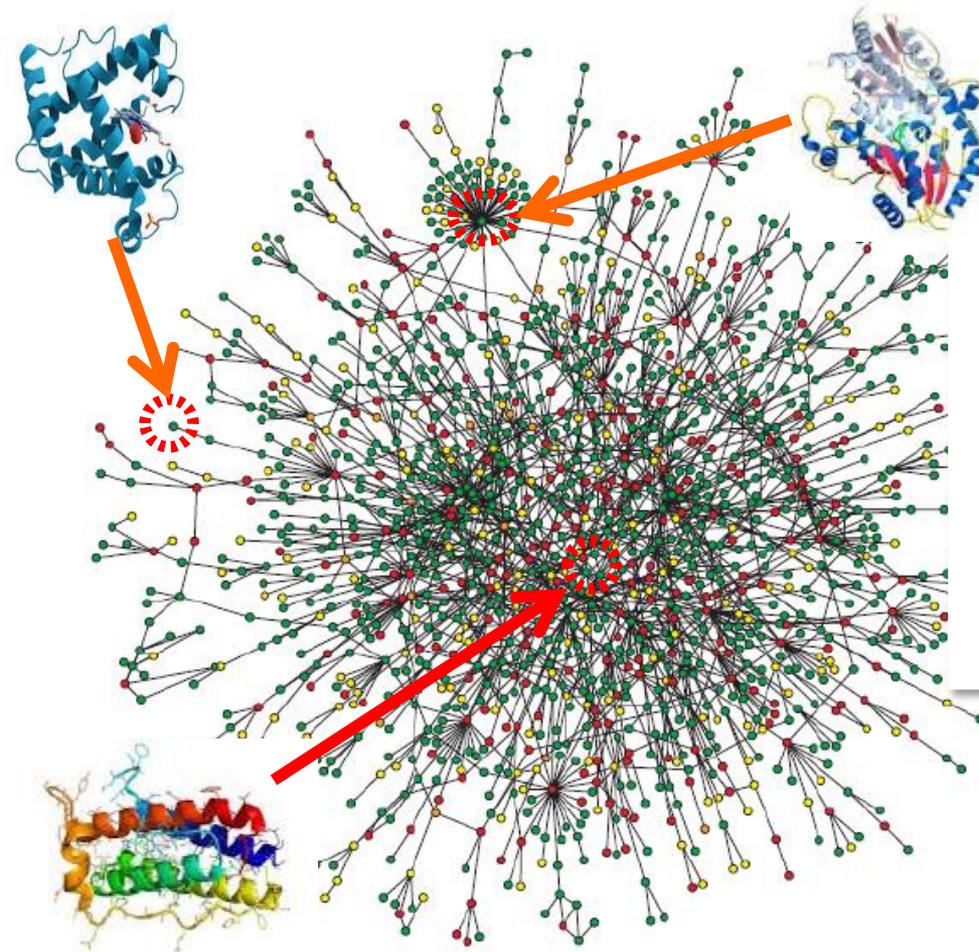


Centralidad

¿Qué tan *relevante* es un nodo de la red?

Topología y rol biológico



Redes Complejas

Degree
Eccentricity
Closeness
Betweenness
Bridging centrality
Eigenvalue centrality
Random walk centralities
....

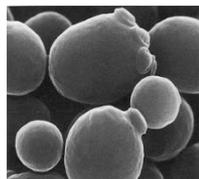


Biologia

Essentiality
Phenotypic variability
Biological function
Cancer related
Disease Related
...

Protein-protein interaction in yeast *S. cerevisiae*, (N=1870 and L=2240).

Jeong et al., Nature, 411, 41 (2001).



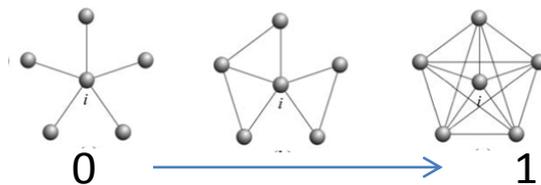
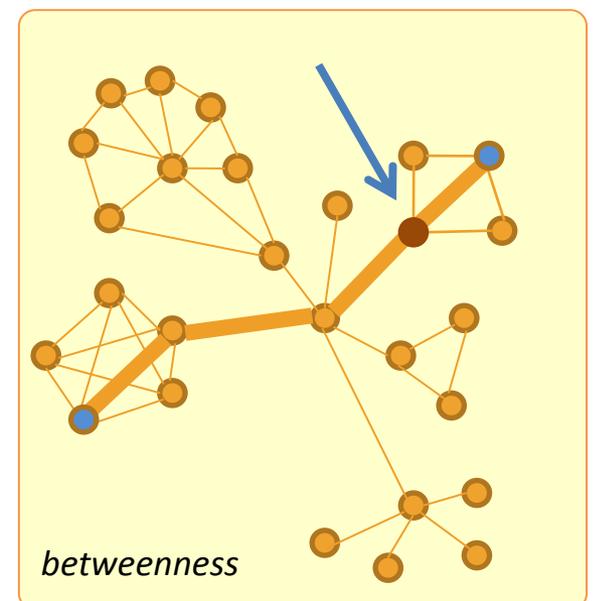
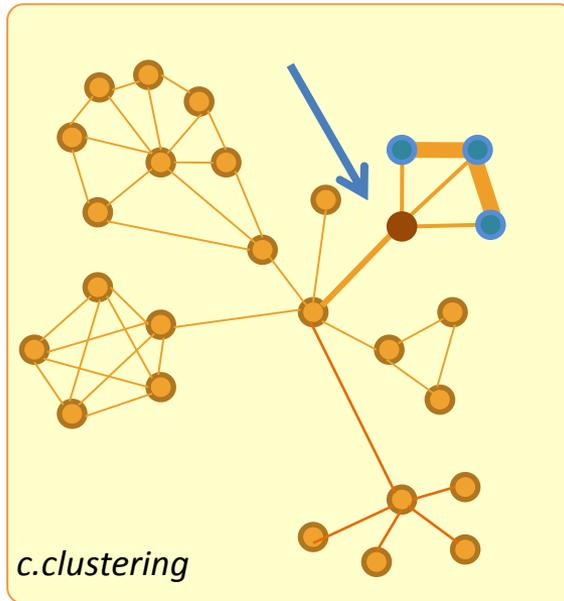
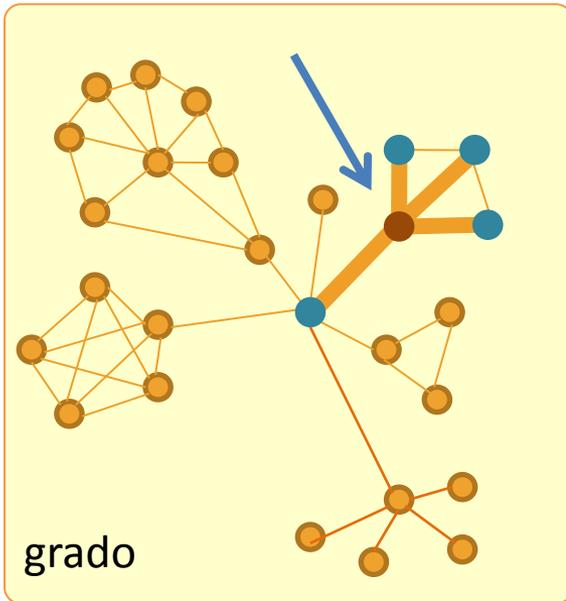
Propiedades topológicas: entorno local/global

Propiedades Locales

grado
 coef. Clustering
 centralidad 'random walk'
 ...

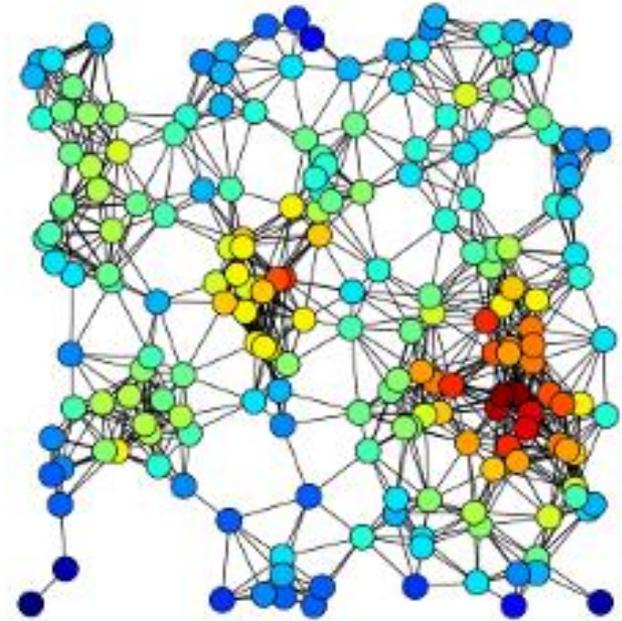
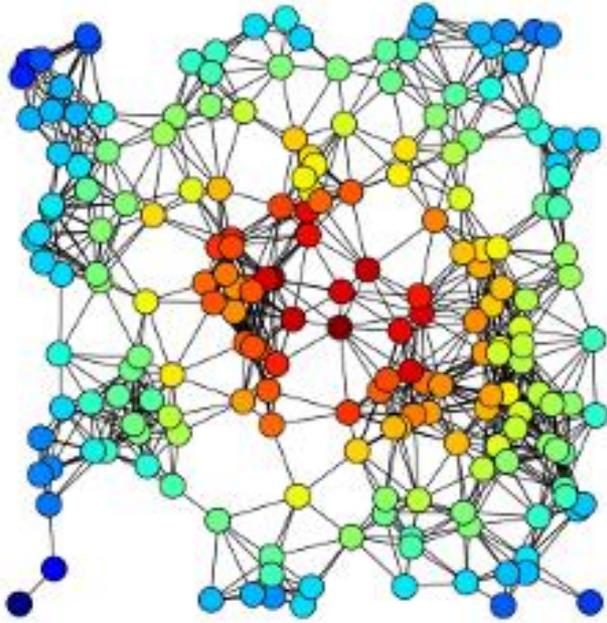
Propiedades Globales

betweenness
 centralidad de autovalores adyacencia
 cercania
 ...



$$be(v) = \sum_{s \neq t \neq v} \frac{\sigma_{s,t}^{shortest-path}(v)}{\sigma_{s,t}^{shortest-path}}$$

Que significa ser **importante**?



Noción de **flujo** sobre la red

Noción de **coesividad** sobre la red

Dime qué modelas con tu red y te diré qué es ser importante

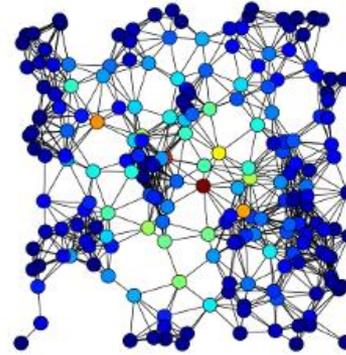
Centralidad

Dime qué modelas con tu red y te diré qué es ser importante

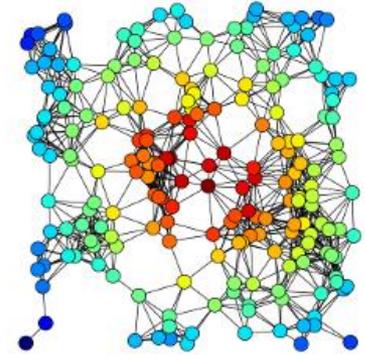
Noción de **flujo** sobre la red

Noción de **coesividad** sobre la red

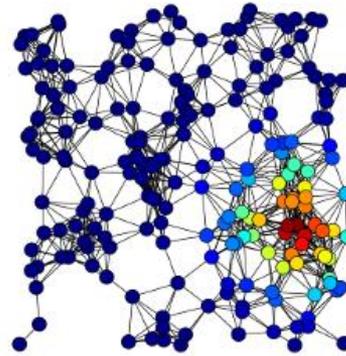
- a. intermedieatz
- b. cercanía
- c. autovalor
- d. grado
- e. centralidad armónica
- f. centralidad de Katz



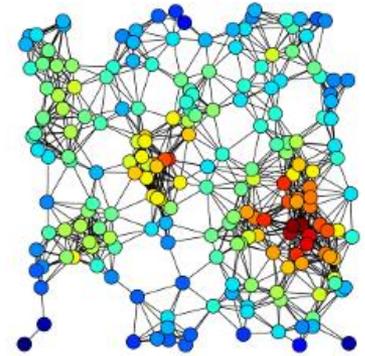
A



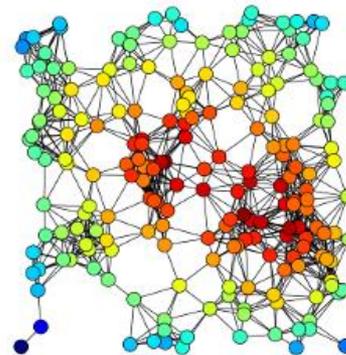
B



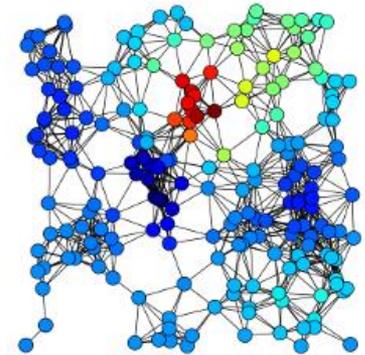
C



D



E



F

Centralidad y flujos

Dime qué modelas con tu red y te diré qué significa ser importante

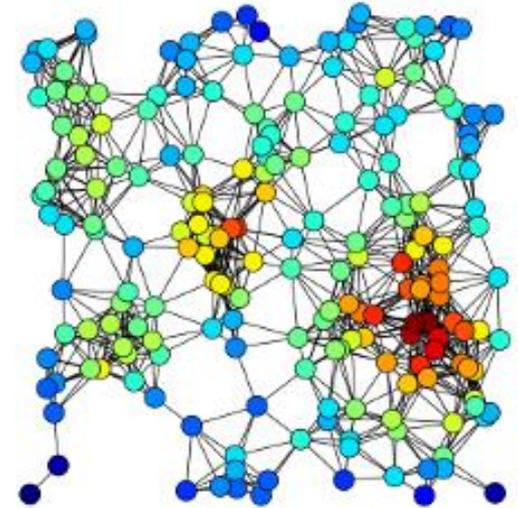
- Flujo de items indivisibles que están en un único lugar a un dado tiempo. Su difusión es por transferencia. Puede haber casos en que:
 - Billetes, libros: se mueven por el grafo **sin restricciones** acerca de repetir enlaces o nodos ya visitados (i.e. *walks*)
 - Ropa usada, *re-gifting*: se mueven por el grafo **sin repetir enlaces** (i.e. *trails*)
 - correo, gps: se mueven por el grafo desde un nodo origen hacia uno destino minimizando distancia recorrida (i.e. geodésicas).
- Flujo de entidades que pueden estar en varios lugares a la vez. Difusión por replicación.
 - Chismes: Transmisión boca-a-boca, en general no repite enlaces, pero si vértices (i.e. *trails*)
 - Campaña mailing: Transmisión por emisora (*broadcasting*), suele ser simultánea.
 - Ideas, creencias, modas, actitudes: influencia que se transmite boca-a-boca por replicación. Puede repetir enlaces (i.e. *walks*)
 - Procesos infecciosos, con inmunidad. Procede por replicación, pero sin visitar vértices (i.e. *paths*).

Centralidad de grado

- Utiliza directamente el **grado** de un nodo como medida de su importancia

$$Cent(n_i) = k_i$$

- En términos de flujo:
 - caracteriza efectos de influencia inmediata.
 - razonable por ejemplo para aplicar en procesos de duplicación paralela (prob de recibir algo que esta distribuido aleatoriamente por la red es proporcional al nro de contactos) o de caminatas al azar.
- En términos de cohesividad:
 - Hubs proveen atajos entre pares de nodos
- Asume linealidad: un nodo con el **doble** de vecinos que otro es **dos** veces mas importante
- Sólo utiliza información **local**

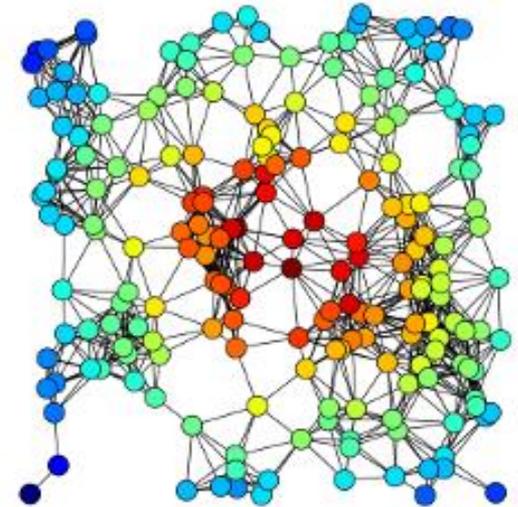


Centralidad de cercanía

- Asume que los nodos mas importantes son aquellos que pueden alcanzar fácilmente a más vecinos (están **cerca** de todos)

$$Closeness\ Centrality(n_i) = \left[\frac{\sum_{j \neq i} d_{geod}(n_i, n_j)}{N - 1} \right]^{-1}$$

(sin vecinos) $0 \leq Closeness\ Centrality \leq 1$ (vecino de todos)



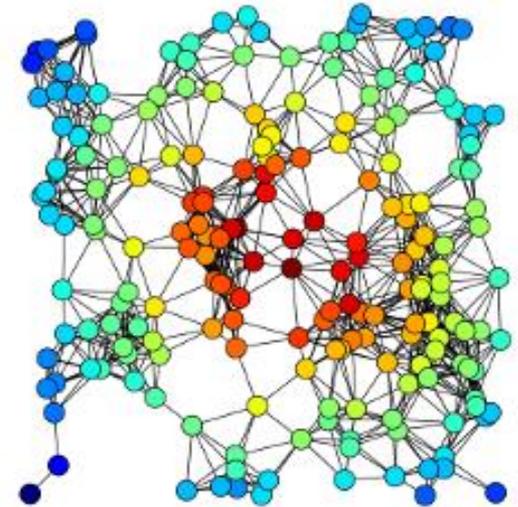
- nodos que en promedio reciben flujo transmitido por la red rápidamente:
 - red de flujo de información: ubicaciones ventajosas
 - red de contactos sexuales : ubicaciones riesgosas
 - en gral para flujos a lo largo de geodésicas o por duplicación en paralelo
 - no seria un buen indice para identificar, por ejemplo, quién recibirá info antes si el proceso de transmisión es tipo rumor (no va por caminos mas cortos)

Centralidad de cercanía

- Asume que los nodos más importantes son aquellos que pueden alcanzar fácilmente a más vecinos (están **cerca** de todos)

$$\text{Closeness Centrality}(n_i) = \left[\frac{\sum_{j \neq i} d_{\text{geod}}(n_i, n_j)}{N - 1} \right]^{-1}$$

(sin vecinos) $0 \leq \text{Closeness Centrality} \leq 1$ (vecino de todos)



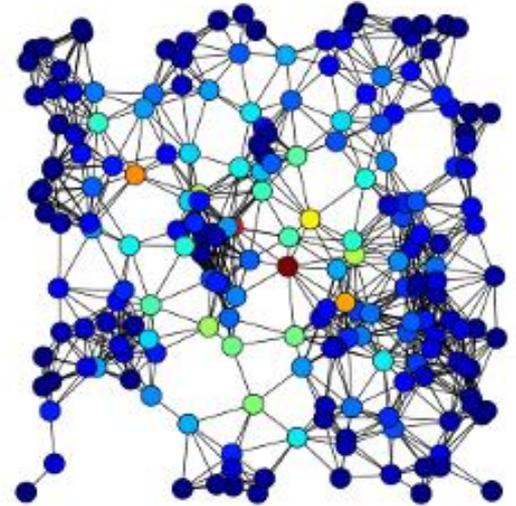
- En general tiene un rango dinámico chico (distancias geodésicas son chicas $1 < d_{\text{geod}} < o(\log N)$)
- No está bien definido para redes con más de una componente.
- Aún si calculáramos cada componente por separado no es trivial dar un ordenamiento global (centralidad en componentes más chicas tiende a ser mayor)

$$\text{Harmonic Closeness Centrality}(n_i) = \frac{1}{N - 1} \sum_{j \neq i} \frac{1}{d_{\text{geod}}(n_i, n_j)}$$

- Bien definido aún para redes con más de una componente.
- Se le da más peso a entorno local

Centralidad de **intermediatez** (*betweenness*)

Si hay un intercambio global de mensajes entre **todos los pares** de nodos de la red y si la información se intercambia por **geodésicas**: el nodo por el que pasarán más mensajes es aquél por el que pasan la mayor cantidad de geodésicas de la red.

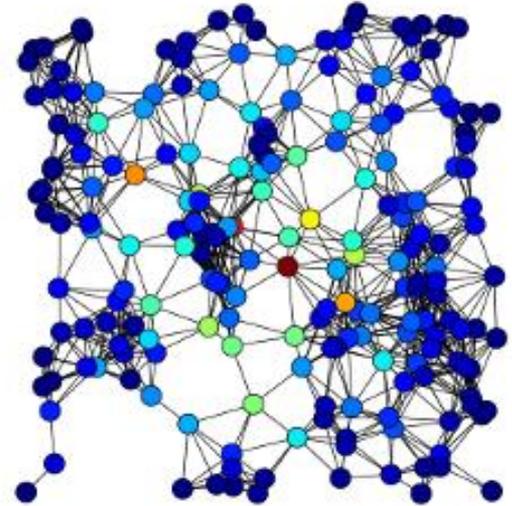


$$bet(n_i) = \sum_{r,s} \frac{x_{rs}^i}{x_{rs}}$$

← nro de geodésicas entre r y s que pasan por i
← nro de geodésicas entre r y s

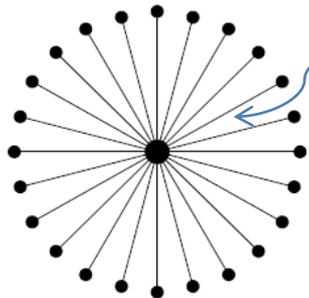
Centralidad de **intermediatez** (*betweenness*)

Si hay un intercambio global de mensajes entre **todos los pares** de nodos de la red y si la información se intercambia por **geodésicas**: el nodo por el que pasarán más mensajes es aquél por el que pasan la mayor cantidad de geodésicas de la red.



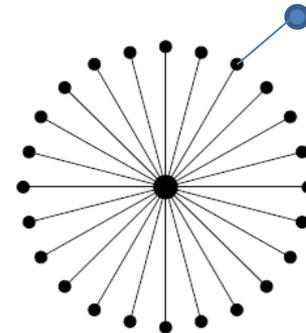
$$bet(n_i) = \sum_{r,s} \frac{x_{rs}^i}{x_{rs}}$$

← nro de geodésicas entre r y s que pasan por i
 ← nro de geodésicas entre r y s



$$bet_{max} = n^2 - (n - 1)$$

Participa en la geodesica de todos los pares menos en la de los $n-1$ perifericos consigo mismos

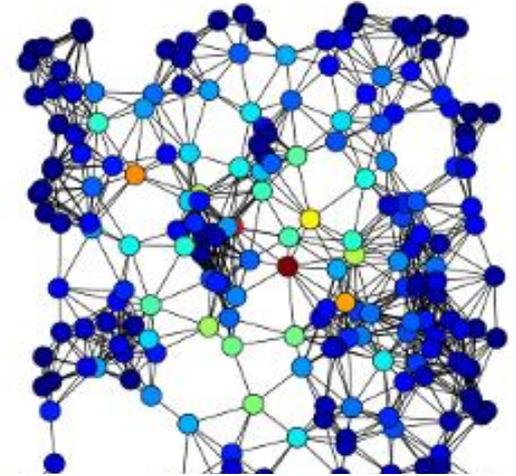


$$bet_{min} = 2n - 1$$

Participa en las geodesicas (ida y vuelta) del azul con el resto $2(n-1)$ mas la del azul consigo mismo.

Centralidad de **intermediatez** (*betweenness*)

Si hay un intercambio global de mensajes entre **todos los pares** de nodos de la red y si la información se intercambia por **geodésicas**: el nodo por el que pasarán más mensajes es aquél por el que pasan la mayor cantidad de geodésicas de la red.



$$bet(n_i) = \sum_{r,s} \frac{x_{rs}^i}{x_{rs}}$$

← nro de geodésicas entre r y s que pasan por i
 ← nro de geodésicas entre r y s

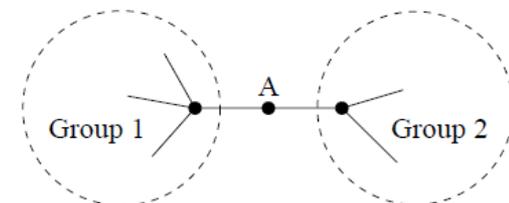
$$bet_{min} = 2n - 1$$

$$bet_{max} = n^2 - (n - 1)$$

$$\left. \begin{array}{l} bet_{min} = 2n - 1 \\ bet_{max} = n^2 - (n - 1) \end{array} \right\} \frac{bet_{max}}{bet_{min}} \sim \frac{1}{2}n \quad (\text{amplio rango dinámico})$$

Es una medida global

Es una medida de intermediatez más que de conectividad



Centralidad de autovalor

Nodos importantes tienen **vecinos** importantes

- x_i (importancia de nodo- i) depende de x_j (importancia de nodo- j vecino). La importancia se *transmite* por la red.
- la manera más simple es asumir linealidad

$$x_i' = \sum_{j=1}^N a_{ij} x_j \quad \mathbf{x}' = \mathbf{A} \mathbf{x}$$

- el nivel de importancia se establece recursivamente en la red

$$\mathbf{x}'' = \mathbf{A} \mathbf{x}' = \mathbf{A} \mathbf{A} \mathbf{x}$$

⋮

Despues de s iteraciones: $\mathbf{x}(s) = \overbrace{\mathbf{A} \mathbf{A} \dots \mathbf{A}}^s \mathbf{x}(0)$

$$\mathbf{x}(s) = \mathbf{A}^s \mathbf{x}(0)$$

Centralidad de autovalor

Nodos importantes tienen **vecinos** importantes

- x_i (importancia de nodo- i) depende de x_j (importancia de nodo- j vecino). La importancia se *transmite* por la red.
- la manera más simple es asumir linealidad

$$x_i' = \sum_{j=1}^N a_{ij} x_j \quad \mathbf{x}' = \mathbf{A} \mathbf{x}$$

- el nivel de importancia se establece recursivamente en la red

$$\mathbf{x}'' = \mathbf{A} \mathbf{x}' = \mathbf{A} \mathbf{A} \mathbf{x}$$

⋮

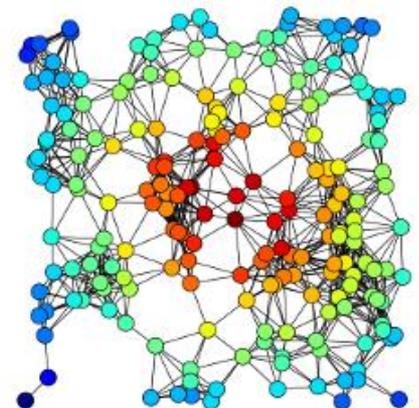
Después de s iteraciones:

$$\mathbf{x}(s) = \overbrace{\mathbf{A} \mathbf{A} \dots \mathbf{A}}^s \mathbf{x}(0)$$

\mathbf{x} solución autoconsistente

$$\mathbf{x} = \mathbf{A}^s \mathbf{x}$$

$s \gg 1$

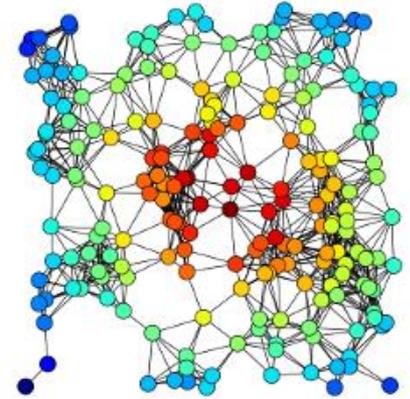


\mathbf{x} es un autovector de \mathbf{A}

Centralidad de autovalor

Busco x solución autoconsistente

$$\mathbf{x}(s) = \mathbf{A}^s \mathbf{x}(0)$$



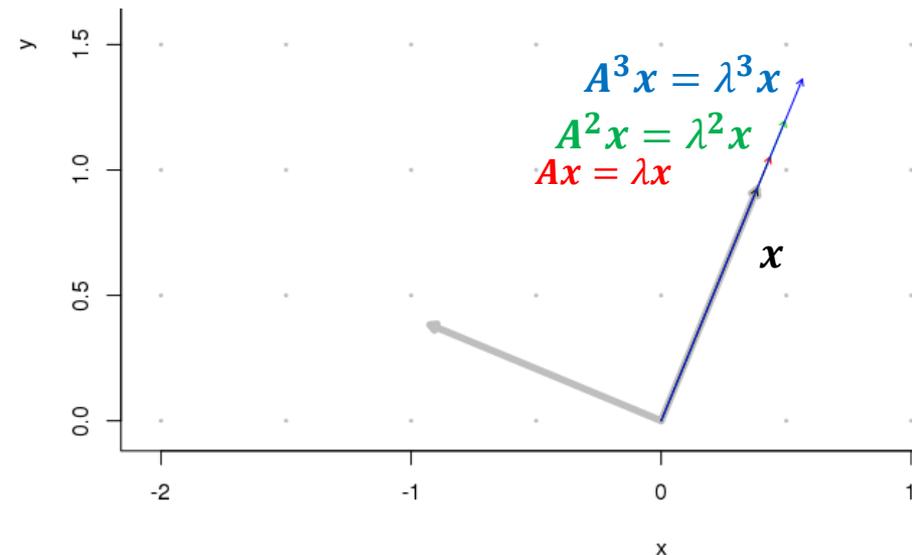
Si inicialmente arranco de un autovector: $\mathbf{x}(0) = \mathbf{v}_k$

$$\mathbf{x}(s) = \overbrace{\mathbf{A}\mathbf{A} \dots \mathbf{A}}^{s-1} \mathbf{A} \mathbf{v}_k = \lambda_k \overbrace{\mathbf{A}\mathbf{A} \dots \mathbf{A}}^{s-1} \mathbf{v}_k$$

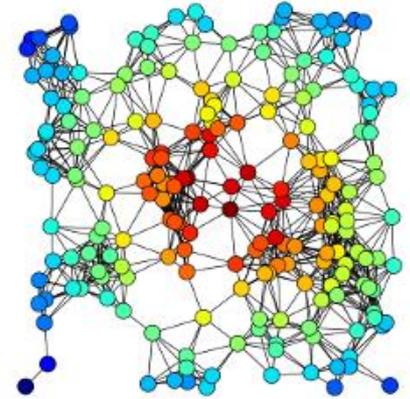
$$\mathbf{x}(s) = \lambda_k^s \mathbf{v}_k$$

Por lo que la **importancia relativa** de cada nodo se podría inferir de las componentes de \mathbf{v}_k

si \mathbf{x} inicial fuera justo un **autovector**, por ejemplo \mathbf{v}_k , de la matriz de adyacencia \mathbf{A}



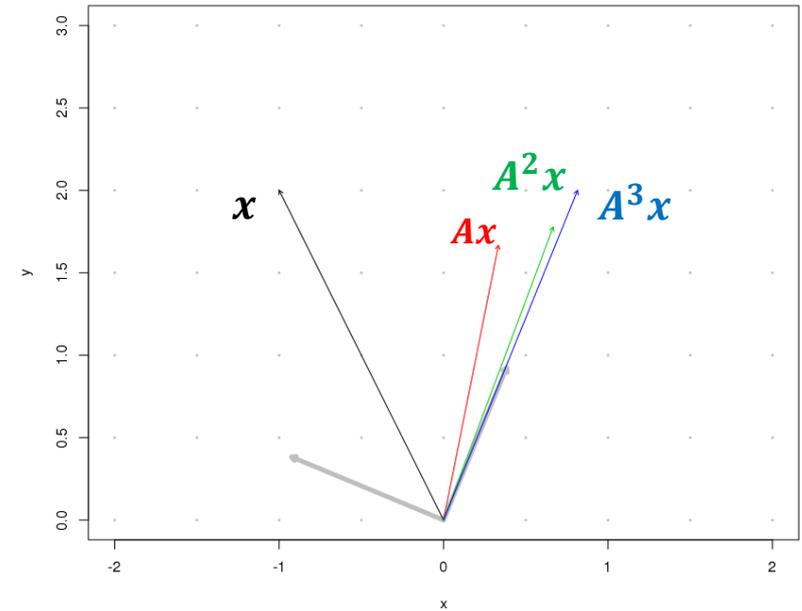
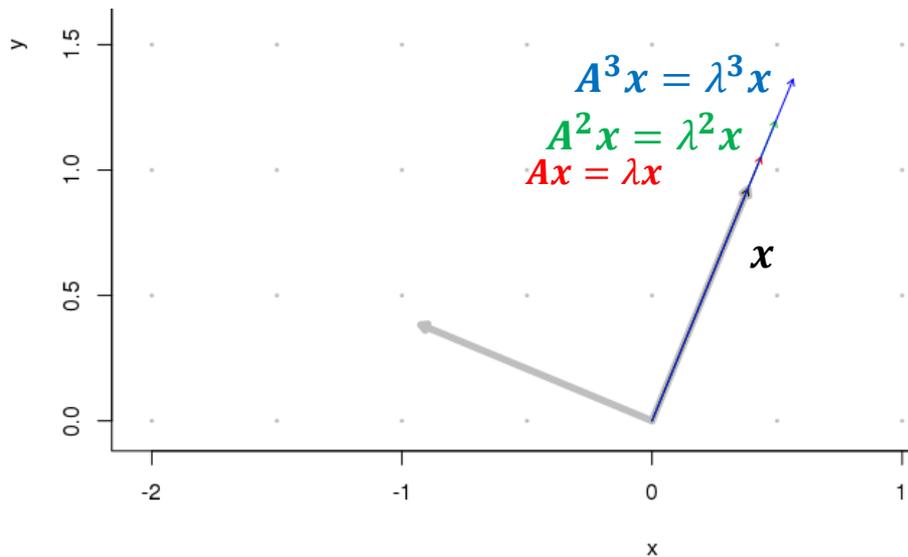
Centralidad de autovalor



$$\mathbf{x}(s) = \mathbf{A}^s \mathbf{x}(0)$$

si \mathbf{x} inicial fuera justo un **autovector**, por ejemplo \mathbf{v}_k , de la matriz de adyacencia \mathbf{A}

En el caso general $\mathbf{x}(0)$ no coincide con una dirección privilegiada asociada con \mathbf{A}



Centralidad de autovalor

Para ver como una **matriz** transforma vectores es útil pensar en término de sus **autovectores** \mathbf{v}_i (se transforman muy fácil: $\mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{v}_i$). El vector $\mathbf{x}(0)$ puede ser escrito como combinación lineal de los $\{\mathbf{v}_i\}$

$$\mathbf{x}(0) = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots = \sum_i c_i \mathbf{v}_i$$

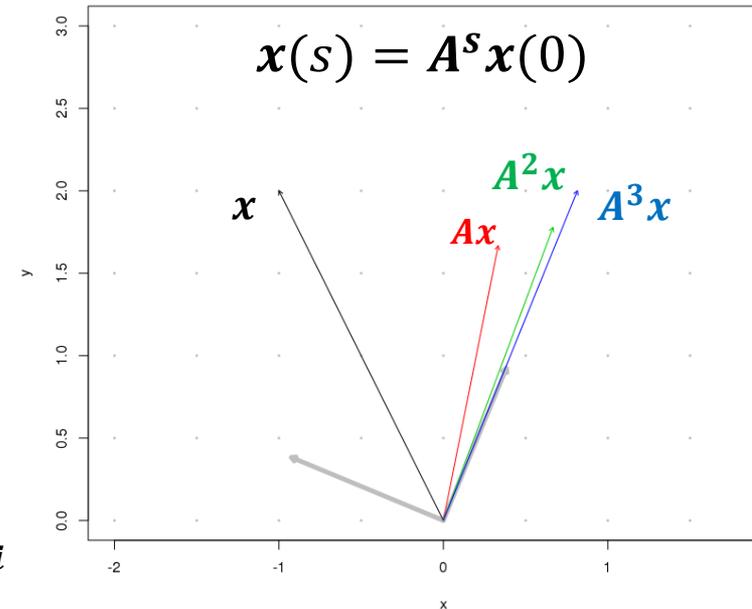
$$\mathbf{A}^s \mathbf{x}(0) = \sum_i c_i \mathbf{A}^s \mathbf{v}_i = \sum_i c_i \lambda_i^s \mathbf{v}_i = \lambda_*^s \sum_i c_i \left(\frac{\lambda_i}{\lambda_*} \right)^s \mathbf{v}_i$$

↑ mayor autovalor

$$\mathbf{x}(s) = \lambda_*^s \sum_i c_i \overbrace{\left(\frac{\lambda_i}{\lambda_*} \right)^s}^{<1} \mathbf{v}_i$$

$$\mathbf{x}(s) \sim \lambda_*^s c_i \mathbf{v}_*$$

$$s \rightarrow \infty$$



encontrar \mathbf{x} autoconsistente se reduce a encontrar el autovector de \mathbf{A} asociado al mayor autovalor

$$\mathbf{A} \mathbf{x} = \lambda_* \mathbf{x}$$

Centralidad de autovalor

$$\mathbf{A}\mathbf{x}^* = \lambda_* \mathbf{x}^*$$

\mathbf{x} se conoce como **centralidad de Bonacich**

$$x_i^* = \frac{1}{\lambda_*} \sum_j A_{ij} x_j^*$$

La centralidad de un nodo depende de la de sus vecinos.

Es \mathbf{x}^* una buena centralidad? Quien me asegura que sus componentes sean todas positivas, por ejemplo?

Teorema de Perron-Frobenius

Sea \mathbf{A} una matriz de $N \times N$ con elementos no negativos. Entonces:

- \mathbf{A} tiene un autovalor real dominante: $\lambda_1 > |\lambda_i|$ para $i = 2, \dots, N$
- λ_1 corresponde a un autovalor para el que se puede elegir un autovector cuyas **componentes son no-negativas**
- Además vale que λ_1 está acotado por la mínima y máxima suma por columna de \mathbf{A}

$$\min_j \left(\sum_{i=1}^N a_{ij} \right) \leq \lambda_1 \leq \max_j \left(\sum_{i=1}^N a_{ij} \right)$$

Centralidad de autovalor

$$A\mathbf{x}^* = \lambda_* \mathbf{x}^*$$

\mathbf{x} se conoce como **centralidad de Bonacich**

$$x_i^* = \frac{1}{\lambda_*} \sum_j A_{ij} x_j^*$$

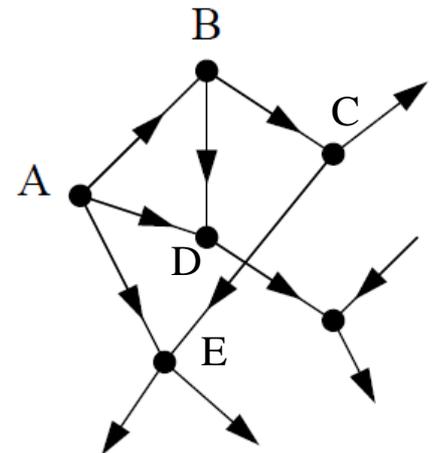
La centralidad de un nodo depende de la de sus vecinos.

Para redes dirigidas (**A** matriz no simétrica) se usa la misma definición de centralidad. Esto equivale a considerar el autovector-por-derecha de mayor autovalor.

Sin embargo puede haber problemas:

- x_A no tiene vecinos, entonces $x_A=0$
- pero también $x_B=0$, $x_C=0$, $x_D=0$, $x_E=0$

Sólo nodos de una **componente fuertemente conexa** de dos o más vértices (o la **out-component** de dicha componente) pueden tener centralidad de Bonacich no nula. **Potencial problema (!)**



Centralidad de Katz

$$x_i = \alpha \sum_j A_{ij} x_j + \beta$$

Matricialmente

centralidad de base.

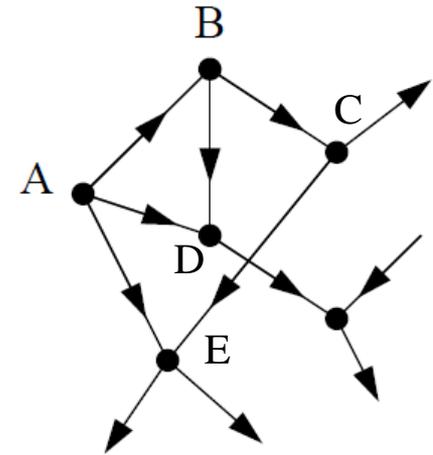
Gran diferencia para nodos con $k_{in}=0$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{x} + \beta \mathbf{1}$$

$$\mathbf{x} - \alpha \mathbf{A} \mathbf{x} = \beta \mathbf{1}$$

$$(\mathbf{I} - \alpha \mathbf{A}) \mathbf{x} = \beta \mathbf{1}$$

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A})^{-1} \beta \mathbf{1}$$

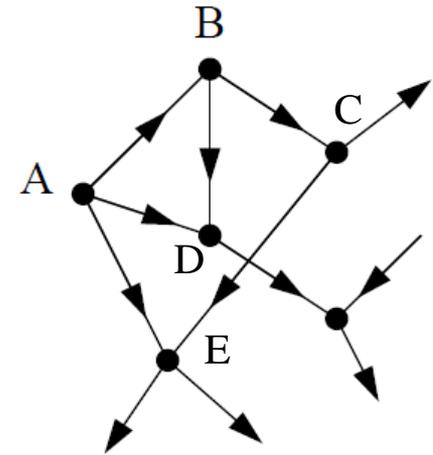


Usualmente se toma $\beta=1$ y se debe especificar α , que controla la importancia relativa del primer término respecto del segundo

Centralidad de Katz

$$x_i = \alpha \sum_j A_{ij} x_j + 1$$

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A})^{-1} \mathbf{1}$$



Se debe especificar α , como lo elijo?

- $\alpha = 0$ resulta en distribución uniforme de centralidad
- $\alpha > 0$ le doy más peso al primer término
- **PERO OJO:** para algún $\alpha > 0$ la matriz $\mathbf{I} - \alpha \mathbf{A}$ puede no ser inversible.
Esto sucede cuando $\det(\mathbf{I} - \alpha \mathbf{A}) = 0$, o lo que es lo mismo, cuando:

$$\det(\mathbf{A} - \alpha^{-1} \mathbf{I}) = 0$$

La condición para identificar α problemáticos asociada a la ecuación de autovalores para \mathbf{A} (!)

Para α creciente el *primer* problema lo encuentro cuando $\alpha^{-1} = \lambda_*$

Por lo tanto, se suele elegir α en el rango $0 < \alpha < 1/\lambda_*$

Centralidad de Katz

Estimación de centralidad

Algebraica:

$$\mathbf{x} = (\mathbf{I} - \alpha\mathbf{A})^{-1}\beta\mathbf{1}$$

Iterando:

$$x_i^{(t+1)} = \alpha \sum_j A_{ij} x_j^{(t)} + \beta$$

Costo
computacional

$o(n^3)$

$o(M.r)$

M:#enlaces, i.e. items no nulos de \mathbf{A}
r: #iteraciones

Centralidad de Katz generalizada

Asignamos una relevancia-a-priori, que puede ser diferente para cada nodo, utilizando **información externa a la red**

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_N)$$

$$x_i^{(t+1)} = \alpha \sum_j A_{ij} x_j^{(t)} + \beta_i \quad \boldsymbol{x} = (\boldsymbol{I} - \alpha \boldsymbol{A})^{-1} \boldsymbol{\beta}$$

Centralidad PageRank

Katz generalizado:

$$x_i^{(t+1)} = \alpha \sum_j A_{ij} x_j^{(t)} + \beta_i$$

Un posible problema: nodos con **alta centralidad** distribuyen ese valor (**alto**) a todos sus contactos k_{out} . Quizas no sea eso lo que uno quiere.

‘diluye’ la contribución del nodo-j



$$x_i^{(t+1)} = \alpha \sum_j \frac{1}{k_j^{out}} A_{ij} x_j^{(t)} + \beta_i$$

Centralidad PageRank

$$x_i^{(t+1)} = \alpha \sum_j \frac{1}{k_j^{out}} A_{ij} x_j^{(t)} + \beta_i$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{D}^{-1} \mathbf{x} + \boldsymbol{\beta}$$

$$[\mathbf{D}^{-1}]_{ii} = 1/k_i^{out}$$

$$(\mathbf{I} - \alpha \mathbf{A} \mathbf{D}^{-1}) \mathbf{x} = \boldsymbol{\beta}$$

$$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A} \mathbf{D}^{-1})^{-1} \boldsymbol{\beta}$$

$$= (\mathbf{D} \mathbf{D}^{-1} - \alpha \mathbf{A} \mathbf{D}^{-1})^{-1} \boldsymbol{\beta}$$

$$= [(\mathbf{D} - \alpha \mathbf{A}) \mathbf{D}^{-1}]^{-1} \boldsymbol{\beta}$$

$$\mathbf{x} = \mathbf{D} (\mathbf{D} - \alpha \mathbf{A})^{-1} \boldsymbol{\beta}$$

razonando como antes $\alpha < \lambda_{\text{dominante}} \text{ de } \mathbf{A} \mathbf{D}^{-1}$.
igual....se suele usar $\alpha=0.85$

Centralidad PageRank sin bias

$$x_i^{(t+1)} = \sum_j \frac{1}{k_j^{out}} A_{ij} x_j^{(t)} + \beta_i$$

$$\mathbf{x} = \mathbf{AD}^{-1}\mathbf{x}$$

$$[\mathbf{D}^{-1}]_{ii} = 1/k_i^{out}$$

- Estoy buscando el autovector de \mathbf{AD}^{-1} asociado a $\lambda=1$
- Para redes no-dirigidas es fácil ver que $x_j = k_j$ es la solución autoconsistente buscada.
- O sea PageRank sin bias equivale a centralidad de grado.

Centralidades de recurrencia (resumen)

$$x_i^{(t+1)} = \alpha \sum_j \frac{1}{k_j^{out}} A_{ij} x_j^{(t)} + \beta_i$$

$$\mathbf{x} = \mathbf{D}(\mathbf{D} - \alpha\mathbf{A})^{-1}\boldsymbol{\beta}$$

PageRank centrality

$$x_i^{(t+1)} = \sum_j \frac{1}{k_j^{out}} A_{ij} x_j^{(t)}$$

$$\mathbf{x} = \mathbf{A}\mathbf{D}^{-1}\mathbf{x}$$

Degree centrality

$$x_i^{(t+1)} = \alpha \sum_j A_{ij} x_j^{(t)} + \beta_i$$

$$\mathbf{x} = (\mathbf{I} - \alpha\mathbf{A})^{-1}\boldsymbol{\beta}$$

Katz centrality

$$x_i^{(t+1)} = \sum_j A_{ij} x_j^{(t)}$$

$$\mathbf{x} = \lambda_*^{-1}\mathbf{A}\mathbf{x}$$

**Eigenvector (Bonacich)
centrality**

Conectores (*hubs*) y Autoridades

- Hasta ahora, en **redes dirigidas**, los algoritmos que vimos asignaban más nivel de **centralidad** a aquellos nodos que **recibían** muchas conexiones
- Puede ser de interés, sin embargo, identificar nodos que actúen como buenos *reviews*. Implica idea de centralidad asociada con k_{out} hacia nodos relevantes (e.g. página con links hacia otras páginas relevantes)

Idea [Kleinberg, Authoritative Sources in a Hyperlinked Environment, 1999 10122 citas]:

Generalizar la centralidad de Bonacich y permitir que cada nodo tenga dos atributos

1. **Autoridad:** que tanto conocimiento, información, etc, tiene un nodo respecto a un tema
2. **Conectividad:** que tanto un nodo es capaz de encontrar información sobre un tema

En este esquema:

- Los nodos son caracterizados simultáneamente como **conectores y autoridades**
- Los mejores conectores apuntan a las mejores autoridades
- Algoritmo HITS (*Hyperlink-induced topic search*)

Conectores (*hubs*) y Autoridades

Para un nodo, dos centralidades: x_i (autoridad), y_i (conector)

$$x_i = \alpha \sum_j A_{ij} y_j$$

$$y_i = \beta \sum_j A_{ji} x_j$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{y}$$

$$\mathbf{y} = \beta \mathbf{A}^T \mathbf{x}$$

pero entonces

$$\mathbf{x} = (\alpha\beta) \mathbf{A} \mathbf{A}^T \mathbf{x}$$

$$\mathbf{y} = (\alpha\beta) \mathbf{A}^T \mathbf{A} \mathbf{y}$$

$$\mathbf{A} \mathbf{A}^T \mathbf{x} = \lambda \mathbf{x}$$

$$\mathbf{A}^T \mathbf{A} \mathbf{y} = \lambda \mathbf{y}$$

con $\lambda = \alpha\beta$

fijarse además que:

- $\mathbf{A} \mathbf{A}^T$ es la matriz de co-citas, por lo que la **centralidad de autoridad** puede verse como **centralidad de autovector de la red de co-citas**
- $\mathbf{A}^T \mathbf{A}$ es la matriz bibliográfica, por lo que la **centralidad de conectividad** puede verse como **centralidad de autovector de la red bibliográfica**

Conectores (*hubs*) y Autoridades

Para un nodo, dos centralidades: x_i (autoridad), y_i (conector)

$$x_i = \alpha \sum_j A_{ij} y_j \qquad y_i = \beta \sum_j A_{ji} x_j$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{y} \qquad \mathbf{y} = \beta \mathbf{A}^T \mathbf{x}$$

pero entonces

$$\mathbf{x} = (\alpha\beta) \mathbf{A} \mathbf{A}^T \mathbf{x} \qquad \mathbf{y} = (\alpha\beta) \mathbf{A}^T \mathbf{A} \mathbf{y}$$

$$\mathbf{A} \mathbf{A}^T \mathbf{x} = \lambda \mathbf{x} \qquad \mathbf{A}^T \mathbf{A} \mathbf{y} = \lambda \mathbf{y} \qquad \text{con } \lambda = \alpha\beta$$

fijarse además que:

$$\mathbf{A}^T (\mathbf{A} \mathbf{A}^T \mathbf{x}) = \lambda \mathbf{A}^T \mathbf{x}$$

$$\mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{x}) = \lambda (\mathbf{A}^T \mathbf{x})$$

por lo que $\mathbf{y} = (\mathbf{A}^T \mathbf{x})$ o sea, basta con resolver la ecuación de autovalores de $\mathbf{A}^T \mathbf{A}$

Conectores (*hubs*) y Autoridades

Para un nodo, dos centralidades: x_i (autoridad), y_i (conector)

$$x_i = \alpha \sum_j A_{ij} y_j$$

$$y_i = \beta \sum_j A_{ji} x_j$$

$$\mathbf{x} = \alpha \mathbf{A} \mathbf{y}$$

$$\mathbf{y} = \beta \mathbf{A}^T \mathbf{x}$$

pero entonces

$$\mathbf{x} = (\alpha\beta) \mathbf{A} \mathbf{A}^T \mathbf{x} \quad \mathbf{y} = (\alpha\beta) \mathbf{A}^T \mathbf{A} \mathbf{y}$$

$$\mathbf{A} \mathbf{A}^T \mathbf{x} = \lambda \mathbf{x} \quad \mathbf{A}^T \mathbf{A} \mathbf{y} = \lambda \mathbf{y} \quad \text{con } \lambda = \alpha\beta$$

- $\mathbf{A} \mathbf{A}^T$ es una matriz simétrica, semi-definida positiva, por lo que sus autovalores son mayores que cero.
- De hecho, sus autovalores son los valores-singulares de \mathbf{A} elevados al cuadrado
- Perron-Frobenius: el autovector del autovalor dominante tiene componentes no-negativas