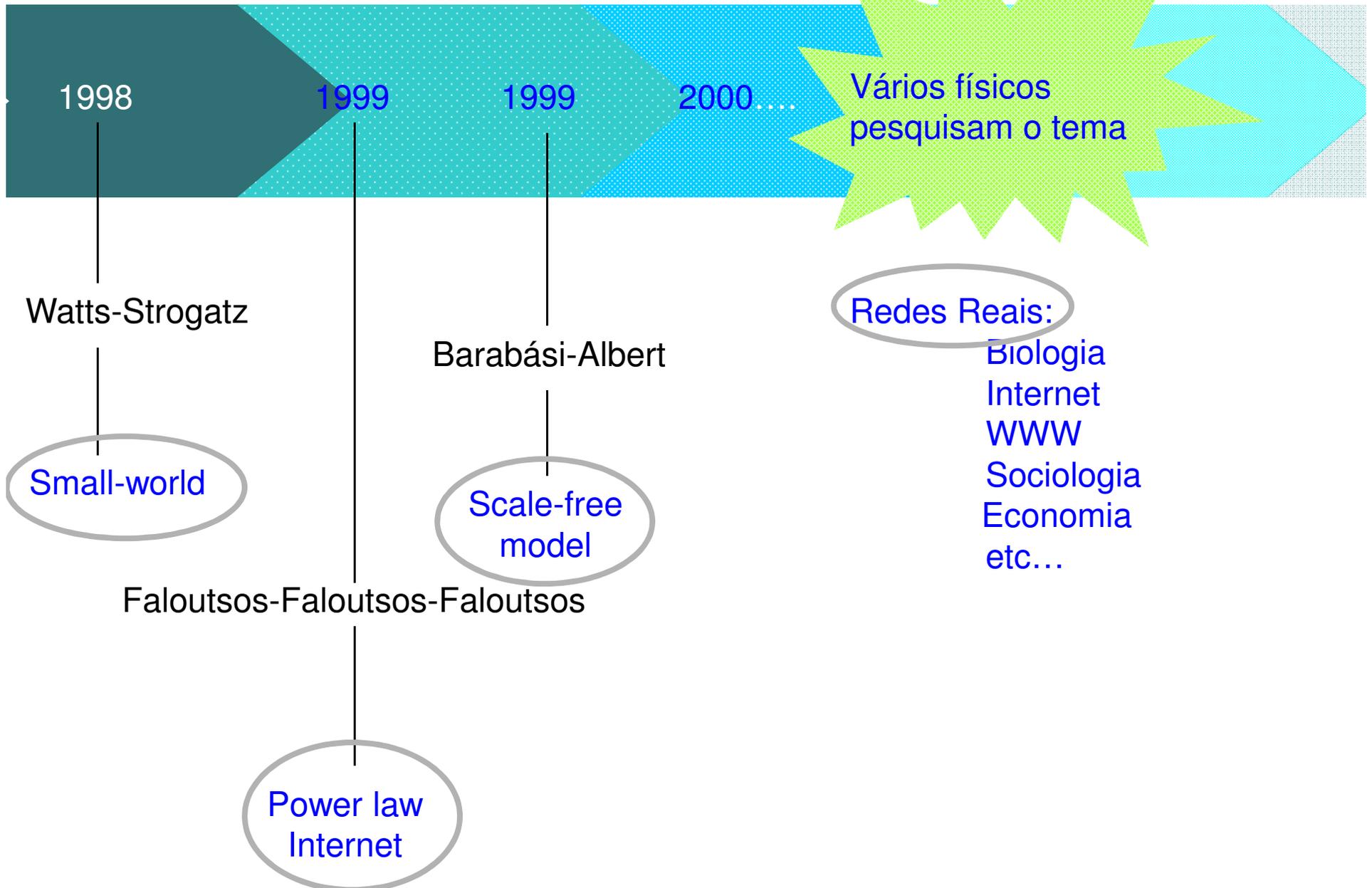
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

Linha do Tempo em Redes*



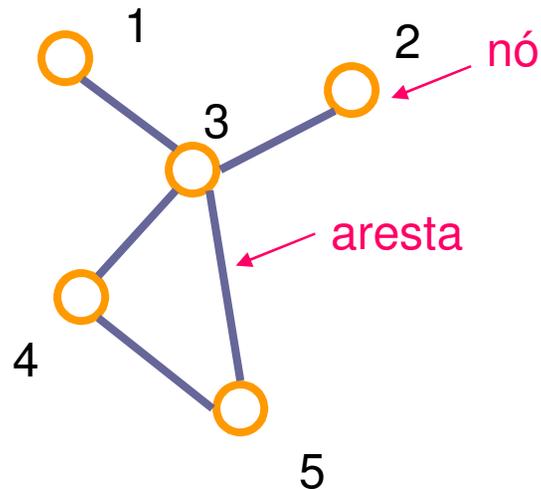
Preparado por: J.F.F. Mendes, Univ. Aveiro

Linha do Tempo em Redes



O que são as redes?

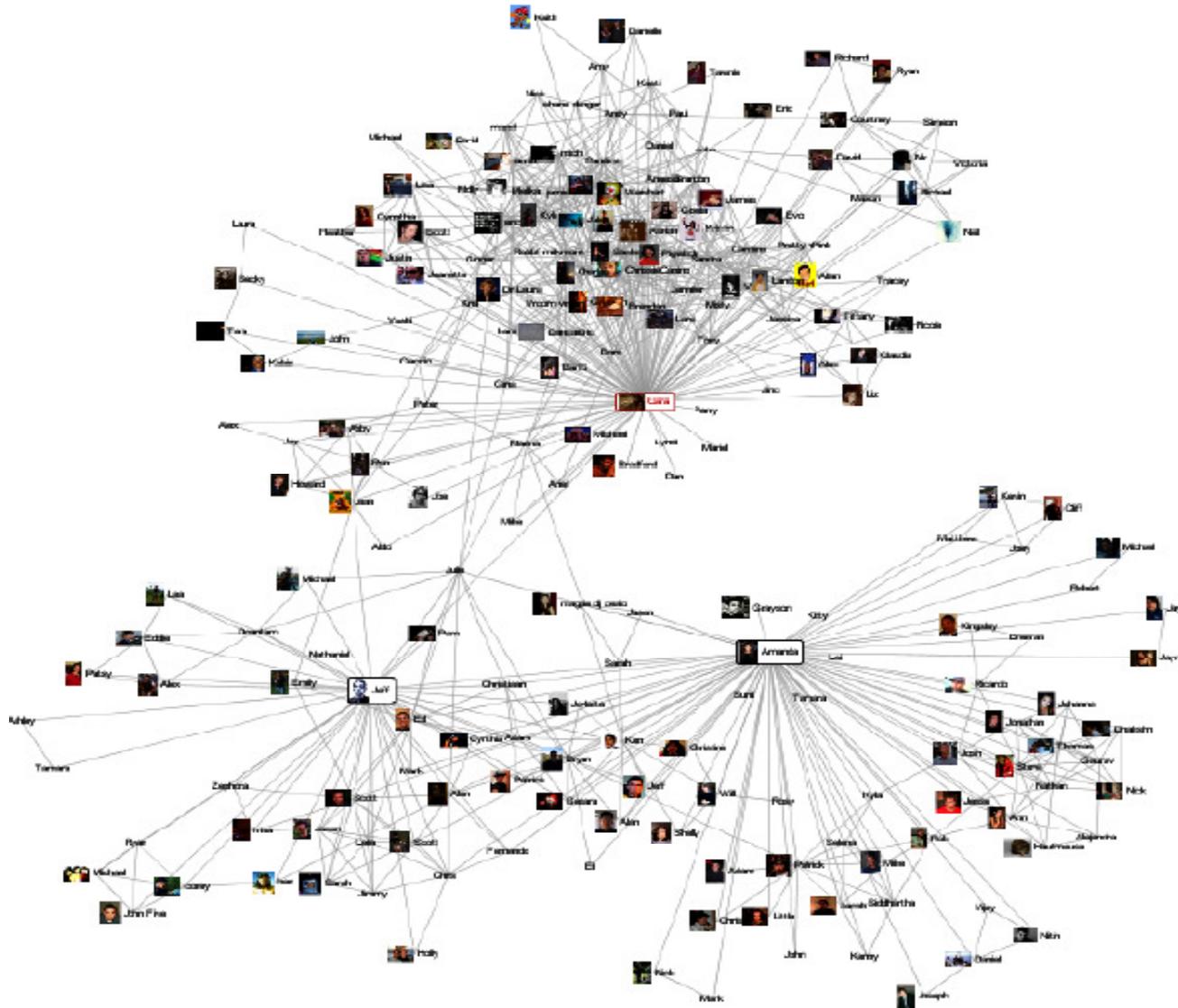
- Redes são coleções de pontos e linhas.



“Rede” \equiv “Grafo”

pontos	linhas	
vertices	Arcos, arestas	matematica
nós	Links, arestas	CC
atores	ligações, relações	sociologia

Rede Social Online: Friendster



"Vizster: Visualizing Online Social Networks." Jeffrey Heer and danah boyd. *IEEE Symposium on Information Visualization (InfoViz 2005)*.

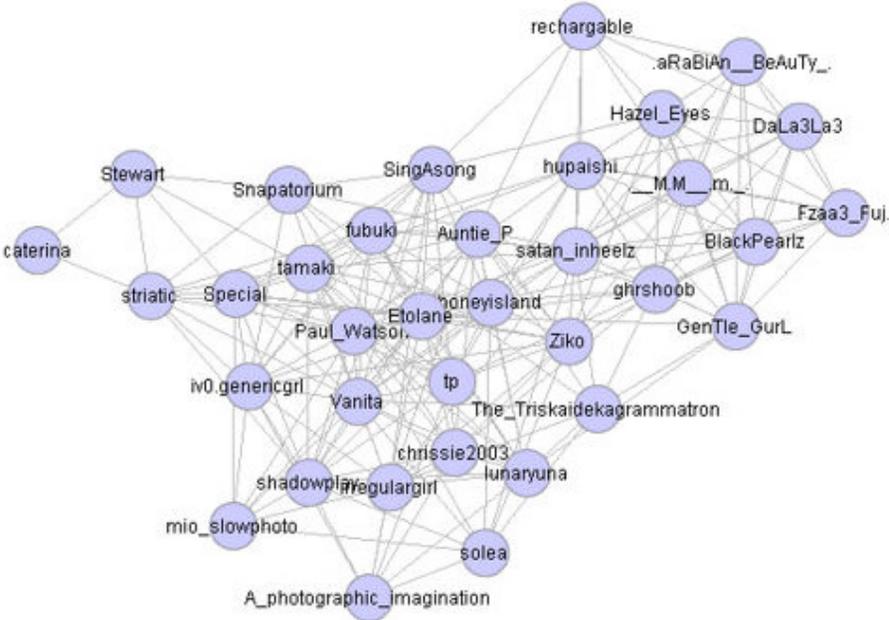
Rede Social Online: Flickr



Home The Tour Sign Up Explore |

Flickr

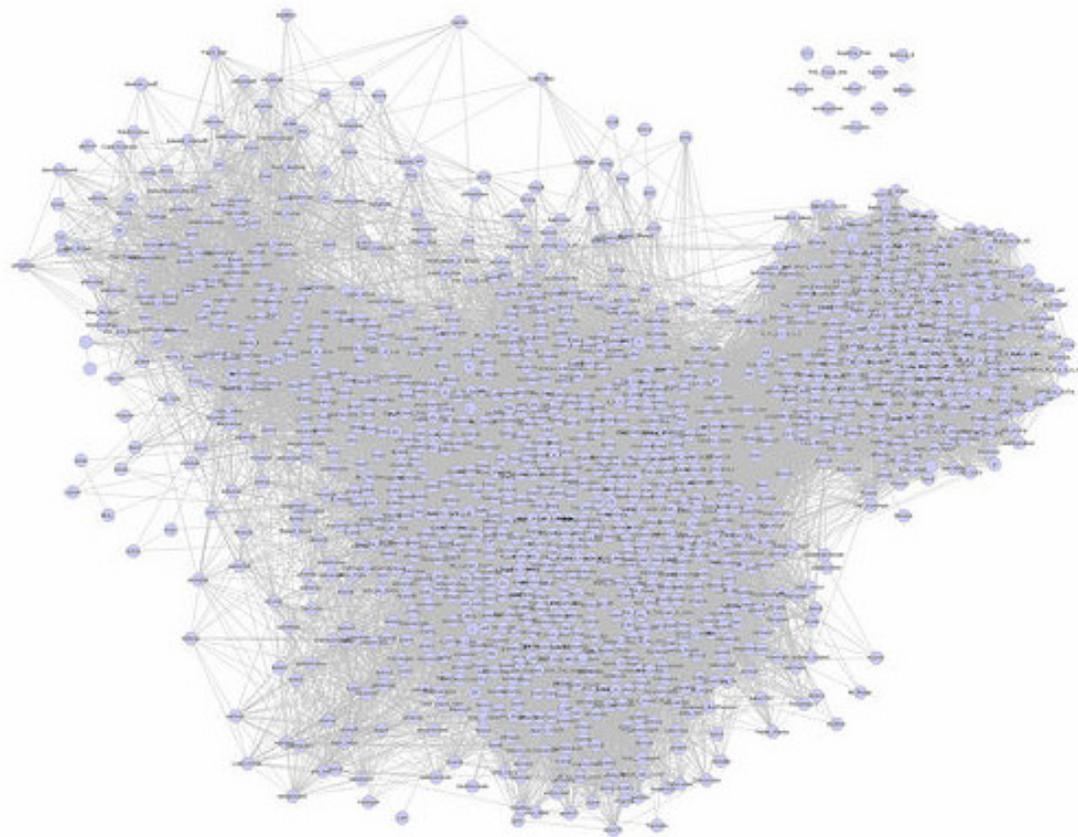
mc-300



The center of the network, including only people with at least 300 mutual contacts (35

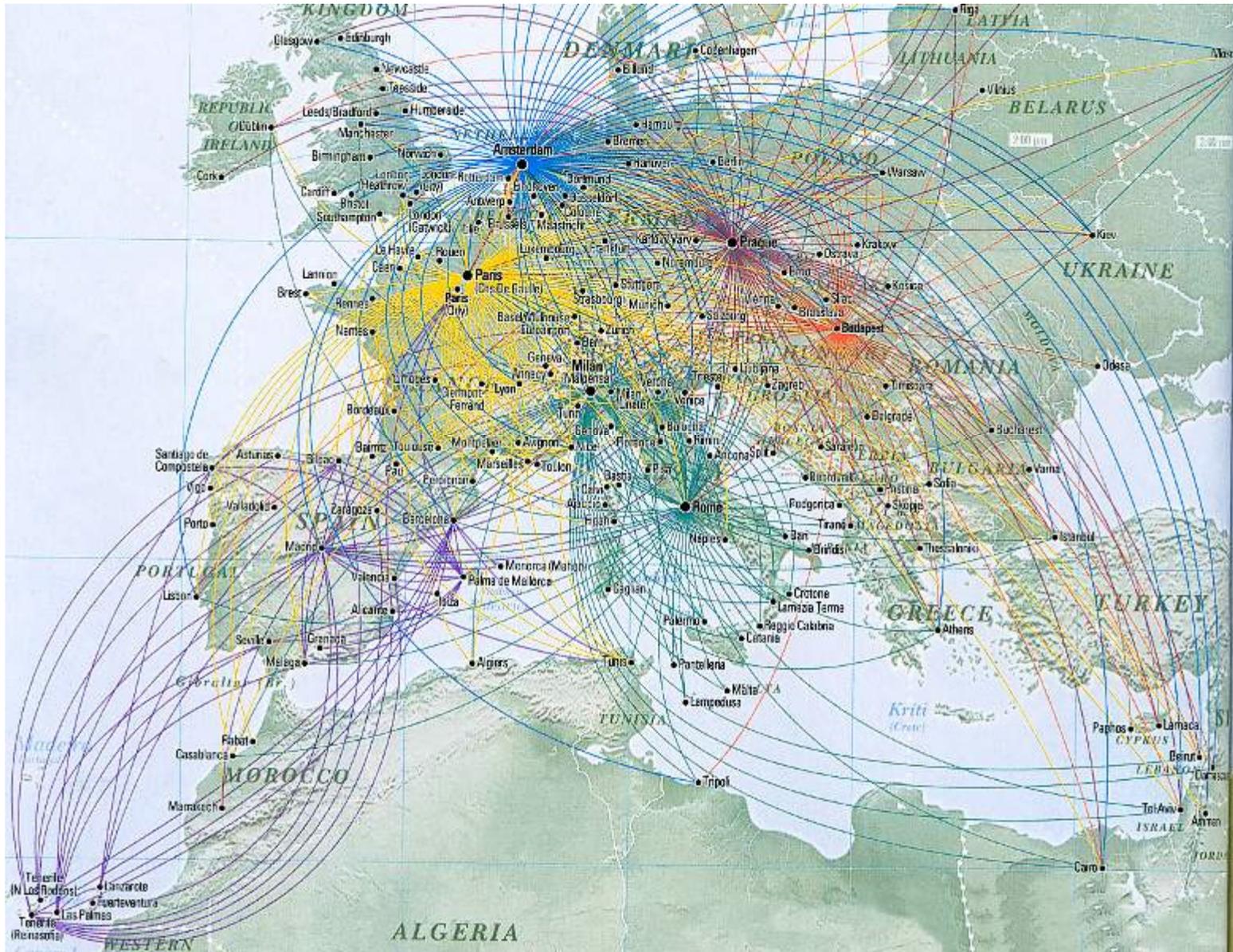
Rede Social Online: Flickr

mc-50



The network, including only people with at least 50 mutual contacts (1211 people). Three of the clusters start to merge, but the UAE cluster remains distinct. At this level, there are eleven individuals representing small, disconnected networks.

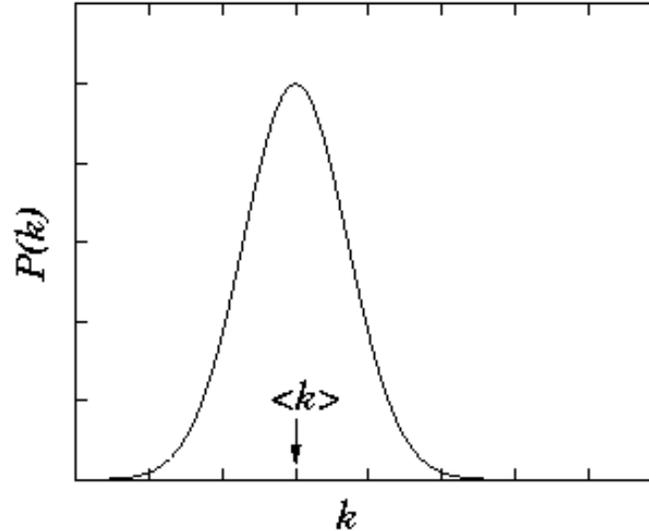
Rede de transporte: linhas aéreas



Redes Randômicas

- Nós conectados aleatoriamente
- Número de arestas incidentes em cada nó segue uma distribuição de Poisson.
- Grafo randômico é uma instaciação de uma família de grafos randômicos.

Poisson distribution



Tópicos

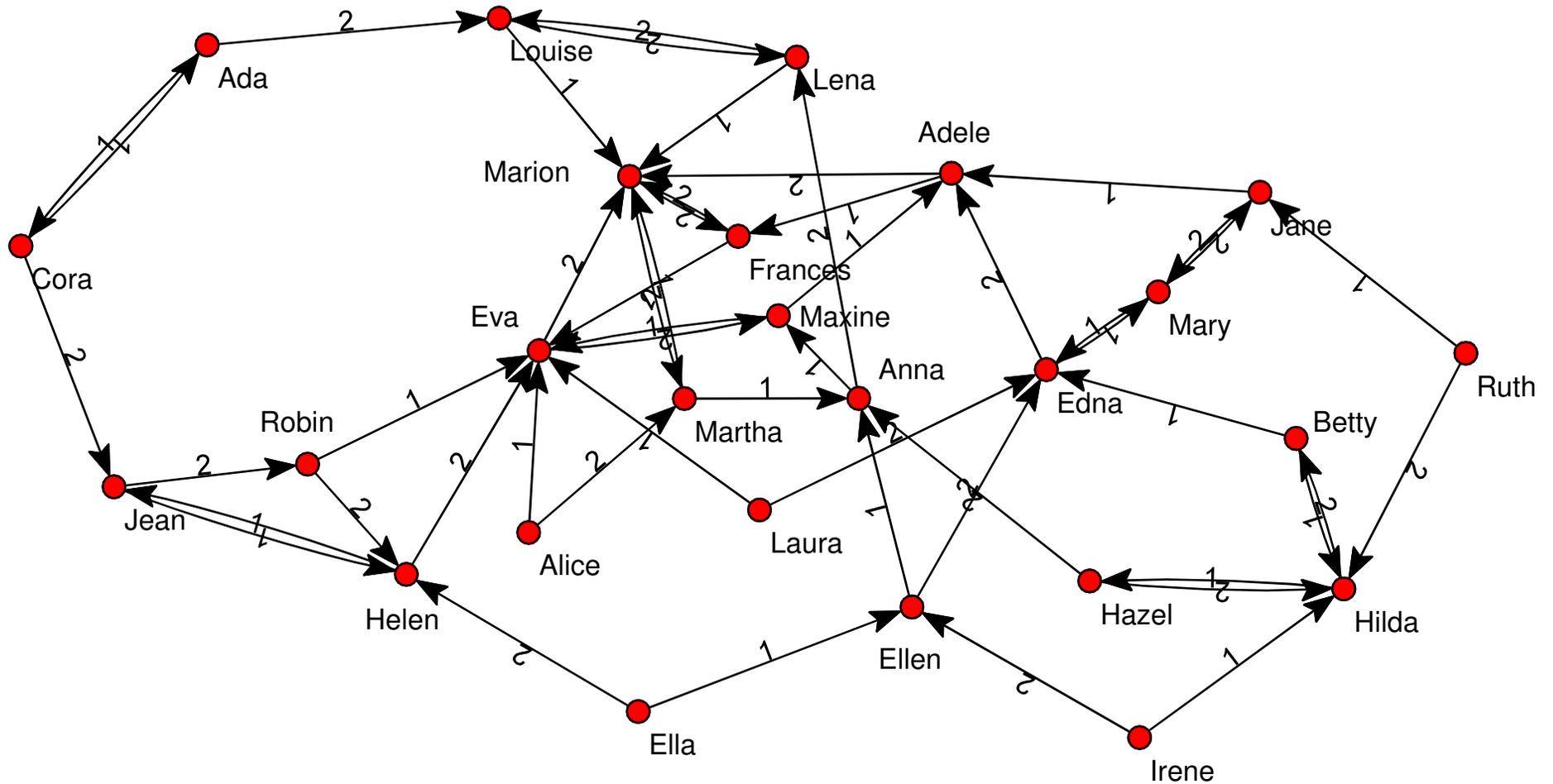
- Redes e Grafos
- Métricas Básicas de Redes
- Grafos Bipartites
- Pajek

Elementos da Redes (Networks) : arestas

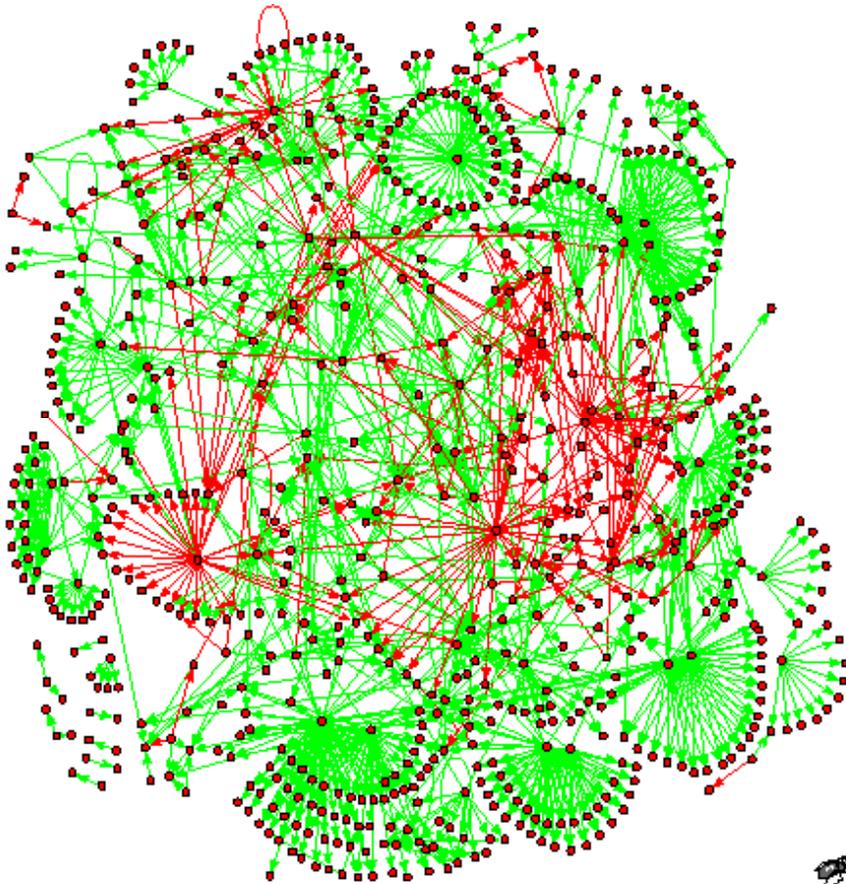
- Orientado
 - $A \rightarrow B$
- Não orientado
 - $A \leftrightarrow B$ or $A - B$
- Atributos das arestas
 - Ponderação (ex. Frequencia de conexões entre 2 sites)
 - Ranking-ordenamento: no roteador a ordem dos links de saída
 - tipo (ex. em sociologia; amigo, parente, co-autor)
 - Propiedades que dependem da estrutura do resto grafo, ex.: *betweenness*

Redes orientadas

- Distribuição de amigas de jantar *The sociometry reader, 1960)*
- Primeira e segunda escolha. Exemplo em CC???



Arestas ponderadas podem valores positivos/negativos



- Ex. Um processo em um nó ativa ou inibe outro nó, um gene ativa ou inibe outro
- Uma pessoa confia/desconfia de outra, uma máquina confia ou não confia em outra (como?)
 - Pergunta de pesquisa: Como um nó 'propaga' percepções negativas em uma rede? Por exemplo, em uma social, o inimigo de meu inimigo é meu amigo? Quais aplicações em CC?

Grafos Ponderados (Weighted Graphs)

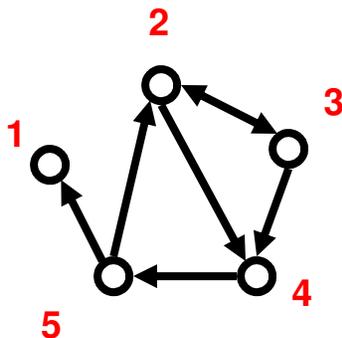
- Cada vértice é anotado por um *weight ed capacity*
- Orientado ou não-orientado
- Usado para modelar
 - Custo de transmissão, latência
 - Capacidade do link
 - hubs and authorities (Google PageRank algorithm)

Matriz de Adjacência

- Representação de arestas como matrix
 - $A_{ij} = 1$ se nó i tem uma aresta para j
 $= 0$ if nó i não tem aresta para j
 - $A_{ii} = 0$ a não ser no caso de *self-loops*
 - $A_{ij} = A_{ji}$ se a rede é não-orientada, ou se i e j compartilham arestas recíprocas



Exemplos:

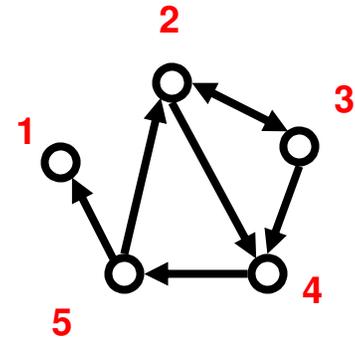


$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Lista de Adjacência

■ Lista de arestas

- 2 3
- 2 4
- 3 2
- 3 4
- 4 5
- 5 2
- 5 1



■ Lista de Adjacência

- Mais adequado para redes grandes e esparsas
- Fácil recuperação de nós vizinhos
 - 1:
 - 2: 3 4
 - 3: 2 4
 - 4: 5
 - 5: 1 2

Nós

■ Propriedades dos nós da rede

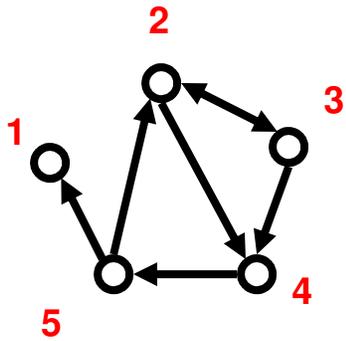
■ Das conexões imediatas (locais)

- indegree
quantas arestas orientadas que chegam a um nó
- outdegree
quantas arestas orientadas originam em um nó
- Grau: Degree (in ou out):
número de arestas que incidem em um nó



■ Do grafo inteiro (globais)

- centrality (betweenness, closeness)



Grau de Nós da Matrix de Valores

■ Outdegree = $\sum_{j=1}^n A_{ij}$

exemplo: o outdegree para nó 3 é 2, certo?

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\sum_{j=1}^n A_{3j}$$

■ Indegree = $\sum_{i=1}^n A_{ij}$

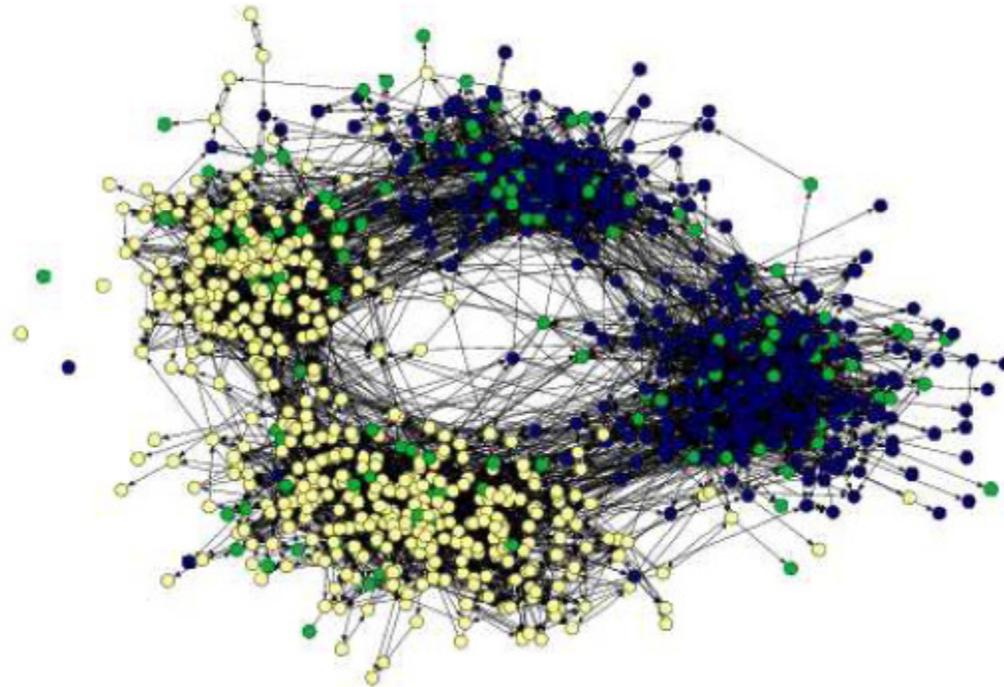
exemplo: o indegree para nó 3 é 1,

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\sum_{i=1}^n A_{i3}$$

Outros atributos de nós

- Faça a escolha.....
 - Localização geográfica
 - função
 - Gosto musical...



Métricas das Redes: sequência dos graus e distribuição dos graus

- Sequência de graus: Uma lista ordenada dos (in,out) graus de cada nó

- Sequência In-degree:

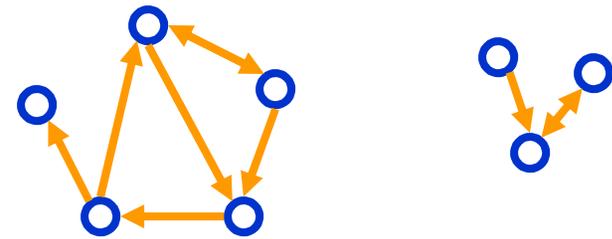
- [2, 2, 2, 1, 1, 1, 0]

- Sequência Out-degree:

- [2, 2, 2, 2, 1, 1, 1, 0]

- Sequencia de graus não-orientados

- [3, 3, 3, 2, 2, 1, 1, 1]



- Distribuição dos graus: contagem de frequência da ocorrência de cada grau

- Distribuição de In-degree:

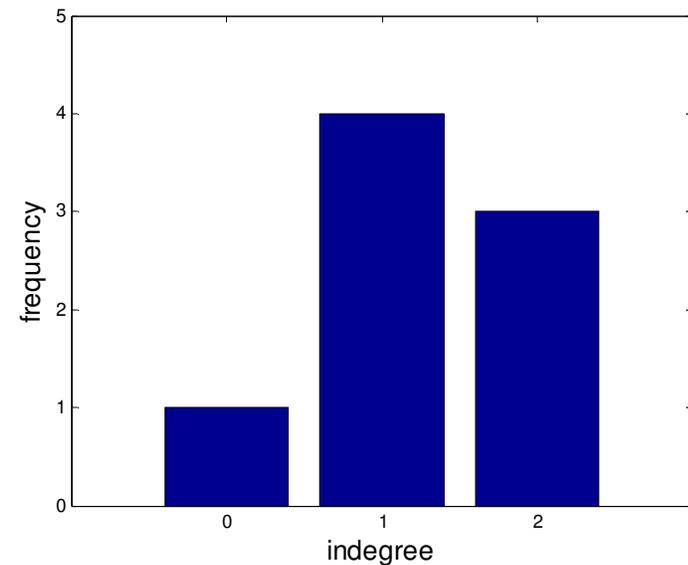
- [(2,3) (1,4) (0,1)]

- Distribuição de Out-degree:

- [(2,4) (1,3) (0,1)]

- Distribuição não orientada:

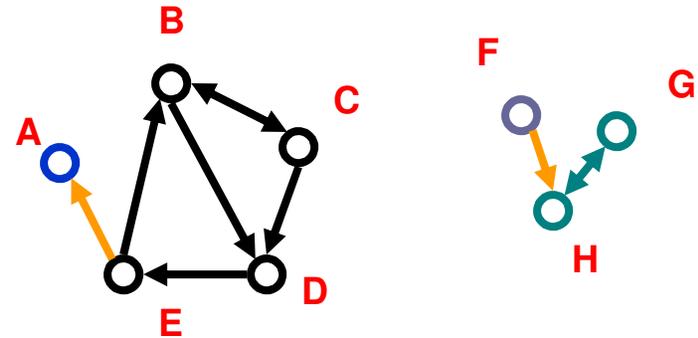
- [(3,3) (2,2) (1,3)]



Métricas das Redes: componentes conectados

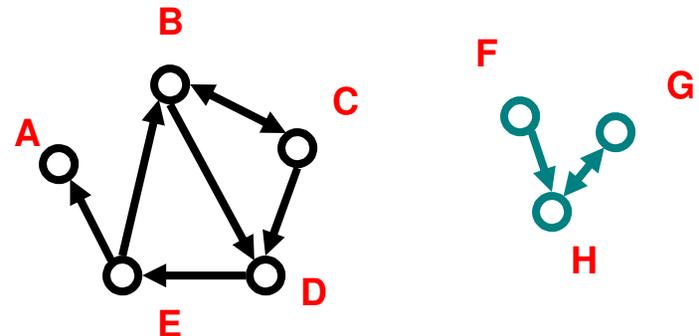
- Componentes fortemente conectados: Strongly connected components (SCC): cada nó dentro do componente pode ser alcançado de outro nó do componente seguindo arestas orientadas.

- SCC
 - BCDE
 - A
 - GH
 - F



- Componentes fracamente conectados (Weakly connected components WCC): cada nó pode ser alcançado a partir de qualquer outro nó seguindo arestas em qualquer direção.

- WCC
 - ABCDE
 - GHF

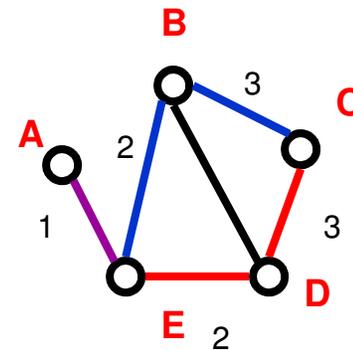


- Em redes não orientadas, simplesmente refere-se a componentes conectados

Métricas das Redes: caminho mínimo - shortest paths

- Caminho mínimo: a menor sequência de arestas conectando dois nós.
- Nem sempre única

- A e C são conectados por 2 shortest paths
 - A - E - B - C
 - A - E - D - C



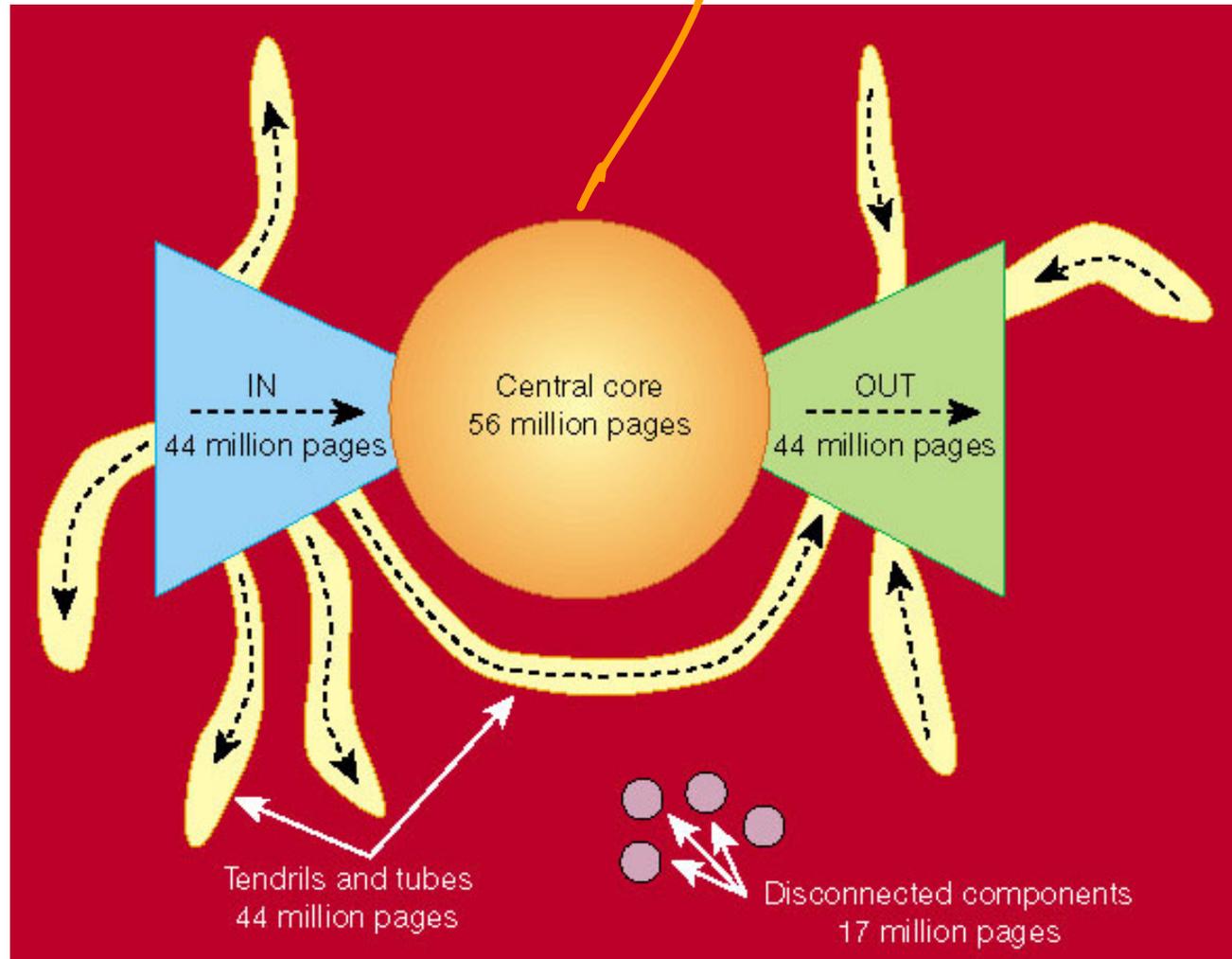
- Diametro: a maior distância geodésica no grafo

- A distância entre A e C é o máximo para o grafo: 3

- Obs: o termo 'diametro' é às vezes usado como sendo a distância do caminho mínimo médio. Wikipedia: the **distance** between two [vertices](#) in a [graph](#) is the number of edges in a shortest path connecting them. This is also known as the **geodesic distance**.

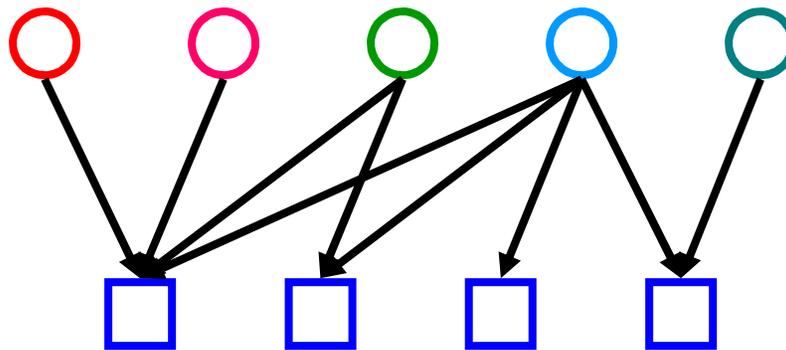
Componentes Gigantes e o grafo da Web

- Se o maior componente engloba uma fração significativa de um grafo, então é chamado “the **giant component**”



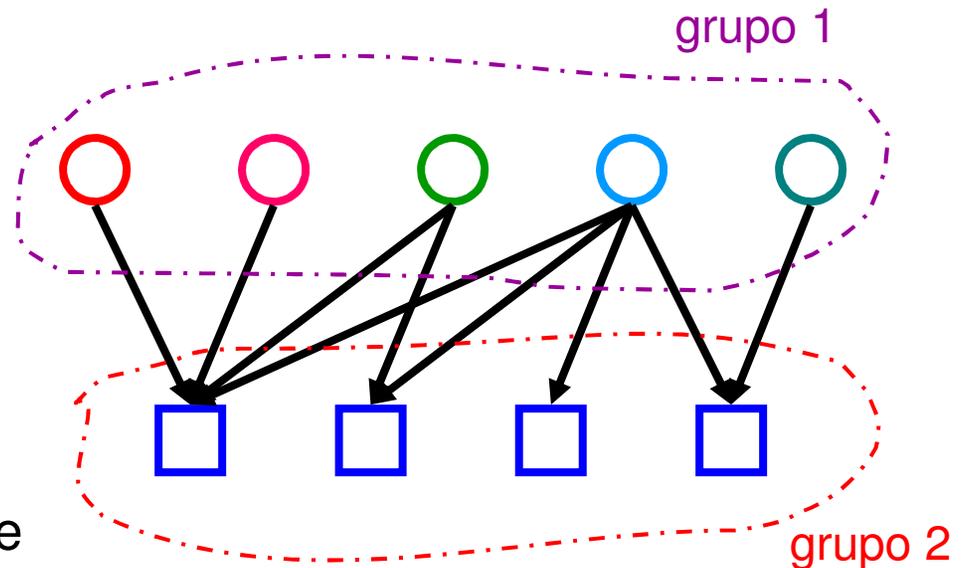
bipartite (two-mode) networks: grafos bipartidos

- arestas ocorrem somente entre dois grupos de nós, mas não dentro desses grupos
- exemplo, pode-se ter indivíduos e *eventos*
 - directores e conselhos diretores
 - Consumidores e seus itens de compra
 - *Spammers* e *mailers* legítimos



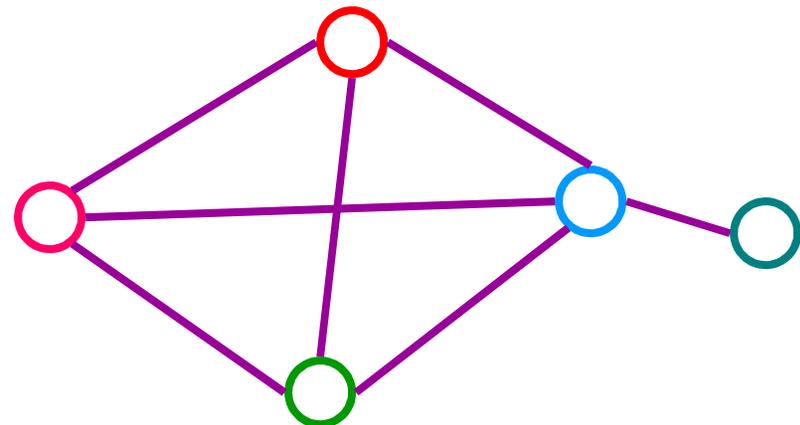
Mapeamento de um grafo bipartido (duas modalidades) em um grafo de uma modalidade

- Rede duas-modalidades



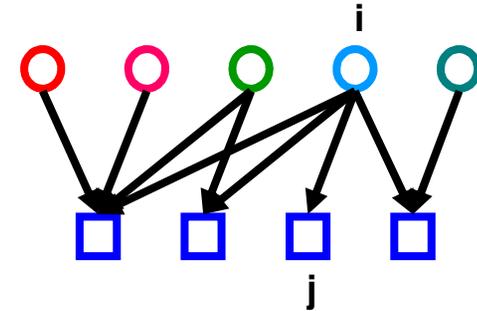
- Projeção em uma modalidade

- Dois nodos do primeiro grupo são conectados se eles ligam aos mesmos nós no segundo grupo.
- Perda de informação
- Ocorrência de cliques



Agora em notação matricial

- B_{ij}
 - = 1 se nó i do primeiro grupo liga ao nó j do segundo grupo
 - = 0 otherwise

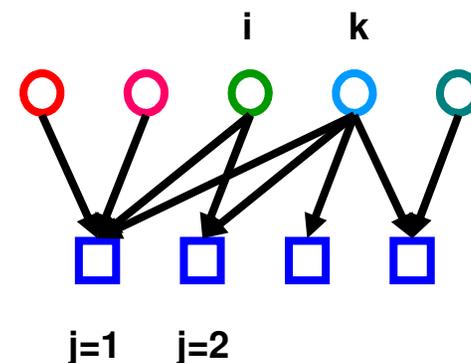


- B não é usualmente uma matrix quadrada
 - Ex. : n consumidores e m productos

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Gerando o grafo de uma-modalidade

- i e k estão conectados se ambos conectam a j
- $A_{ik} = \sum_j B_{ij} B_{kj}$
- $A = B B^T$
- A matriz transposta troca B_{xy} e B_{yx}
- Se B é uma matrix $n \times m$, B^T é uma matrix $m \times n$



$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

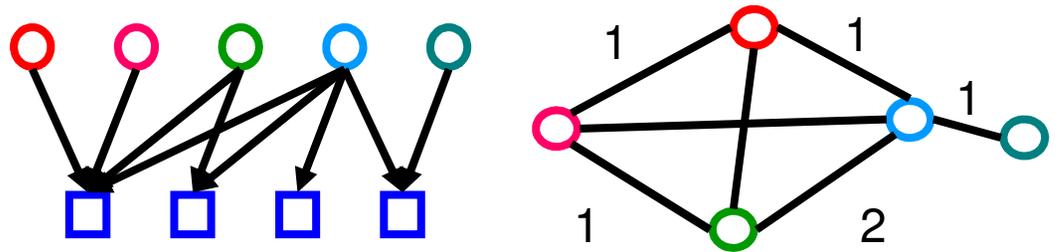
$$B^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Multiplicação de matrizes

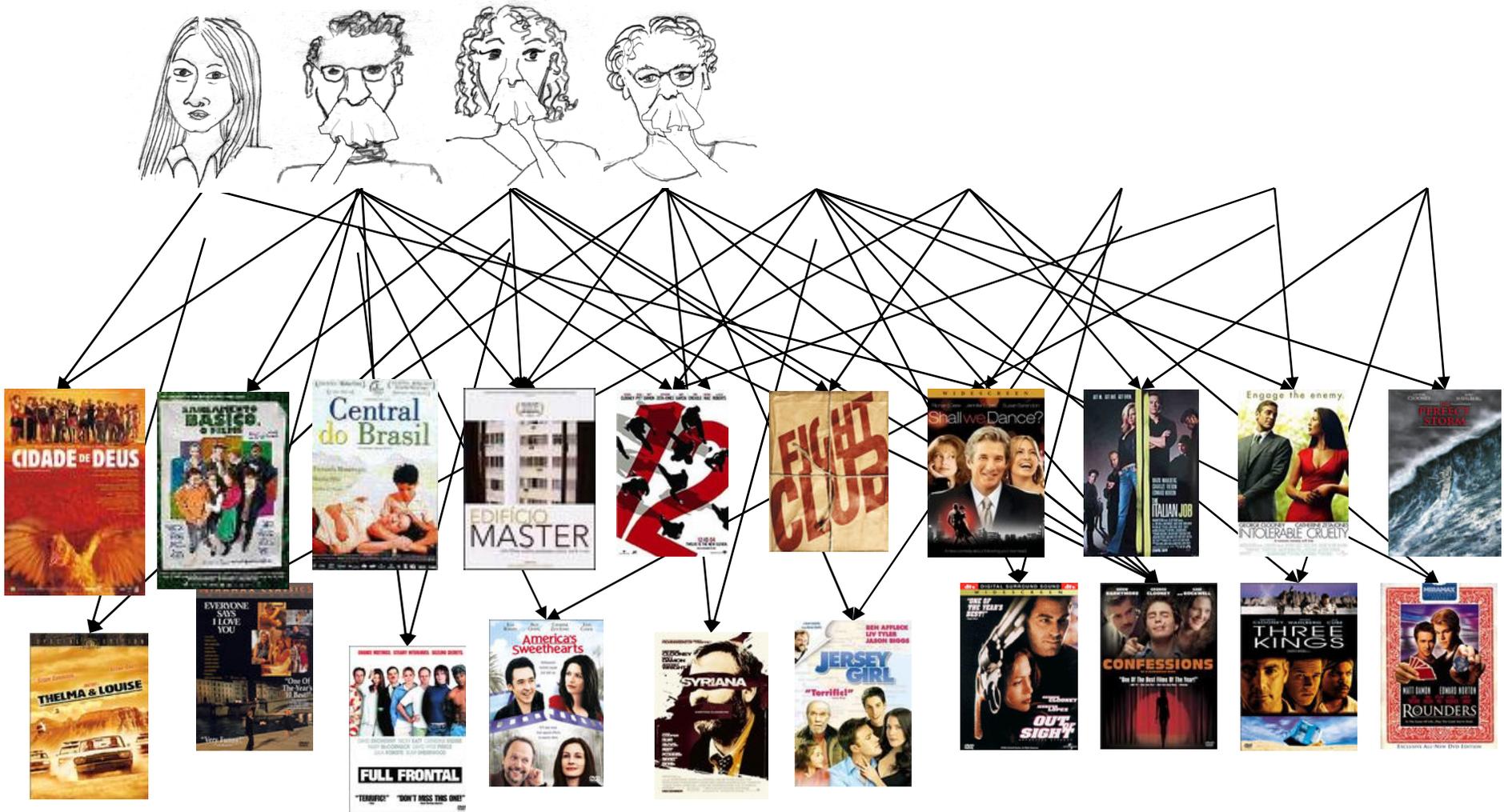
- Forma geral de multiplicação de matrizes $Z_{ij} = \sum_k X_{ik} Y_{kj}$
- Seja $Z = A$, $X = B$, $Y = B^T$

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 \\ 1 & 1 & 2 & 4 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 1*1 + 1*1 + 1*0 + 1*0 = 2$$

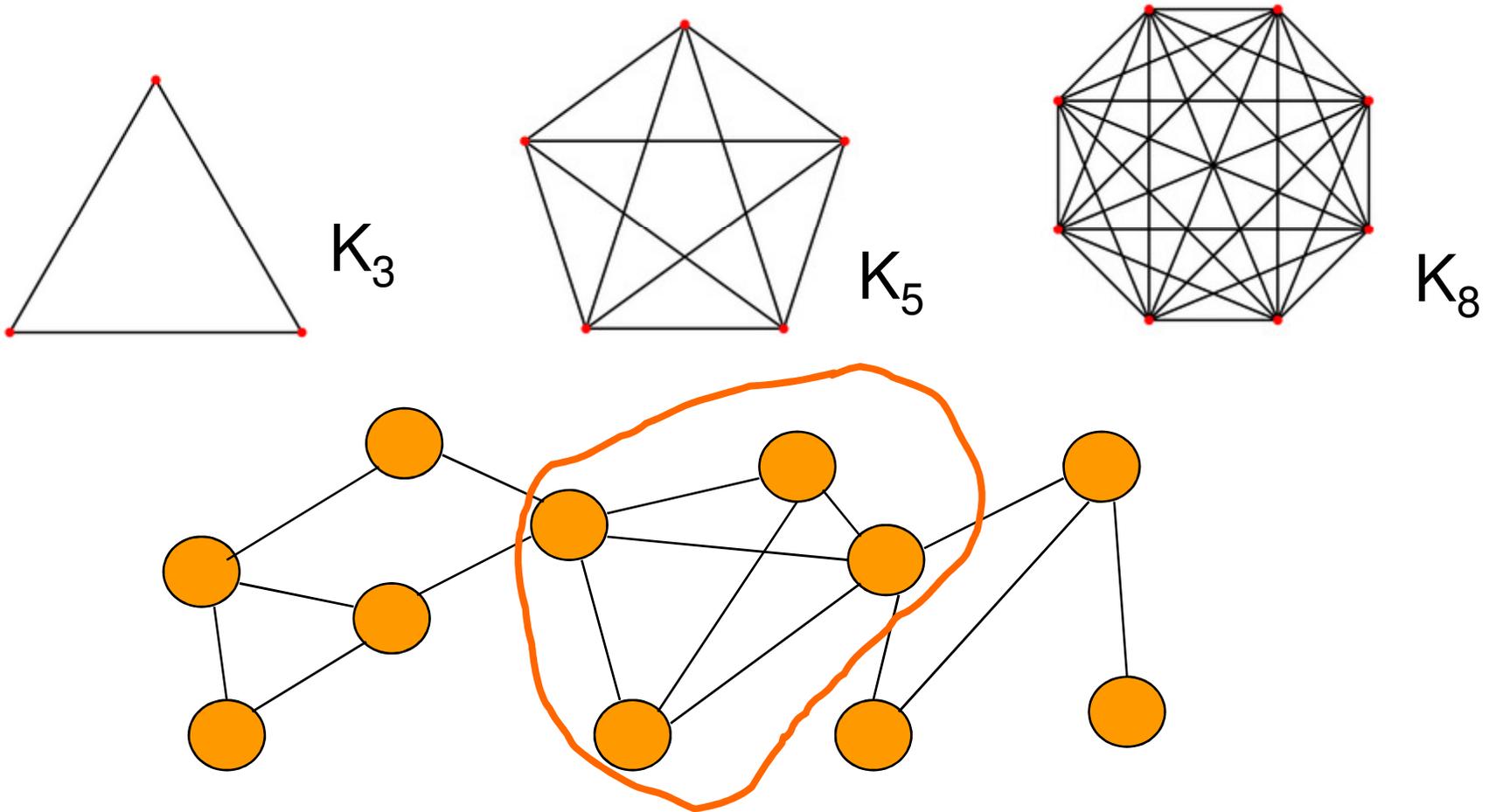


Rede de atores e filmes



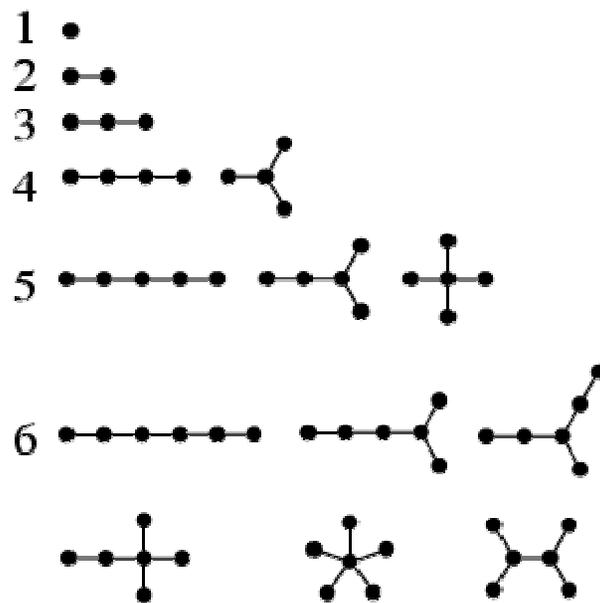
Cliques e grafos completos

- K_n é um grafo completo (clique) com n vértices
 - Cada vértice é conectado a todos outros vértices.



Árvores

- Grafos não-orientados que não contem ciclos



Clustering Coefficient

- Transitividade

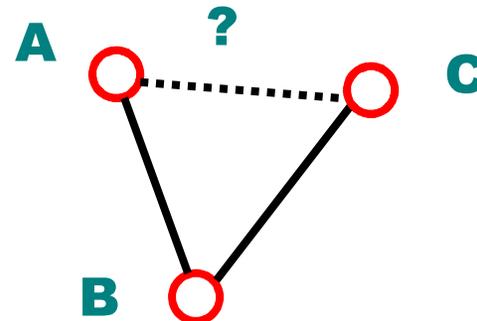
- se A está conectado a B e B é conectado a C qual a probabilidade que A seja conectado a C?

- Amigos dos meus amigos são provavelmente meus amigos...

- Clustering coefficient

$$CC = \frac{\text{numero de triângulos conectados ao nó } i}{\text{Número de triplas centradas no nó } i}$$

$$CC = [0 : 1]$$



Local clustering coefficient (Watts&Strogatz 1998)

- Para um vertice i
 - A fração de pares de vizinhos do nó que estão eles próprios conectados
 - Seja n_i o número de vizinhos do vértice i

$$CC_i = \frac{\text{number of connections between } i\text{'s neighbors}}{\text{maximum number of possible connections between } i\text{'s neighbors}}$$

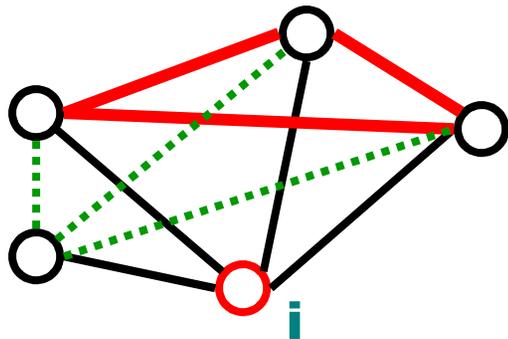
$$CC_i \text{ directed} = \frac{\# \text{ directed connections between } i\text{'s neighbors}}{n_i * (n_i - 1)}$$

$$CC_i \text{ undirected} = \frac{\# \text{ undirected connections between } i\text{'s neighbors}}{n_i * (n_i - 1)/2}$$

Local clustering coefficient (Watts&Strogatz 1998)

- Média sobre todos nós n

$$C = \frac{1}{n} \sum_i C_i$$



— link presente
- - - link ausente

$$n_i = 4$$

max número de conexões:

$$4 * 3 / 2 = 6$$

3 conexões presentes

$$C_i = 3 / 6 = 0.5$$

Buscar, instalar e trazer exemplos do Pajek

- <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

Networks / Pajek 
**Program for
Large Network
Analysis - for Windows**

**How to
run Pajek on Linux**

<http://vlado.fmf.uni-lj.si/pub/networks/pajek/howto/linux.htm>