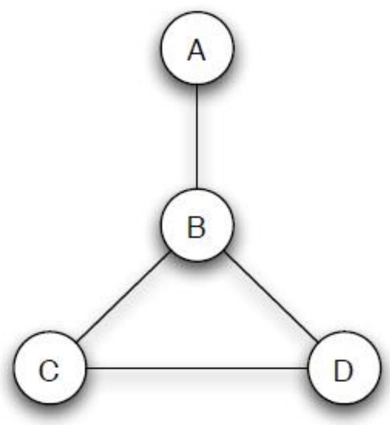
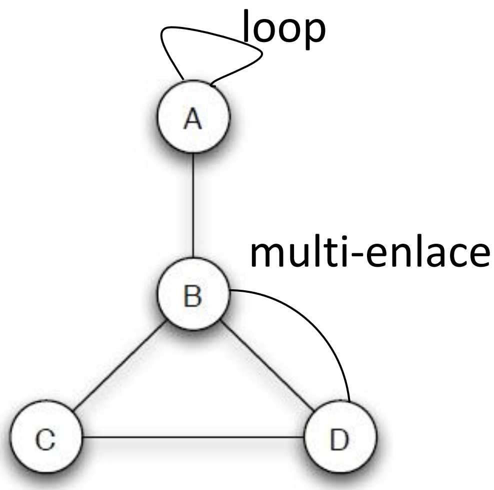
# Conceptos Básicos

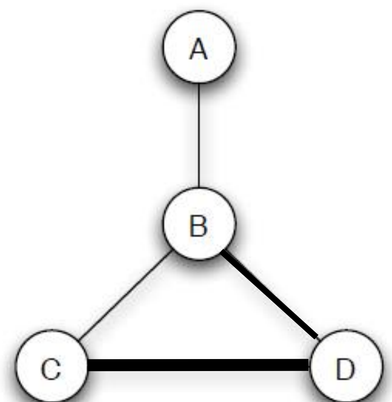
# Un grafo es…

- Un grafo *G* es un conjunto de nodos y enlaces

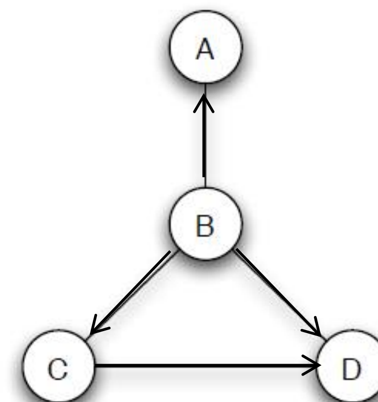



grafo simple
no-dirigido



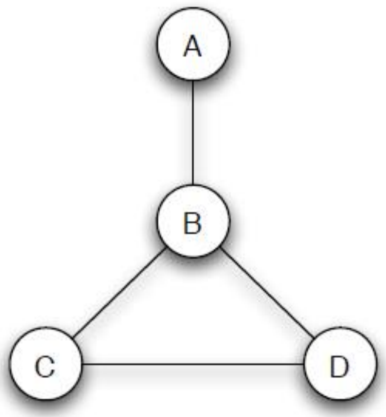

Grafo pesado



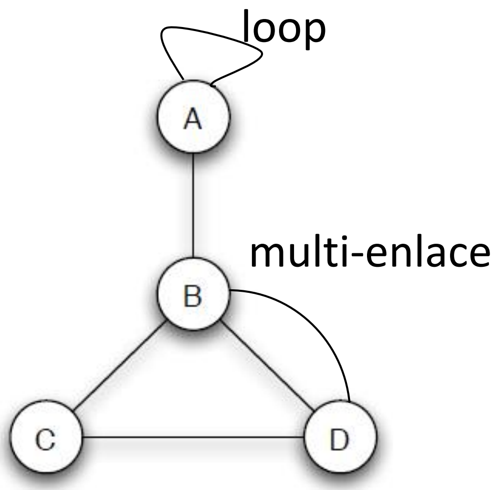

Grafo dirigido

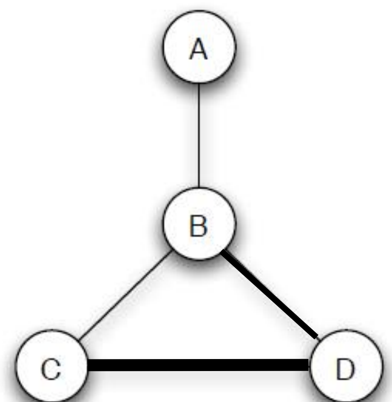
# Un grafo es…

- Un grafo *G* es un conjunto de nodos y enlaces

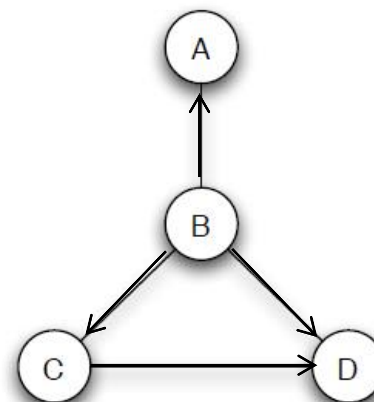



grafo simple
no-dirigido




Grafo pesado

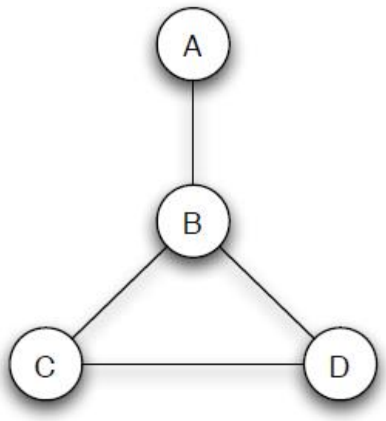



Grafo dirigido

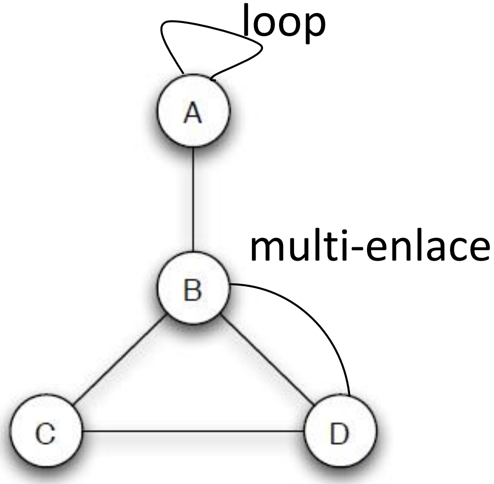
**Grafo simple no dirigido**, ***G(N,E)***, es un par de conjuntos $\boldsymbol{N} \neq 0$ y ***E*** . ***N*** es un conjunto de elementos $\boldsymbol{N}=\{n_1, n_2, \ldots, n_N\}$ llamados nodos o vértices. Los elementos de $\boldsymbol{E}=\{e_1, e_2, \ldots, e_M\}$ son <u>pares no-ordenados</u> de elementos distintos de ***N***, llamados enlaces.

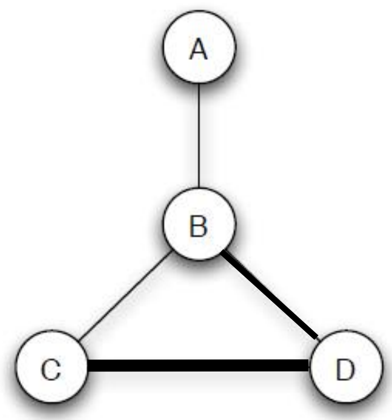
# Un grafo es…

- Un grafo *G* es un conjunto de nodos y enlaces

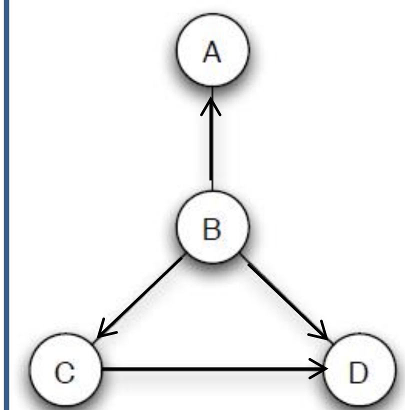



grafo simple
no-dirigido




Grafo pesado

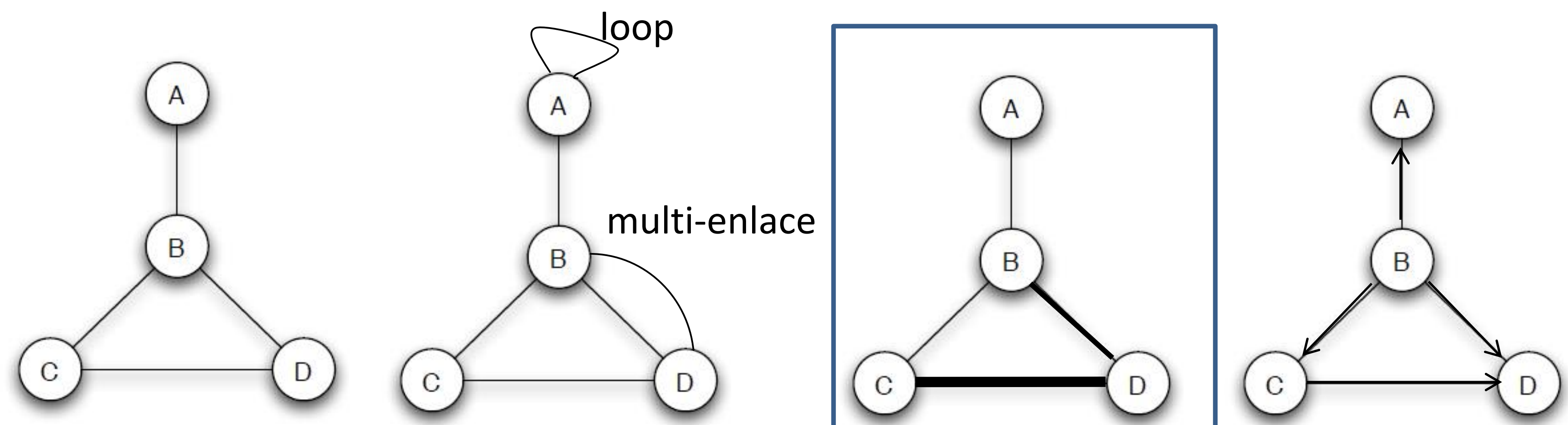



Grafo dirigido

**Grafo simple dirigido**, ***G(N,E)***, es un par de conjuntos $\boldsymbol{N} \neq 0$ y ***E*** . ***N*** es un conjunto de elementos $\boldsymbol{N}=\{n_1, n_2, \ldots, n_N\}$ llamados nodos o vértices. Los elementos de $\boldsymbol{E}=\{e_1, e_2, \ldots, e_M\}$ son <u>pares ordenados</u> de elementos distintos de ***N***, llamados enlaces.

# Un grafo es…

- Un grafo *G* es un conjunto de nodos y enlaces




grafo simple
no-dirigido

Grafo pesado

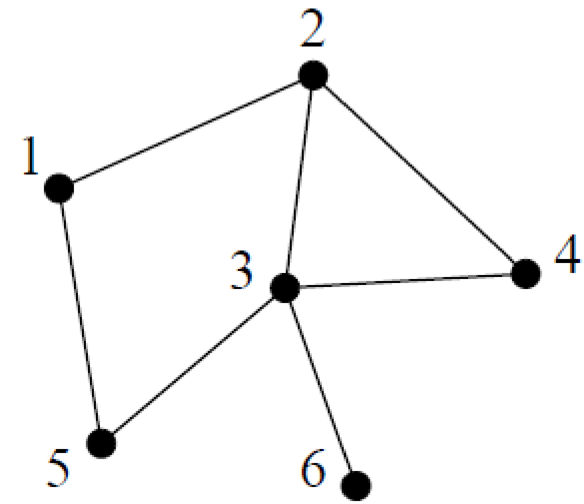
Grafo dirigido

**Grafo pesado**, ***G(N,E)***, es un par de conjuntos $\boldsymbol{N} \neq 0$ y ***E*** . ***N*** es un conjunto de elementos $\boldsymbol{N}=\{n_1, n_2, \ldots, n_N\}$ llamados nodos o vértices. Los elementos de $\boldsymbol{E}=\{e_1, e_2, \ldots, e_M\}$ son pares de elementos distintos de ***N***, llamados enlaces. Además existe la **función de peso** $\boldsymbol{w}$:$\boldsymbol{E} \times \mathbb{R}$

# Cómo representar un grafo simple

**Lista de enlaces**

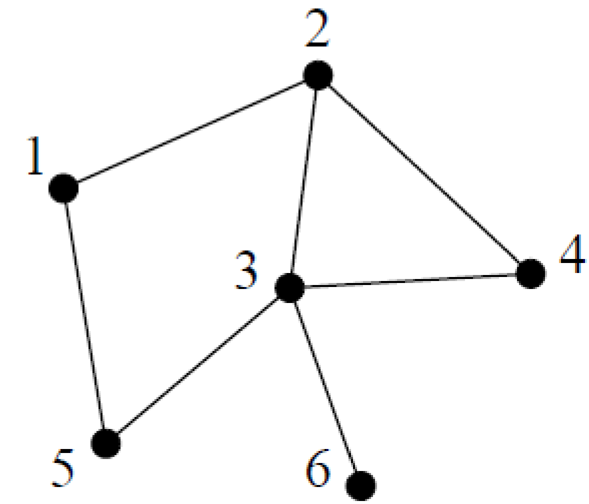
(1,2)
(1,5)
(2,3)
(2,4)
(3,4)
(3,5)
(3,6)



# Cómo representar un grafo simple

**Matriz de adyacencia**

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




$A_{ij}=$ 1 si existe enlace entre nodos *i* y *j*
0 si no

***A*** es una matriz **simétrica.**

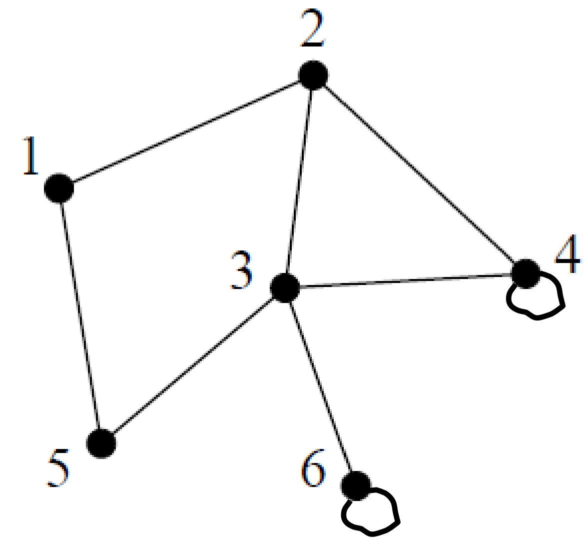
$\boldsymbol{A} = \boldsymbol{A}^T$

$A_{ij} = A_{ji}$

# Cómo representar un grafo con loops

**Matriz de adyacencia**

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




$A_{ij}=$ 1 si existe enlace entre nodos *i* y *j*
0 si no

Los ***loops*** se corresponden con
$A_{ii} = 2$

***A*** es una matriz **simétrica.**

$\boldsymbol{A} = \boldsymbol{A}^T$

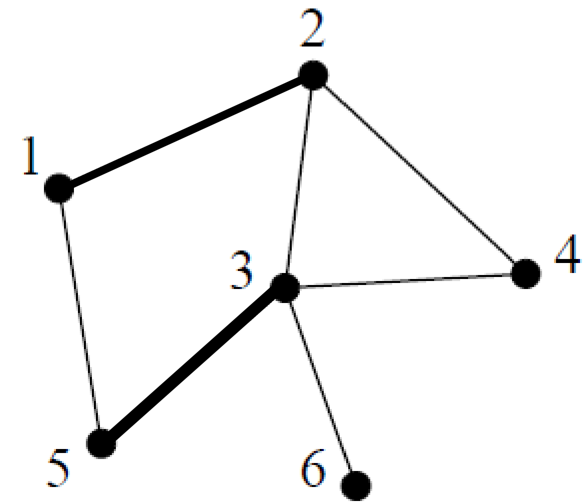
$A_{ij} = A_{ji}$

- Un enlace tipo *loop* tiene dos conexiones con el nodo-i
- Cada enlace aparece 2 veces en ***A***

# Cómo representar un grafo pesado

**Matriz de adyacencia**

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$




$A_{ij}$= $w_{ij}$ si existe enlace entre nodos *i* y *j*
0 si no

***A*** es una matriz **simétrica.**

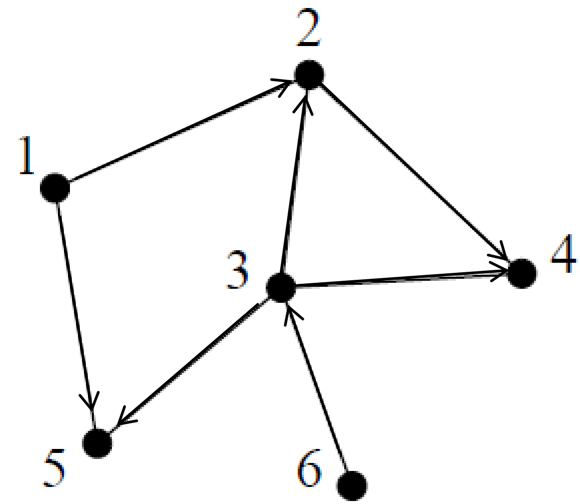
$\boldsymbol{A} = \boldsymbol{A}^T$

$A_{ij} = A_{ji}$

# Cómo representar un grafo dirigido

**Matriz de adyacencia**

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

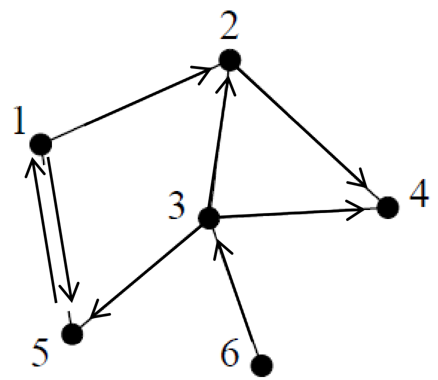



$A_{ij}=$ 1 si existe enlace que **llega a *i* desde *j***
0 si no

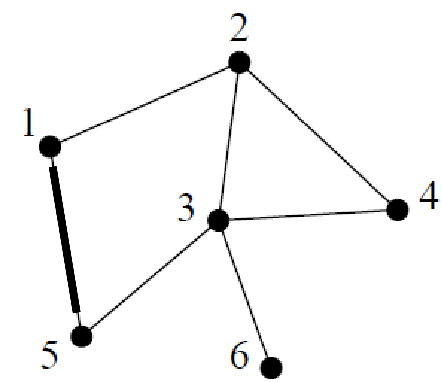
***A*** <u>no es</u> una matriz **simétrica.**

# Versión no-dirigida de grafo dirigido

¿Cuál es la version no-dirigida de un grafo dirigido?



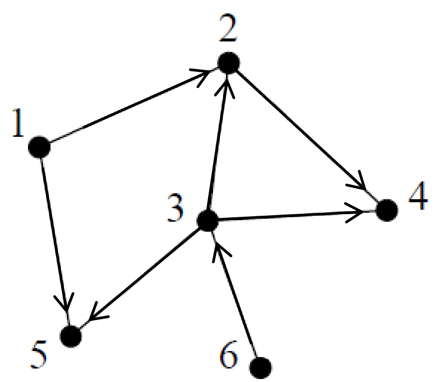

***A***




$\boldsymbol{A' = A + A^T}$

La red no-dirigida tiene menos información [$N^2$ vs N(N+1)/2 grados de libertad].
Existen otras maneras de *proyectar* recapitulando otros aspectos de la red dirigida…

# Proyección co-citas de red dirigida




Coeficiente de **cocitación** entre nodos *i* y *j*:

- Número de nodos que tienen enlaces que apuntan a nodos *i* y *j* al mismo tiempo

$$C_{ij} = \sum_{k=1}^{n} A_{ik}A_{jk} = \sum_{k=1}^{n} A_{ik}(\boldsymbol{A}^T)_{kj}$$

$$C = AA^T$$

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

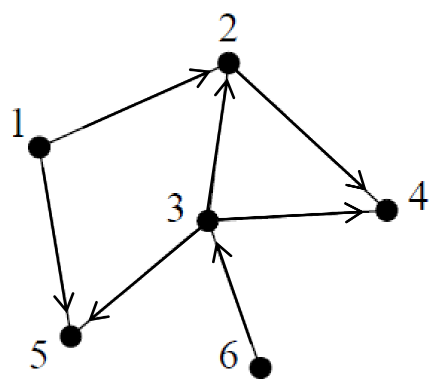
- Notar que

$$C_{ii} = \sum_{k=1}^{n} A_{ik}A_{ik} = \sum_{k=1}^{n} A_{ik} \geq 0$$



$C_{ii}$ = num_enlaces_entrantes = *in_degree* = $k_{in}$

# Proyección co-citas de red dirigida



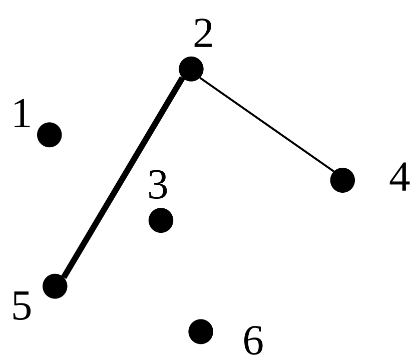

Coeficiente de **cocitación** entre nodos *i* y *j*:

- Número de nodos que tienen enlaces que apuntan a nodos *i* y *j* al mismo tiempo

$$C_{ij} = \sum_{k=1}^{n} A_{ik} A_{jk} = \sum_{k=1}^{n} A_{ik} (\boldsymbol{A}^T)_{kj}$$

$$C = AA^T$$

$$\tilde{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$



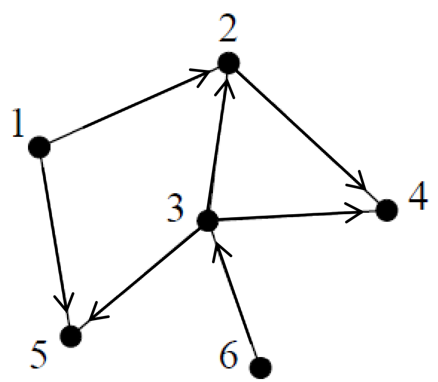

Puedo armar una red no-dirigida si considero

$$\tilde{C} = C - diag(C)$$



Red de cocitación: enlaces cuantifican noción de *similitud* inferida a partir de patrón de citas recibidas

# Proyección acople bibliográfico




Coeficiente de **acople bibliográfico** entre nodos *i* y *j* :
Número de nodos referenciados, a la vez, por los nodos *i* y *j*

$$B_{ij} = \sum_{k=1}^{n} A_{ki} A_{kj} = \sum_{k=1}^{n} (\boldsymbol{A}^T)_{ik} A_{kj}$$

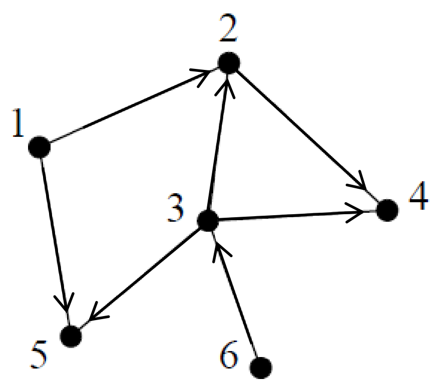
$$B = A^T A$$

$$B = \begin{pmatrix} 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$B_{ii} = \sum_{k=1}^{n} A_{ki} A_{ki} = \sum_{k=1}^{n} A_{ki}$$

$B_{ii}$ = num_enlaces_salientes = *out_degree* = $k_{out}$

# Cómo representar un grafo dirigido



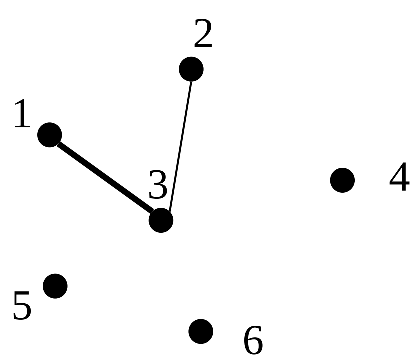

Coeficiente de **acople bibliográfico** entre nodos *i* y *j* :

Número de nodos referenciados, a la vez, por los nodos *i* y *j*

$$B_{ij} = \sum_{k=1}^{n} A_{ki}A_{kj} = \sum_{k=1}^{n} (\boldsymbol{A}^T)_{ik}A_{kj}$$

$$B = A^T A$$

$$B = \begin{pmatrix} 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

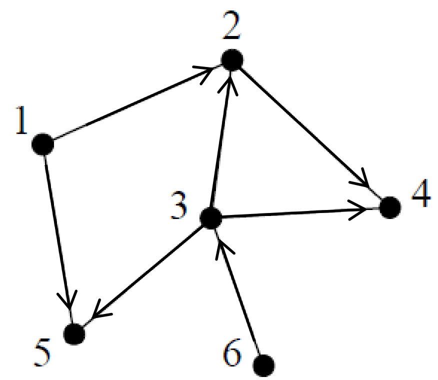


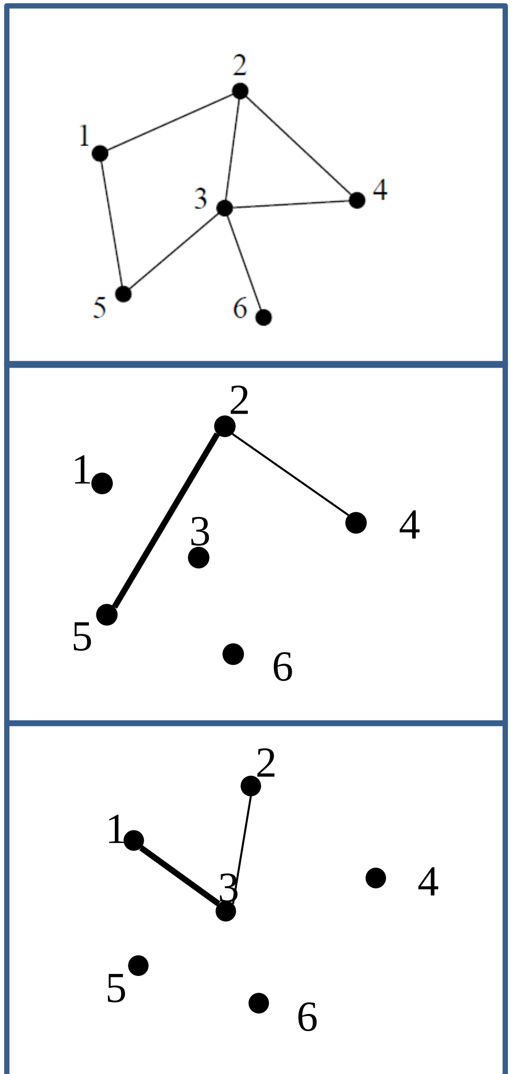

$$\tilde{B} = B - diag(B)$$

Red de acople bibliografico: noción de *similitud* inferida a partir de patrón de citas realizadas

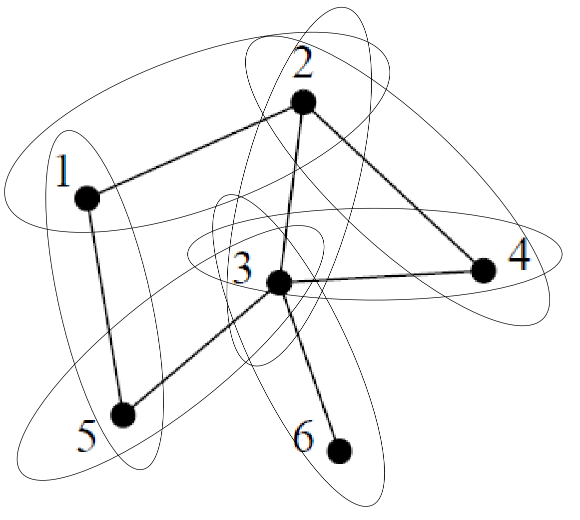
# Cómo representar un grafo dirigido

¿Cuál es la version no-dirigida de un grafo dirigido?



# Hipergrafos



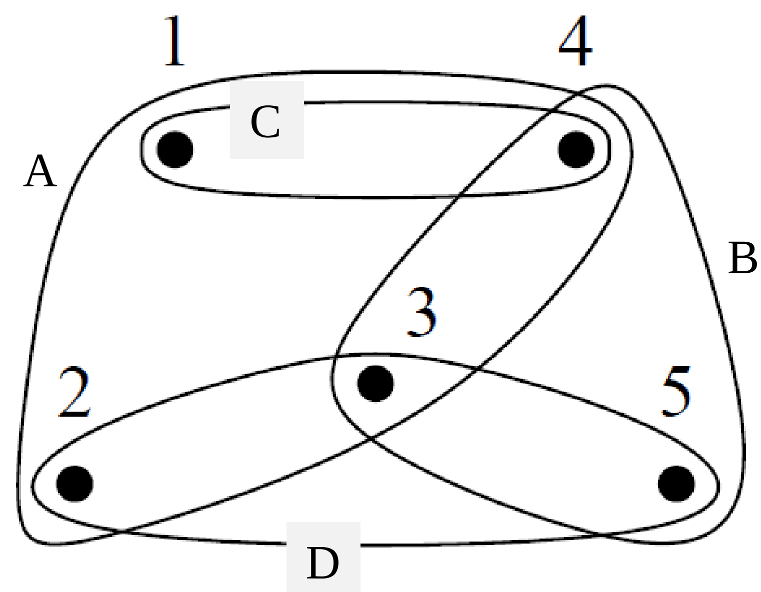

Cada enlace se corresponde con una asociación entre **pares** de vértices.

Es posible extender el concepto de enlace para conjunto s de más de dos elementos?

# Hipergrafos

Descripción que involucre explicitar relaciones entre más de dos entidades (hiperenlaces)

- Concepto de familias
- Reacciones bioquímicas que involucren más de dos especies
- …




**Matriz de incidencia**

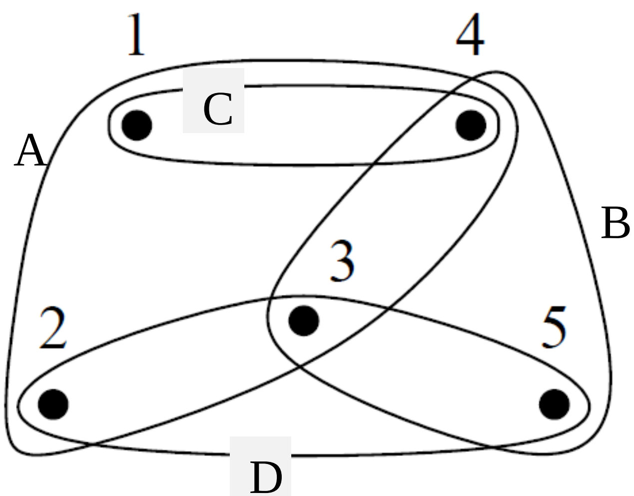
$$I = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{cccc} A & B & C & D \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}$$

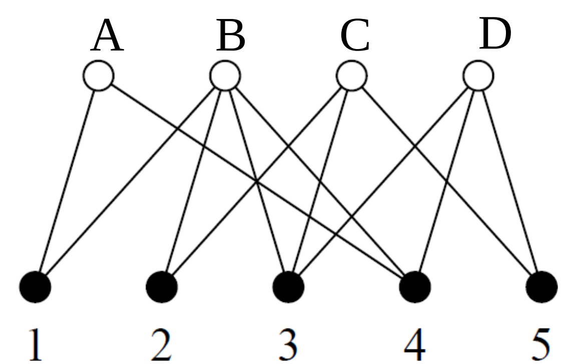
La matriz $\boldsymbol{I}$ involucra a dos tipos de entidades

$I_{i\alpha}$ 1 si nodo *i* esta incluido en conjunto $\alpha$
0 si no

# Hipergrafos <-> redes de filiación

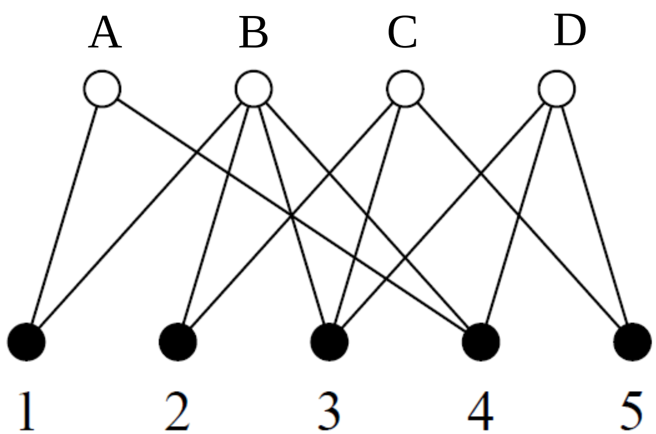
En gral la info contenida en hipergrafos puede representarse mediante **redes de filiación** compuesta por dos tipos de nodos: elementos e hiperenlaces que dan lugar a **focos**




$$
I = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\begin{array}{cccc} A & B & C & D \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}
$$

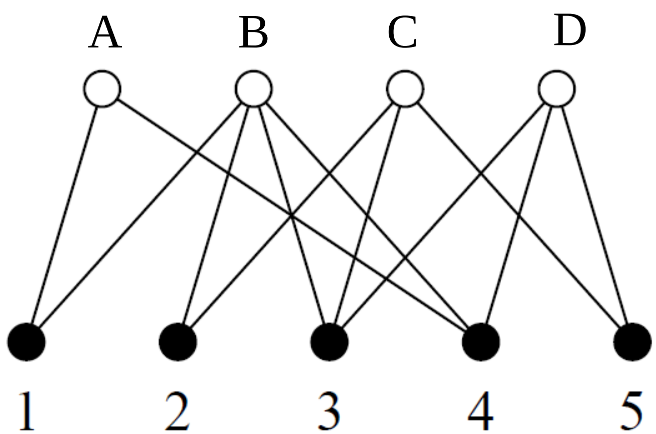
# Red Bipartita




$$
\boldsymbol{A} \;=\; \begin{array}{c|ccccccccc}
 & 1 & 2 & 3 & 4 & 5 & A & B & C & D \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
4 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
A & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
B & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
D & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0
\end{array}
$$

**Red bipartita**: grafo *G(V, A, E)* constituido por tres conjuntos: $V \neq 0$, $A \neq 0$ y $E$. Los elementos de $V=\{v_1, v_2, \ldots, v_N\}$ y $A=\{a_1, a_2, \ldots, a_L\}$ son distintos y se denominan nodos tipo-*V* y tipo-*A* de la red respectivamente. Los elementos de $E=\{e_1, e_2, \ldots, e_M\}$, denominados enlaces, son pares diferentes de elementos no-ordenados , uno de *tipo-V* otro de *tipo-A.*

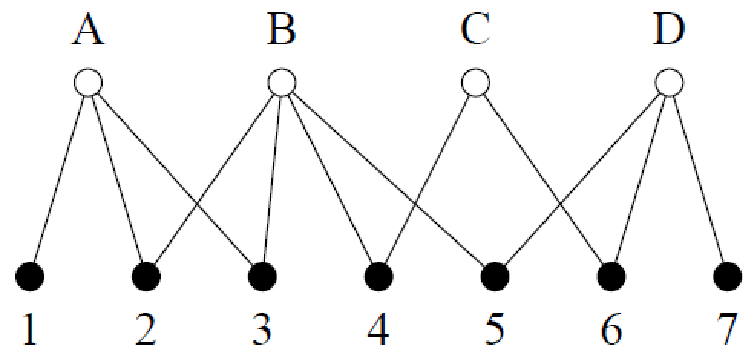
# Red Bipartita

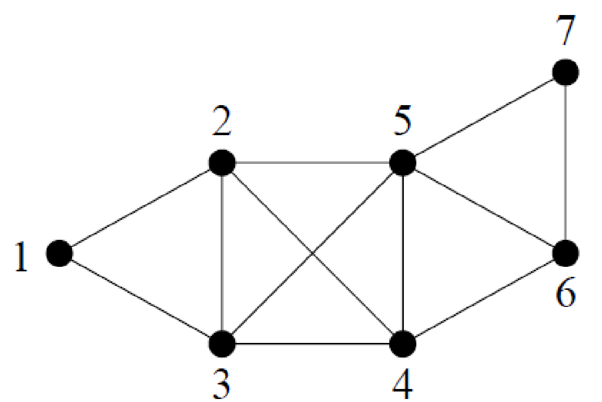


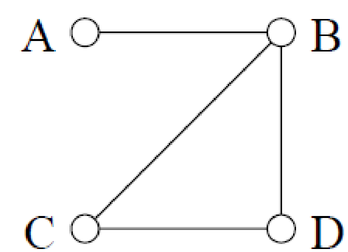

$$
\boldsymbol{A} = \begin{array}{c|ccccc|cccc}
 & 1 & 2 & 3 & 4 & 5 & A & B & C & D \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & & & & \\
2 & 0 & 0 & 0 & 0 & 0 & & & & \\
3 & 0 & 0 & 0 & 0 & 0 & & & I & \\
4 & 0 & 0 & 0 & 0 & 0 & & & & \\
5 & 0 & 0 & 0 & 0 & 0 & & & & \\
\hline
A & & & & & & 0 & 0 & 0 & 0 \\
B & & & & & & 0 & 0 & 0 & 0 \\
C & & & I^T & & & 0 & 0 & 0 & 0 \\
D & & & & & & 0 & 0 & 0 & 0
\end{array}
$$

**Red bipartita**: grafo $G(V, A, E)$ constituido por tres conjuntos: $V \neq 0$, $A \neq 0$ y $E$. Los elementos de $V=\{v_1, v_2, \ldots, v_N\}$ y $A=\{a_1, a_2, \ldots, a_L\}$ son distintos y se denominan nodos tipo-$V$ y tipo-$A$ de la red respectivamente. Los elementos de $E=\{e_1, e_2, \ldots, e_M\}$, denominados enlaces, son pares diferentes de elementos no-ordenados, uno de *tipo-V* otro de *tipo-A.*

# Proyecciones de una Red Bipartita




$$P_{ij} = \sum_{\alpha=1}^{L} I_{i\alpha} I_{j\alpha} = \sum_{\alpha=1}^{L} I_{i\alpha} \, (I^T)_{\alpha j}$$

$$P = I \, . \, I^T$$

$$P'_{\alpha\beta} = \sum_{i=1}^{N} I_{i\alpha} I_{i\beta} = \sum_{i=1}^{N} \, (I^T)_{\alpha i} I_{i\beta}$$

$$P' = I^T . \, I$$



Cual es el significado de $P_{ij}$?

# The human disease network

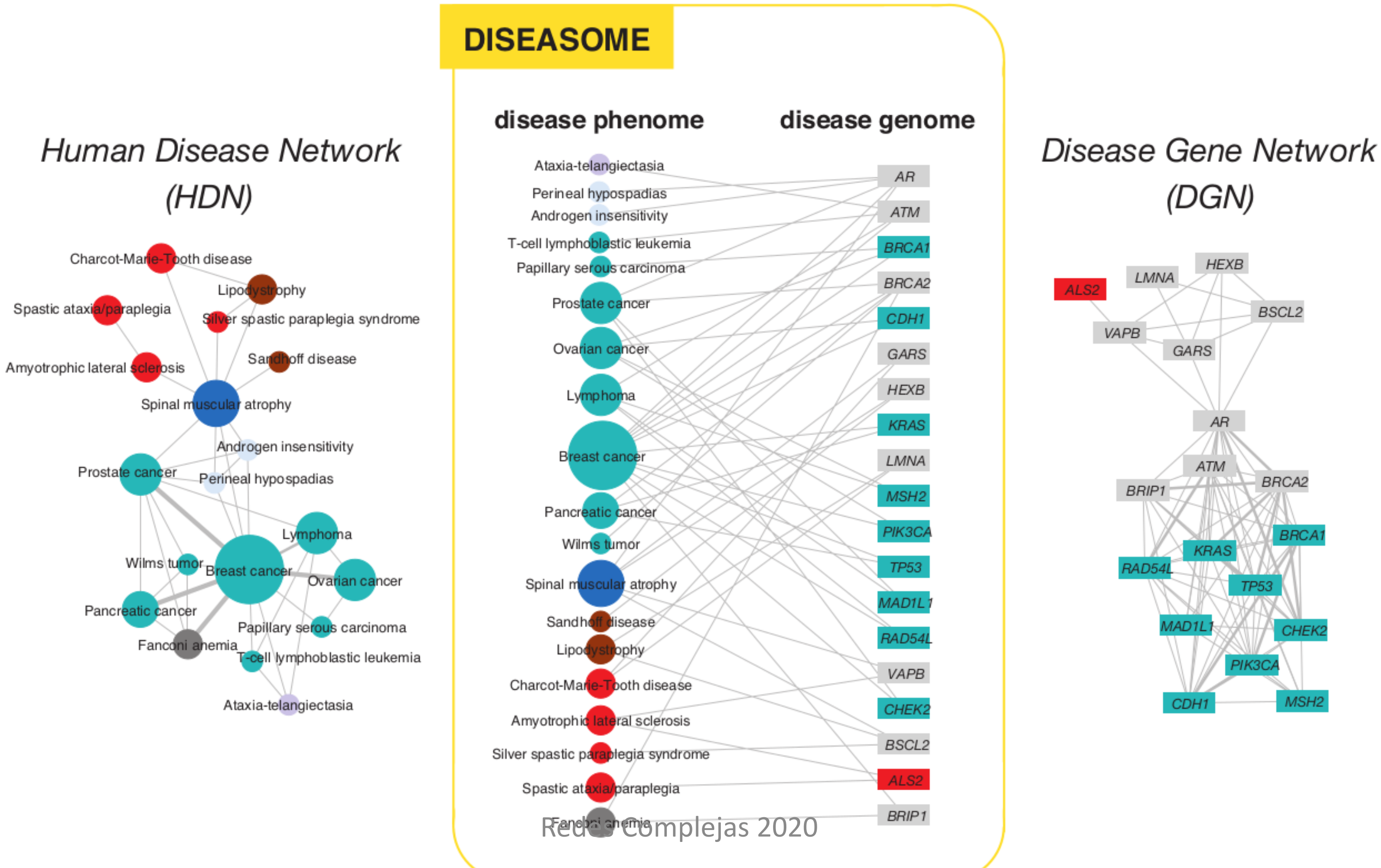
**Kwang-Il Goh*†‡§, Michael E. Cusick†‡¶, David Valle‖, Barton Childs‖, Marc Vidal†‡¶**, and Albert-László Barabási*†‡****

*Center for Complex Network Research and Department of Physics, University of Notre Dame, Notre Dame, IN 46556; †Center for Cancer Systems Biology (CCSB) and ¶Department of Cancer Biology, Dana–Farber Cancer Institute, 44 Binney Street, Boston, MA 02115; ‡Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115; §Department of Physics, Korea University, Seoul 136-713, Korea; and ‖Department of Pediatrics and the McKusick–Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205

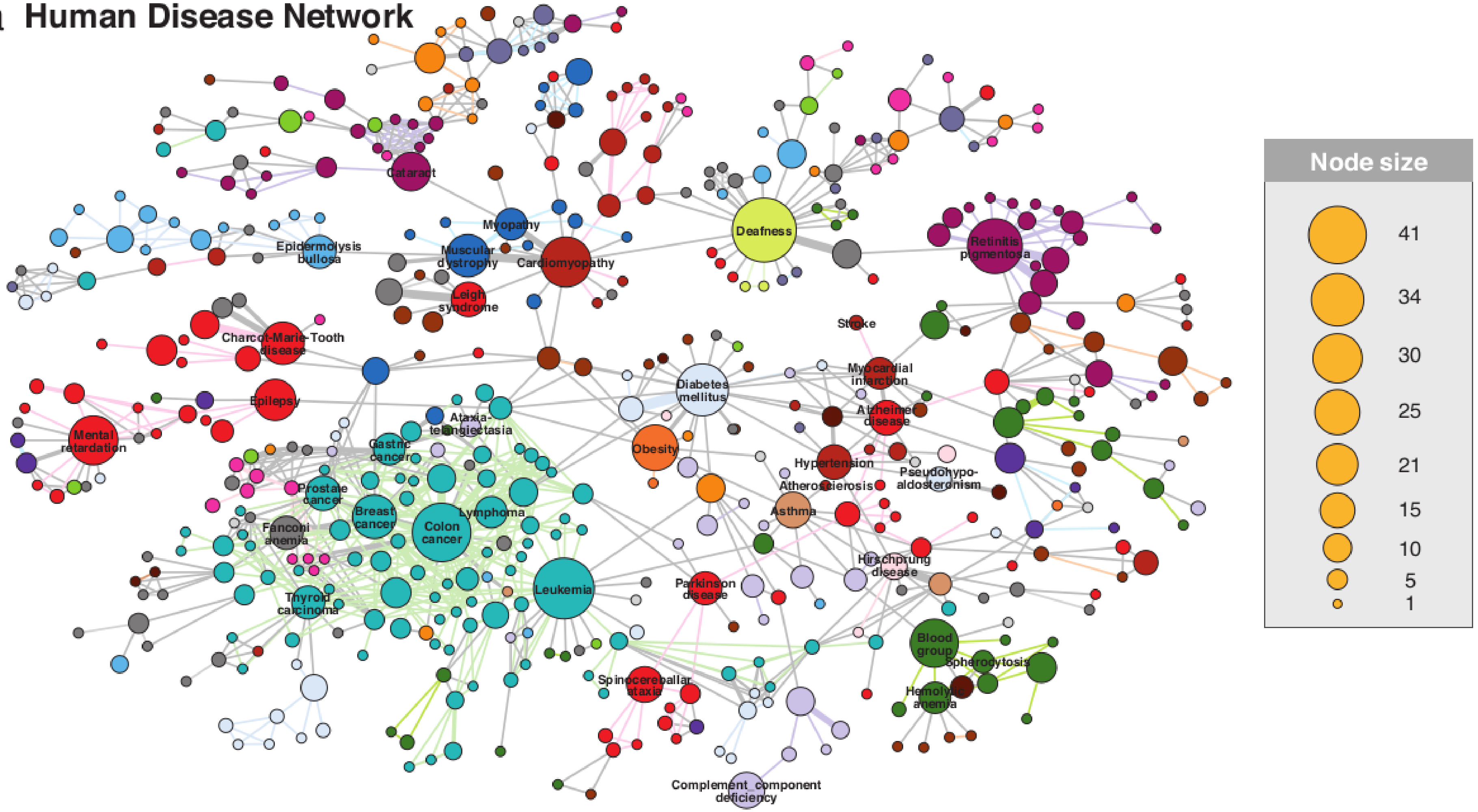
Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved April 3, 2007 (received for review February 14, 2007)

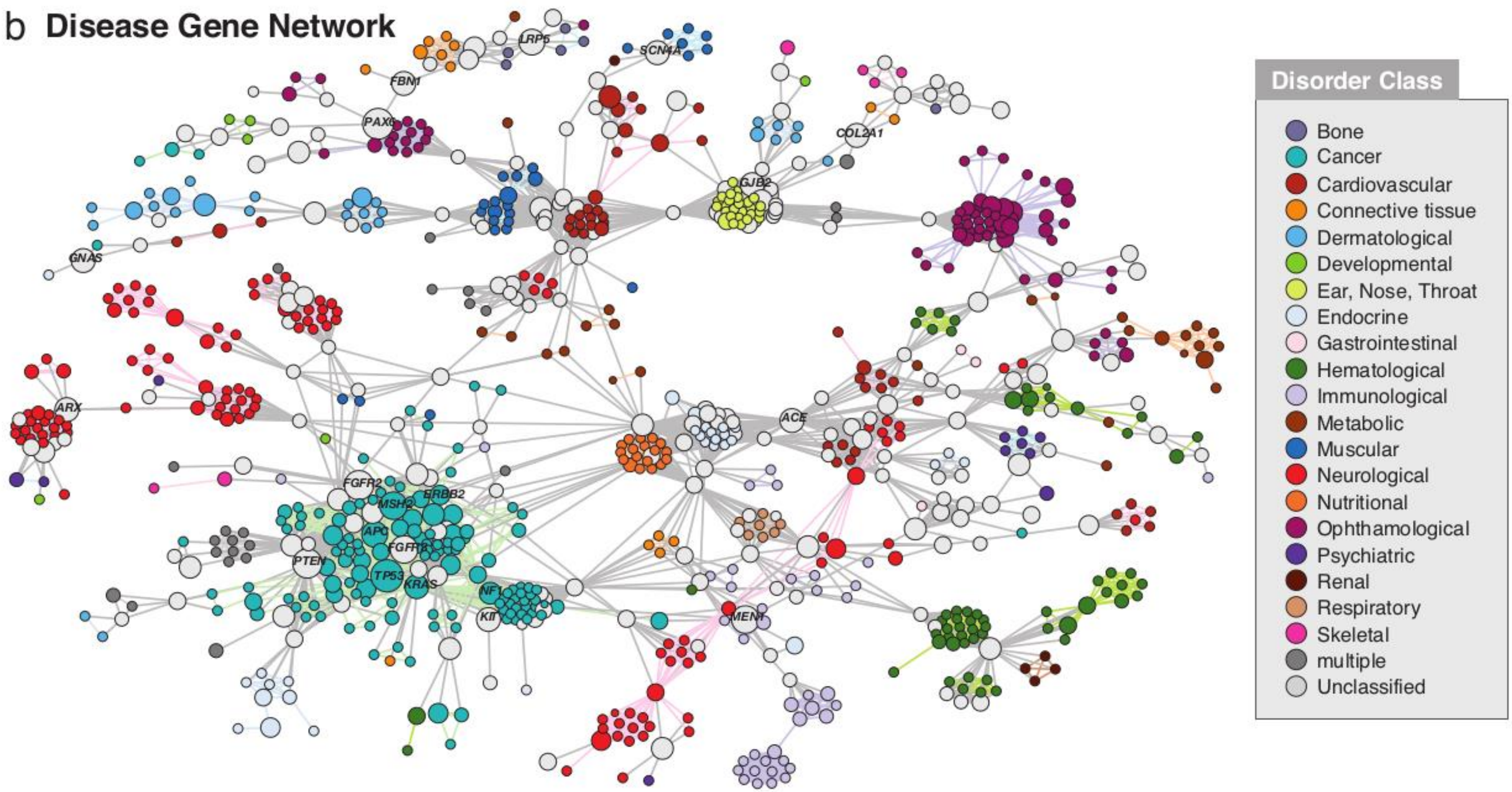
**A network of disorders and disease genes linked by known disorder–gene associations offers a platform to explore in a single graph-theoretic framework all known phenotype and disease gene associ-**

known genetic disorders, whereas the other set corresponds to all known disease genes in the human genome (Fig. 1). A disorder and a gene are then connected by a link if mutations in that gene are



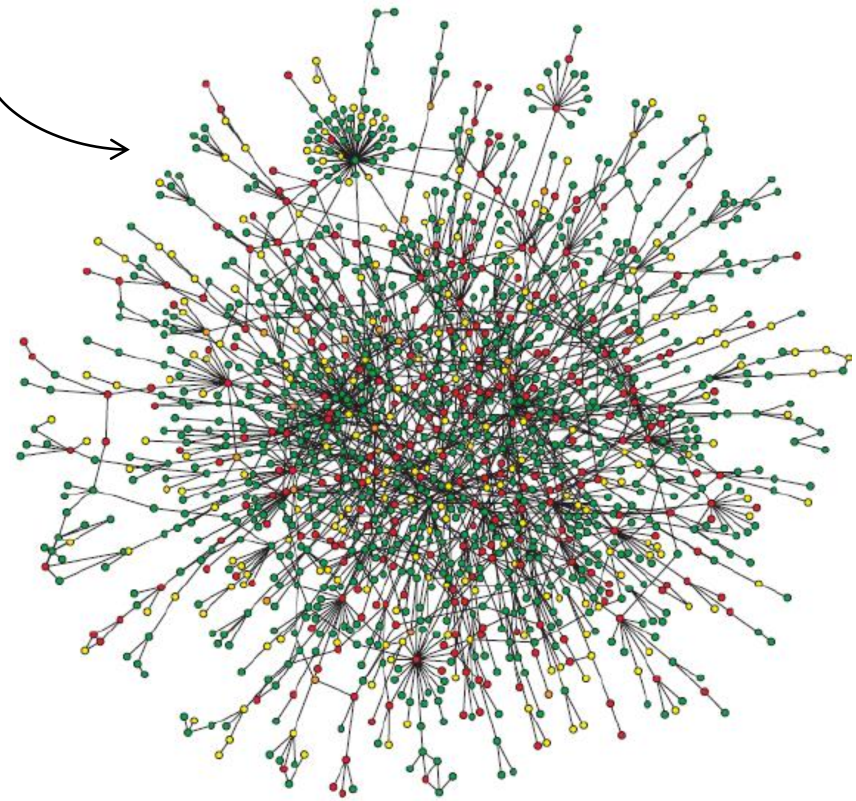
a **Human Disease Network**




b Disease Gene Network
FBN1
LRP5
SCN4A
PAX6
COL2A1
GJB2
GNAS
ARX
ACE
FGFR2
ERBB2
MSH2
APC
FGFR3
PTEN
TP53
KRAS
NF1
KIT
MEN1
Disorder Class
Bone
Cancer
Cardiovascular
Connective tissue
Dermatological
Developmental
Ear, Nose, Throat
Endocrine
Gastrointestinal
Hematological
Immunological
Metabolic
Muscular
Neurological
Nutritional
Ophthamological
Psychiatric
Renal
Respiratory
Skeletal
multiple
Unclassified

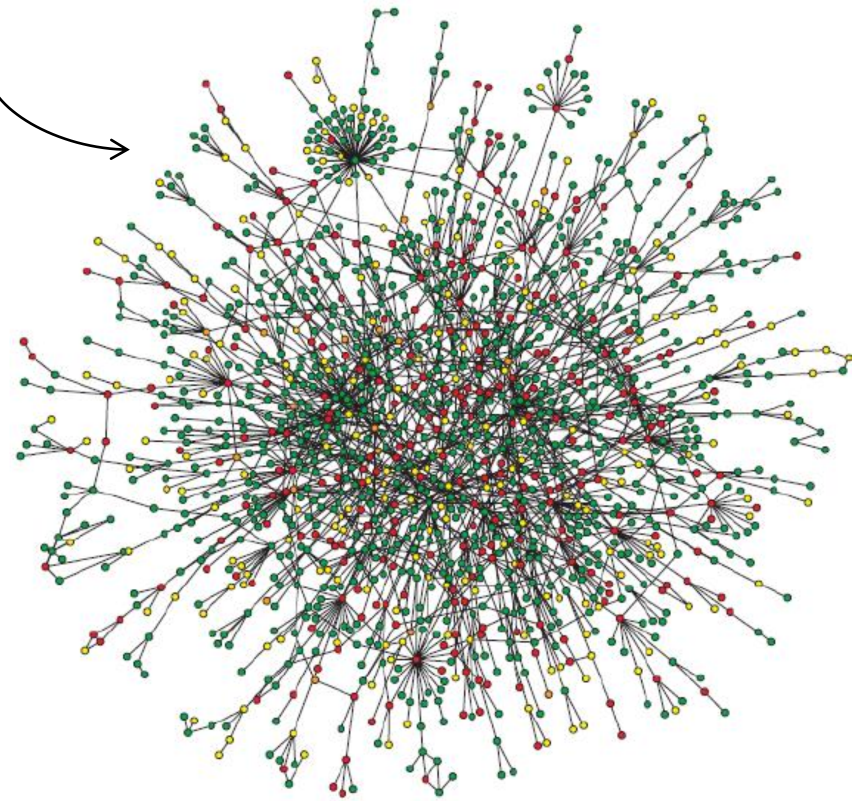
# Caracterizando redes complejas

Propiedades cuantitativas para analizar y entender como están organizadas este tipo de *cosas*



# Propiedades de redes complejas

Propiedades cuantitativas para analizar y entender como están organizadas este tipo de *cosas*



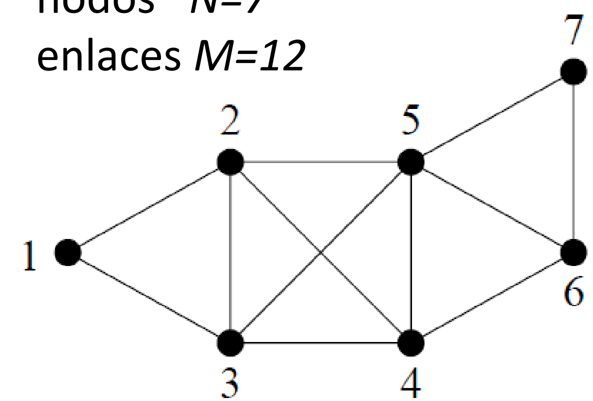
- Distribución de grado
- Asortatividad/homofilia
- Clustering
- Motifs
- Modularidad
- Distancias
- Centralidad
- Eficiencia
- Robustez
- ….

# Grado

**Grado de un nodo**: número de enlaces incidentes

$$k_i = \sum_{j=1}^{N} A_{ij}$$

nodos *N=7*
enlaces *M=12*




Además, como cada *enlace* tiene dos extremos*:

$2M = \sum_{i=1}^{N} k_i$ por lo que $M = \frac{1}{2}\sum_{i,j} A_{ij}$

*$k_1$=2, $k_2$=4, $k_3$=4 …*

*2 * 12 = 2+4+4+4+5+3+2*

**Grado medio**

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2M}{N}$$

*<k>=24/7=3.43*



*Esto hace sentido aun para grafos con loops (si asignamos adyacencia 2 a elementos diagonales con loops)

# Densidad

Máximo número de enlaces para un grafo de N vértices:

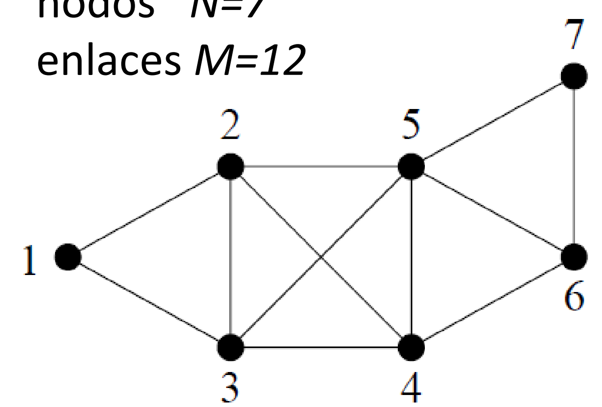
$$M_{max} = \binom{N}{2} = \frac{N * (N-1)}{2}$$

**Densidad del grafo**

$$\rho = \frac{M}{M_{max}}$$

$$= \frac{M}{\binom{N}{2}} = \frac{2 * M}{N * (N-1)} = \frac{< k >}{N-1}$$

nodos *N=7*
enlaces *M=12*




$k_1$=2, $k_2$=4, $k_3$=4 ...

$<k>$=24/7=3.43

$M_{max}$=21

$\rho$=0.67

# Redes densas o ralas

$$\rho = \frac{2 * M}{N * (N-1)} = \frac{< k >}{N-1}$$

Supongamos que analizamos una red que crece en el tiempo: N=N(t), M=M(t)

Una red es **rala** si $\underset{N \to \infty}{\rho(t) \to 0}$ Internet, WWW, redes de amistad

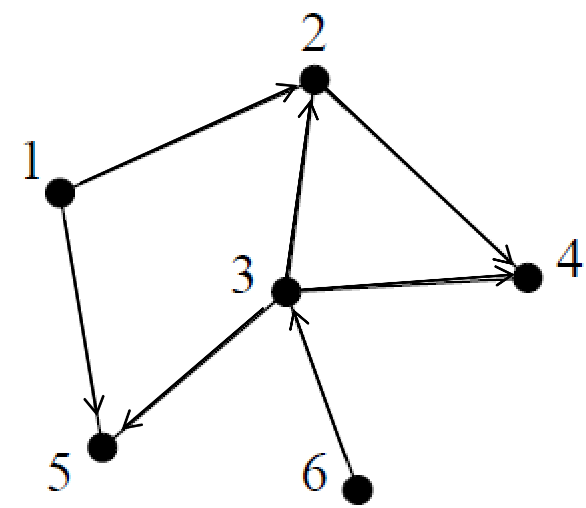
**densa** si $\underset{N \to \infty}{\rho(t) \to cte}$ redes tróficas

# Grado de redes dirigidas

$$k_i^{\ in} = \sum_j A_{ij}$$

$$k_j^{\ out} = \sum_i A_{ij}$$

nodos *N=6*




Cada **enlace** tiene un *extremo-in* y otro *out*

$$M = \sum_{i=1}^{N} k_i^{\ in} = \sum_{i=1}^{N} k_i^{\ out} = \sum_{i,j} A_{ij}$$

$k^{in}_3 = 1$
$k^{out}_3 = 3$

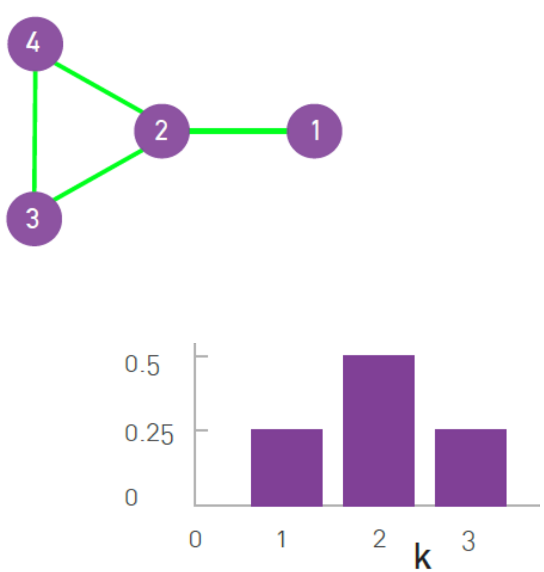
**Grado medio**

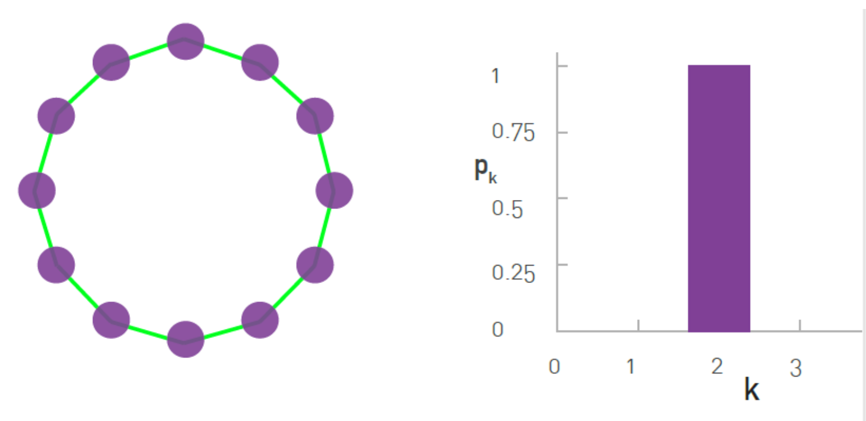
$$\left.\begin{aligned} \langle k^{in} \rangle &= \tfrac{1}{N} \sum_{i=1}^{N} k_i^{\ in} \\ \langle k^{out} \rangle &= \tfrac{1}{N} \sum_{i=1}^{N} k_i^{\ out} \end{aligned}\right\} = \tfrac{1}{N} \sum_{i,j} A_{ij} = \frac{M}{N}$$

# Distribución de grado

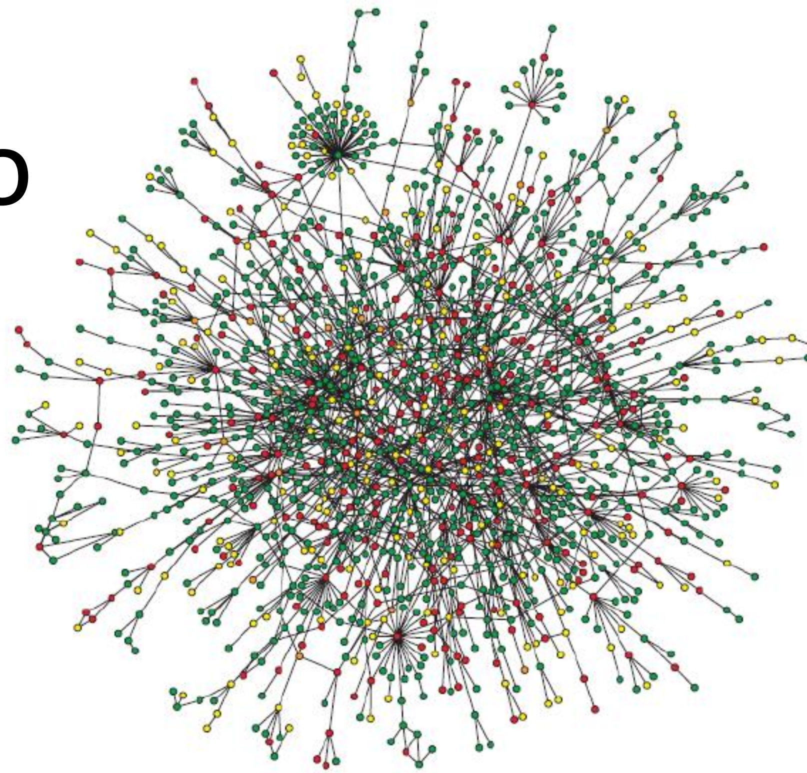
La **distribución de grado**, $p_k$, de nodos de una red provee la probabilidad de que un nodo elegido al azar tenga grado k.

$$p_k = \frac{n_k}{N} \qquad \sum_k^N p_k = 1$$

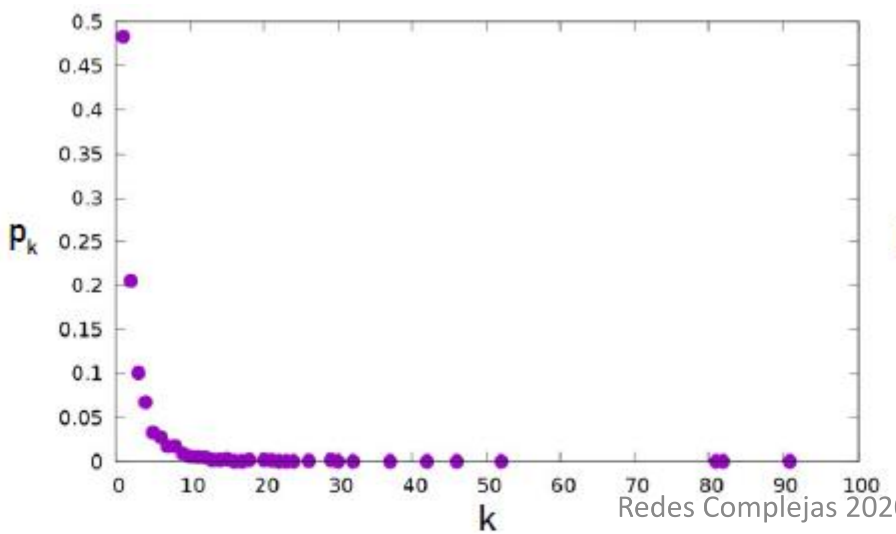


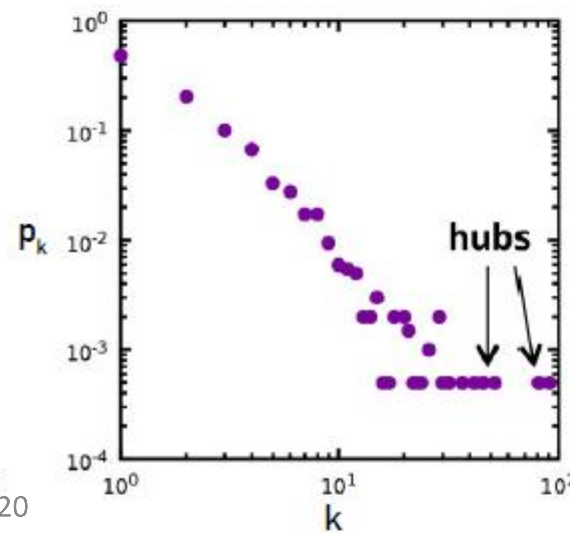



# Distribución de grado



- En redes reales el grado puede variar muchísimo



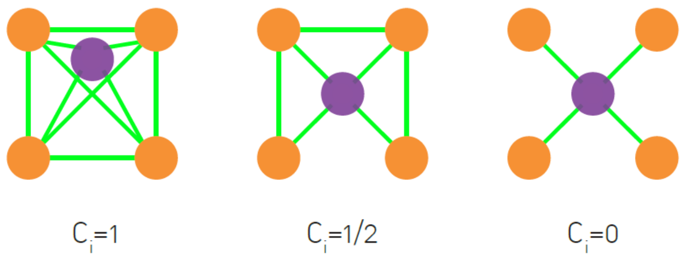
# Coeficiente de *clustering*

En qué medida los vecinos de un nodo son vecinos entre sí ?

**Coeficiente de clustering local:**

pares de vecinos enlazados (triángulos)

$$C_i = \frac{2L_i}{k_i(k_i-1)}$$




$$L_i = \sum_{\substack{j,k \\ i \neq j \neq k}} a_{ij}\, a_{ik} a_{jk} = \sum_{\substack{j,k \epsilon N(i) \\ j \neq k}} a_{jk}$$

- probabilidad que pares de vecinos del nodo i estén conectados
- Medida de densidad local

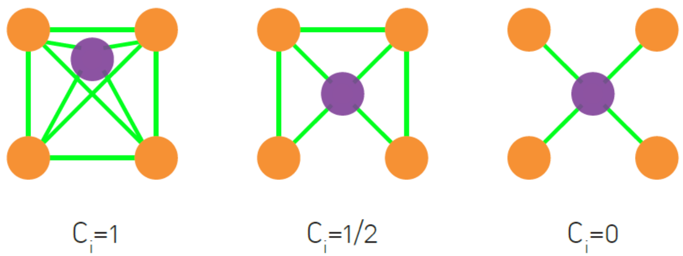
# Coeficiente de *clustering*

En qué medida los vecinos de un nodo son vecinos entre sí ?

**Coeficiente de clustering local:**

pares de vecinos enlazados (triángulos)

$$C_i = \frac{2L_i}{k_i(k_i - 1)}$$




En qué medida pares de vecinos de <u>un nodo tomado al azar en la red</u> son vecinos entre sí ?

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

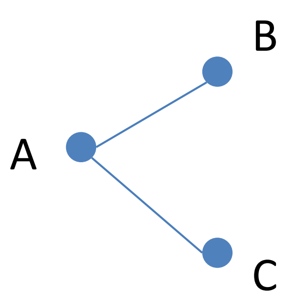
# Coeficiente de *clustering*

**Coeficiente de clustering global**

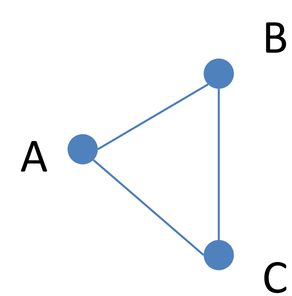
que tan probable es encontrar triangulos (a.k.a. clausura transitiva) en la red ?

$$C_{\Delta} = \frac{3 \times Number\ Of\ Triangles}{Number\ Of\ Connected\ Triples}$$

**triplete conectado**: conjunto de 3 nodos X, Y, Z donde X esta conectado con Y e Y con Z




BAC




BAC
ACB
CBA

# Coeficiente de *clustering*

$$C_i = \frac{2L_i}{k_i(k_i - 1)}$$

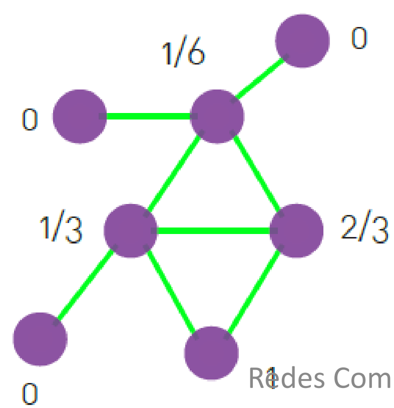
En qué medida los vecinos de un nodo i son vecinos entre sí ?

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

En qué medida pares de vecinos de un nodo tomado al azar en la red son vecinos entre sí ?

$$C_\Delta = \frac{3 \times Number\ Of\ Triangles}{Number\ Of\ Connected\ Triples}$$

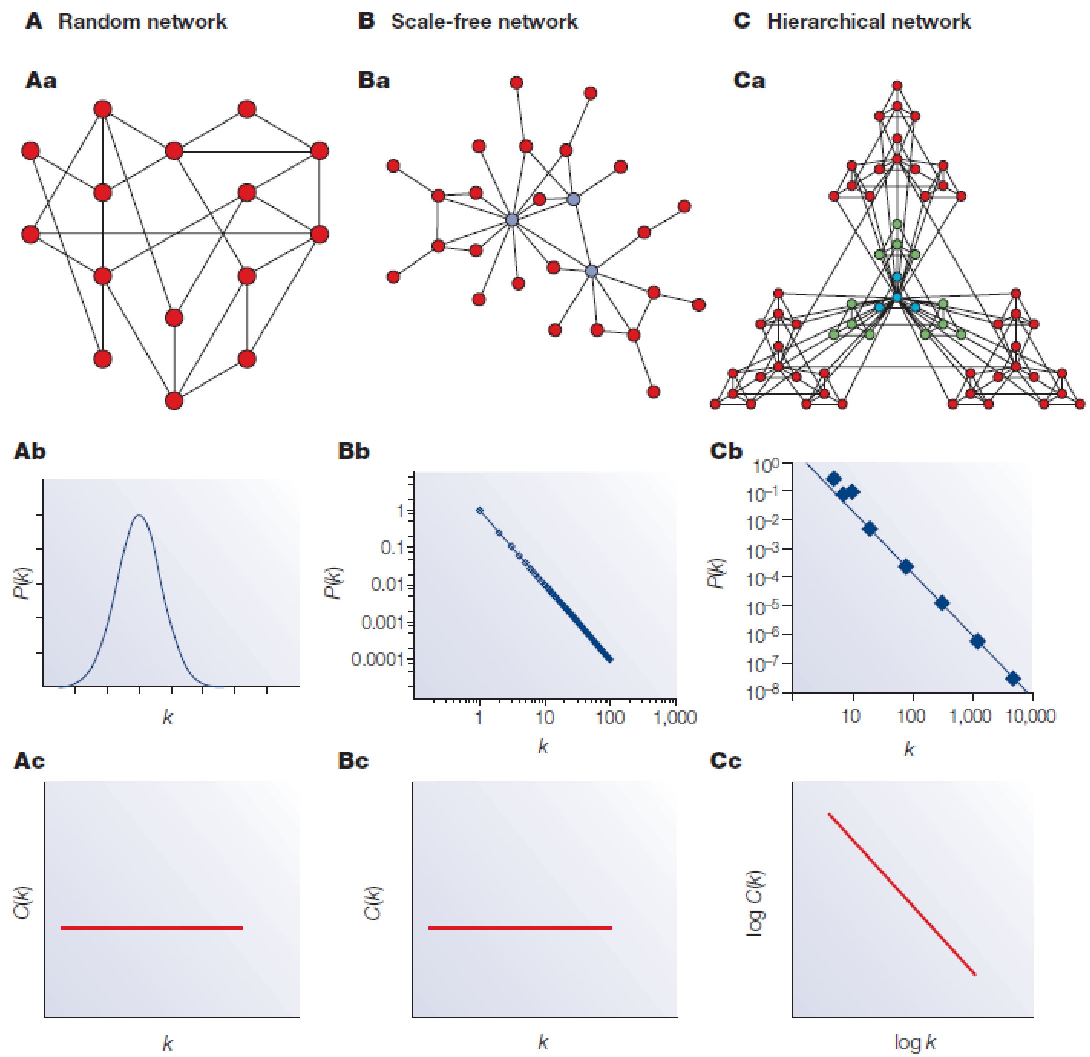
Que tan probable es encontrar triangulos (a.k.a. clausura transitiva) en la red ?




$$\langle C \rangle = \frac{13}{42} \approx 0.310$$

$$C_\Delta = \frac{3}{8} = 0.375$$

# Ejemplo: Modelos de juguete




Barabasi 2004 Nat Gen Network Biology

Red aleatoria

- Nodos enlazados con proba *p*
- *P(k)* Poisson (cola exponencial)
- C(k) independiente de k

Red libre-de-escala

- Crecimiento de red: *rich gets richer*
- *P(k) ~ $k^{-a}$*
- Red no posee estructura modular
- C(k) independiente de k

Red jerarquica

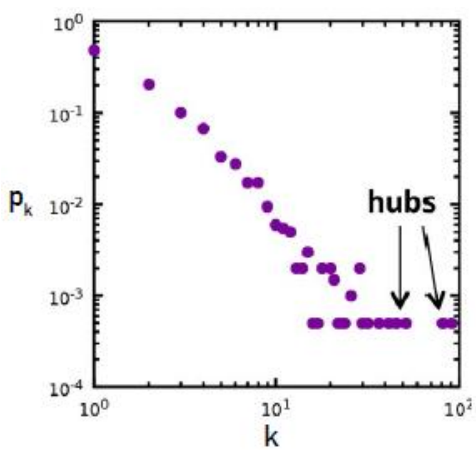
- Estructura jerarquica y modular
- *P(k) ~ $k^{-a}$*
- <C(k)>~0.6 (alto)
- C(k) ~ $k^{-\sim 1}$
- Estructuras densas comunicadas por algunos hubs

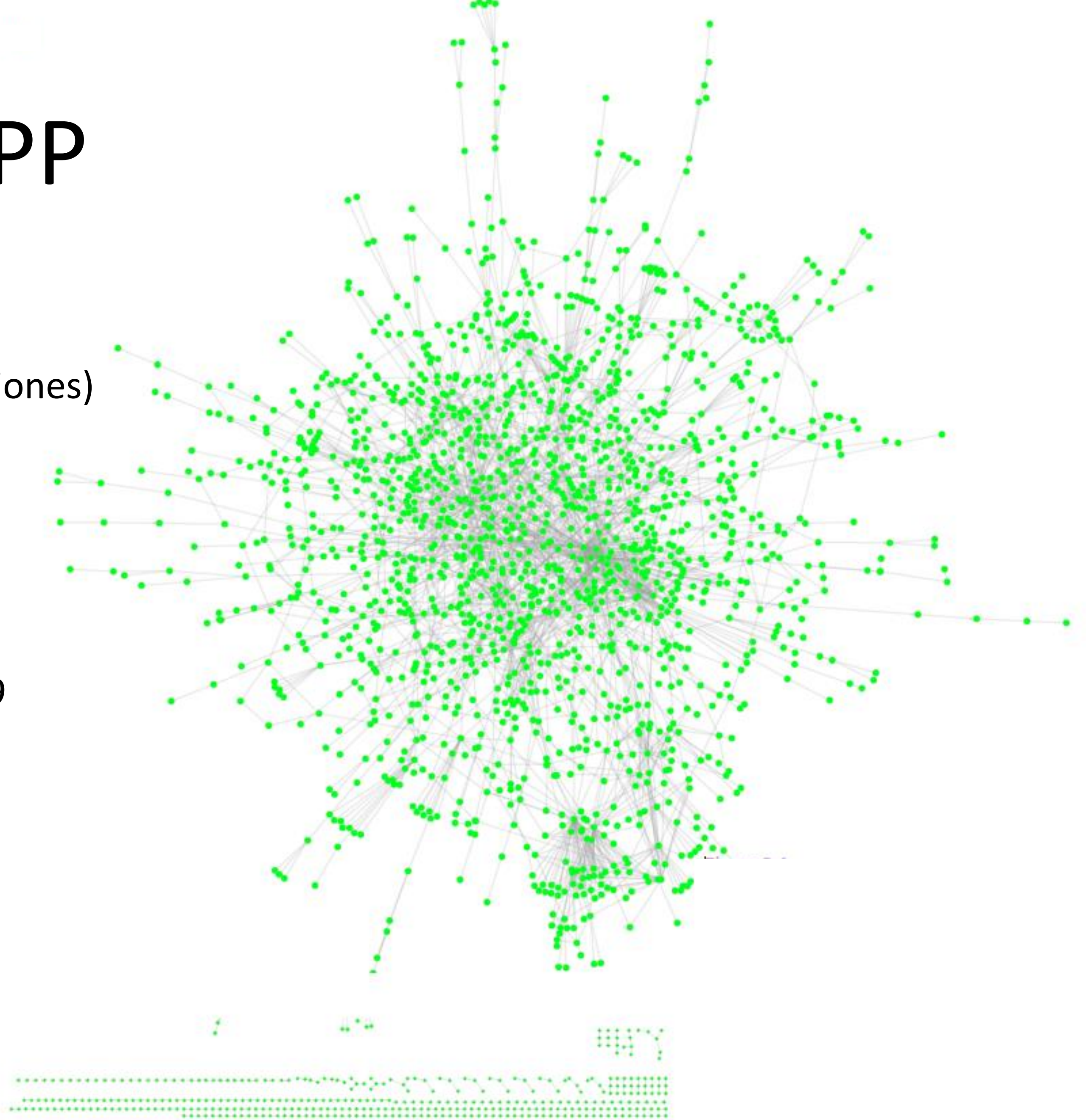
# Ejemplo RIPP

N=2018 nodos (proteinas)
M=2930 enlaces (interacciones)

**Grado medio**

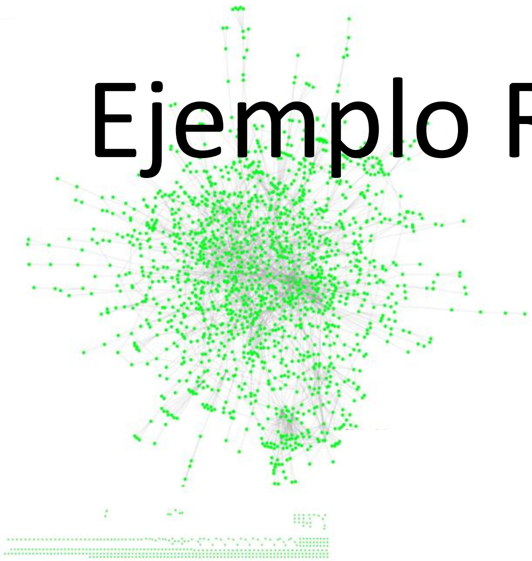
$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2M}{N} = 2.9$$



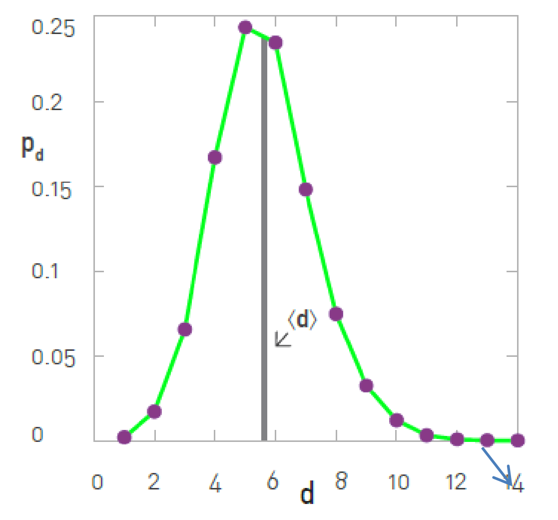
Usando BFS: componente gigante de 81% de nodos

# Ejemplo RIPP



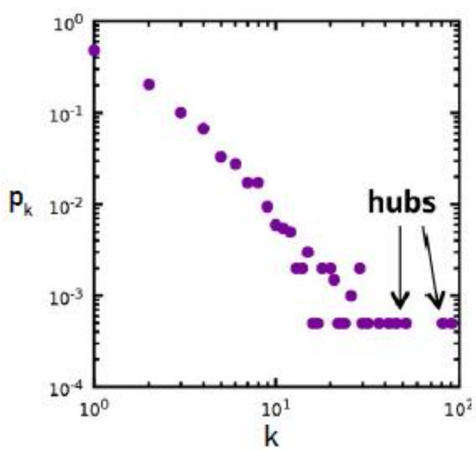
Distribucion de distancias

- distancia media = 5.61
- diámetro de la red=14



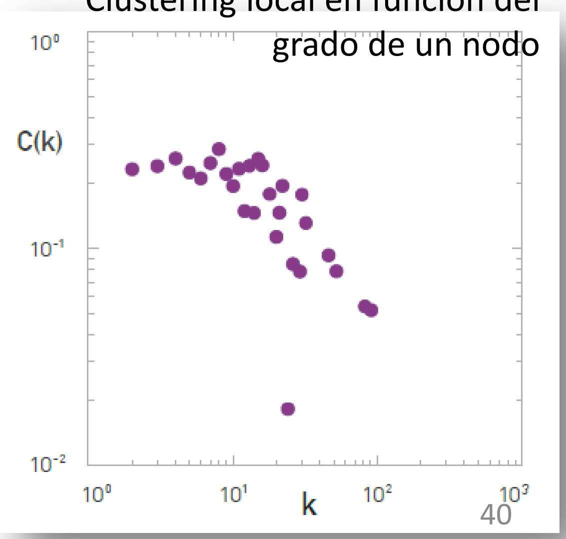

**Grado medio**

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2M}{N} = 2.9$$




- $<C_i>=0.12$
- C(k) decrece con k.
  - Nodos de bajo grado en entornos densos
  - hubs en entornos ralos
  - estructura jerárquica …

Clustering local en función del grado de un nodo



# Hablando del proyecto final…

- Vayan pensando y sopesando ideas ... pero no se apresuren en definir su proyecto
- La teoria de redes es una herramienta, no un fin en si mismo.. Entiendan lo mejor posible el sistema que desean analizar e identifiquen dos o tres preguntas concretas que deseen contestar.
- Jueguen/masajeen/exploren/entiendan lo mejor que puedan los datos, para saber que preguntarles.
- No pregunten por preguntar.
- No subestimen el proceso de obtencion, preparacion y manejo de datos. Es lo equivalente a preparar un experimento. Garbage-in -> garbage-out
- Cosas que en general pueden funcionar:

  Analisis de robustez

  Comparar evolucion temporal

  Correlacion estructura con campos definidos sobre nodos o enlaces

  Combinar / comparar redes

  Relacion entre propiedades topologicas (centralidad / puentes / ... ) y propiedades externas

  Aprendizaje semisupervisado

  Priorizacion

  …