



Minería de Medios Sociales

Máster en Ciencia de Datos e Ingeniería de Computadores

Bloque I: Redes Sociales y Ciencia de Datos en Redes

Sesión I.2: Análisis de Redes Sociales

Oscar Cordon García

Dpto. Ciencias de la Computación e Inteligencia Artificial. Universidad de Granada
ocordon@decsai.ugr.es

ANÁLISIS DE REDES SOCIALES

El **Análisis de Redes Sociales** (*Social Network Analysis*, SNA) se centra en el descubrimiento de patrones de interacción entre **actores sociales** en redes sociales

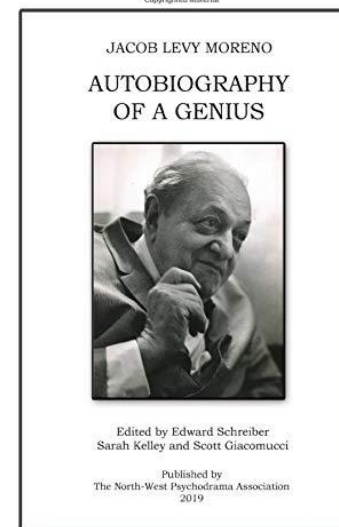
Es un área de investigación metodológica interdisciplinar con contribuciones de la **Sociología**, la **Psicología Social**, la **Antropología**, la Física, las Matemáticas y la Informática, entre otras

Los orígenes del SNA, como una base para el desarrollo de conceptos sociológicos útiles, puede fecharse a **comienzos de la década de 1930**, cuando Moreno desarrolló el **enfoque sociométrico** como una forma de conceptualizar la estructura de las relaciones sociales establecidas entre pequeños grupos de individuos

Estos lazos interpersonales entre miembros de un grupo se representaban usando los denominados **sociogramas**, que pueden definirse como gráficos en los que los individuos se representan como nodos y las relaciones entre ellos como líneas

Esos diagramas resultaron ser muy útiles para descubrir las estructuras ocultas de los grupos mediante la identificación de protagonistas, alianzas y subgrupos, etc.

Moreno, Jacob Levy. Who Shall Survive? Foundations of sociometry, group psychotherapy and sociodrama. New York: Beacon House; 1953



ANÁLISIS DE REDES SOCIALES (*SOCIAL NETWORK ANALYSIS*)

El objetivo principal del **Análisis de Redes Sociales** (SNA) es examinar tanto los contenidos como los **patrones de relación en redes sociales para entender las relaciones entre los actores** y las implicaciones de esas relaciones

Es un área interdisciplinar entre las ciencias sociales, la estadística, la teoría de grafos, la complejidad y la informática

Son tareas habituales del SNA:

- identificar los actores más **influyentes**, **prestigiosos** o **centrales** de la red, mediante medidas estadísticas,
- identificar **hubs** y **autoridades**, usando algoritmos de análisis de enlaces en redes dirigidas,
- determinar patrones de interacción comunes entre actores, mediante **medidas de niveles de interacción**, y
- descubrir grupos de actores cohesionados, con técnicas de detección de **comunidades**

EJEMPLOS DE APLICACIÓN DEL ANÁLISIS DE REDES SOCIALES

- **Marketing viral: maximización de la difusión “boca a boca” de productos** de una compañía dirigiéndose a los clientes de mayor valor en la red (aquellos con una mayor influencia y soporte)
- Análisis de las redes de llamadas telefónicas en compañías de telecomunicaciones para **identificación de perfiles de los usuarios y recomendación de tarifas personalizadas** de acuerdo a dichos perfiles
- Uso para **predicción de la deserción de clientes** (*churn prediction*) identificando cambios en sus patrones de contactos telefónicos
- **Detección de fraudes**, por ejemplo, en comunicaciones organizacionales (conjunto de datos de Enron) para analizar la frecuencia y la dirección de los envíos de e-mails formales/informales que pueden revelar los patrones de comunicación entre empleados y jefes
Estos patrones pueden ayudar a identificar personas implicadas en actividades fraudulentas

TIPOS DE MEDIDAS DE ANÁLISIS DE REDES SOCIALES

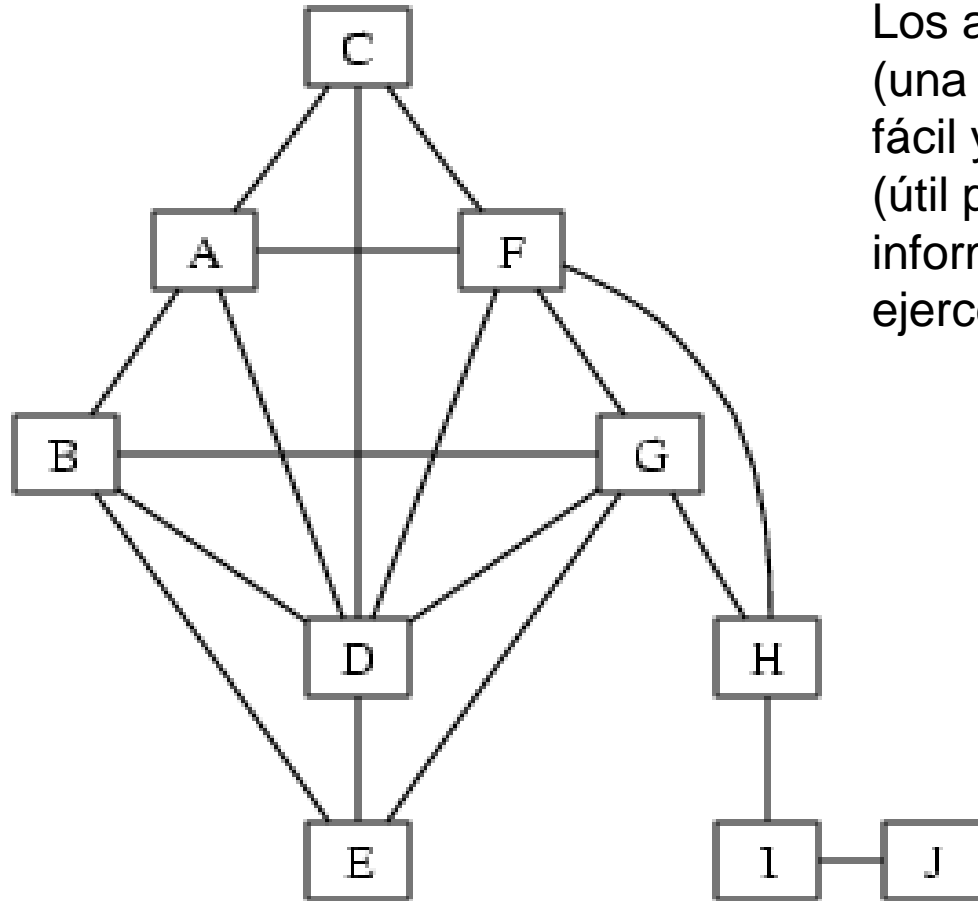
De la misma forma que en el análisis de redes en general, el análisis de la estructura de las redes sociales pretende entender el comportamiento de los sistemas complejos (en este caso, *sistemas sociales*) que generan dichas redes

Existen dos tipos de medidas:

- **Medidas locales (a nivel de actores):** Todas ellas están basadas en el concepto general de **centralidad** (redes no dirigidas) o **prestigio** (redes dirigidas), una medida general de la posición de un actor en la estructura global de la red social

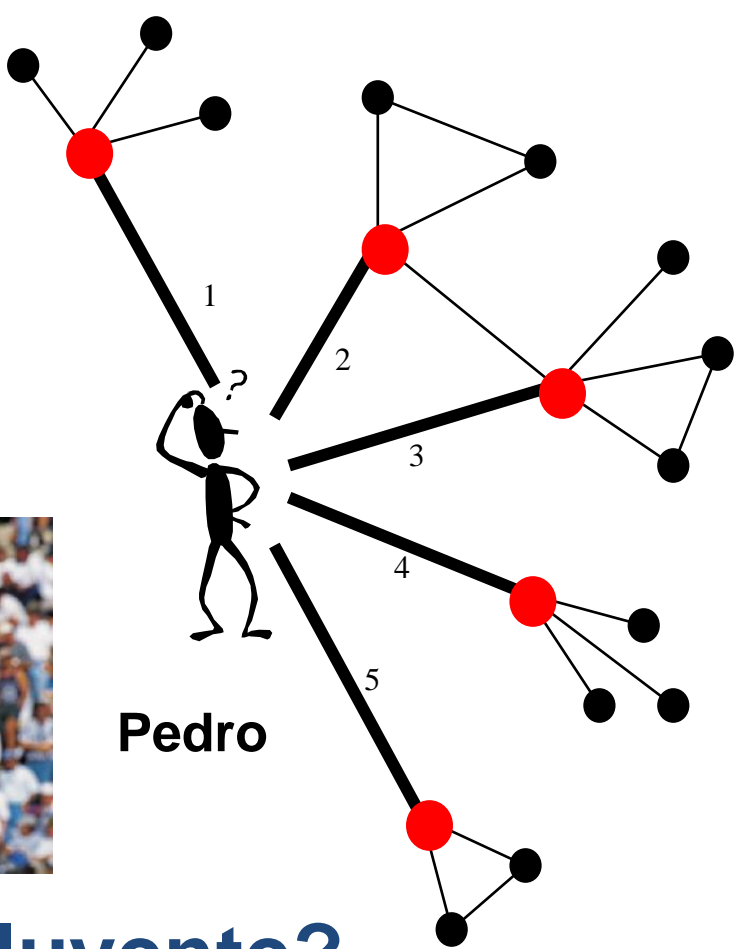
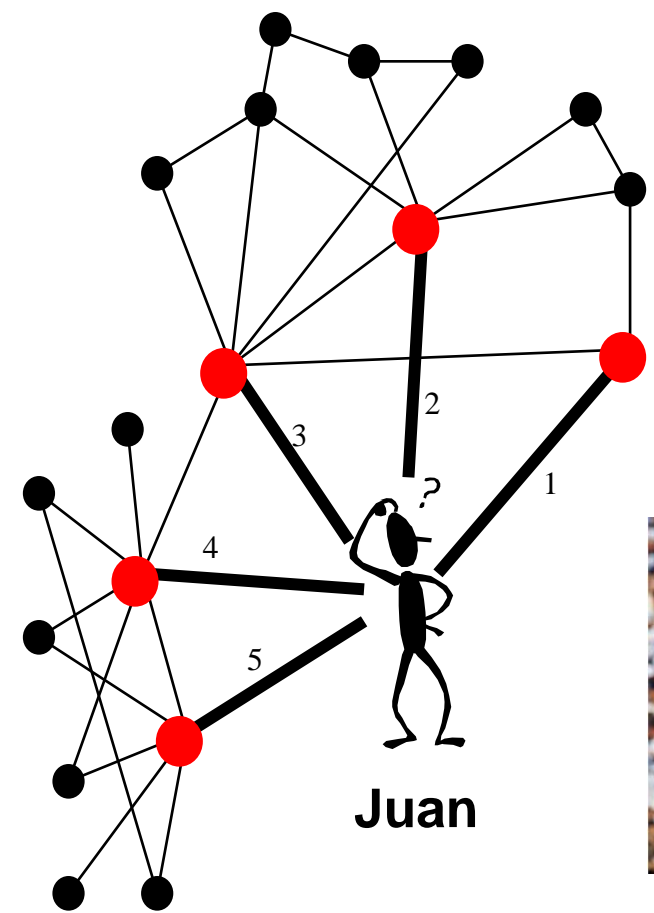
Se usan para identificar los **actores clave** de la red. Muestran como las relaciones se concentran en unos pocos individuos, dando una idea de su *poder social*

- **Medidas globales (a nivel de red):** Proporcionan información más compacta que permite evaluar la estructura global de la red, aportando información sobre propiedades importantes de los fenómenos sociales subyacentes



Los actores con una **“posición más central”** (una mayor centralidad) tienen un acceso más fácil y rápido a los demás actores de la red (útil para acceder a recursos como información) y una mayor capacidad para ejercer un control del flujo entre ellos

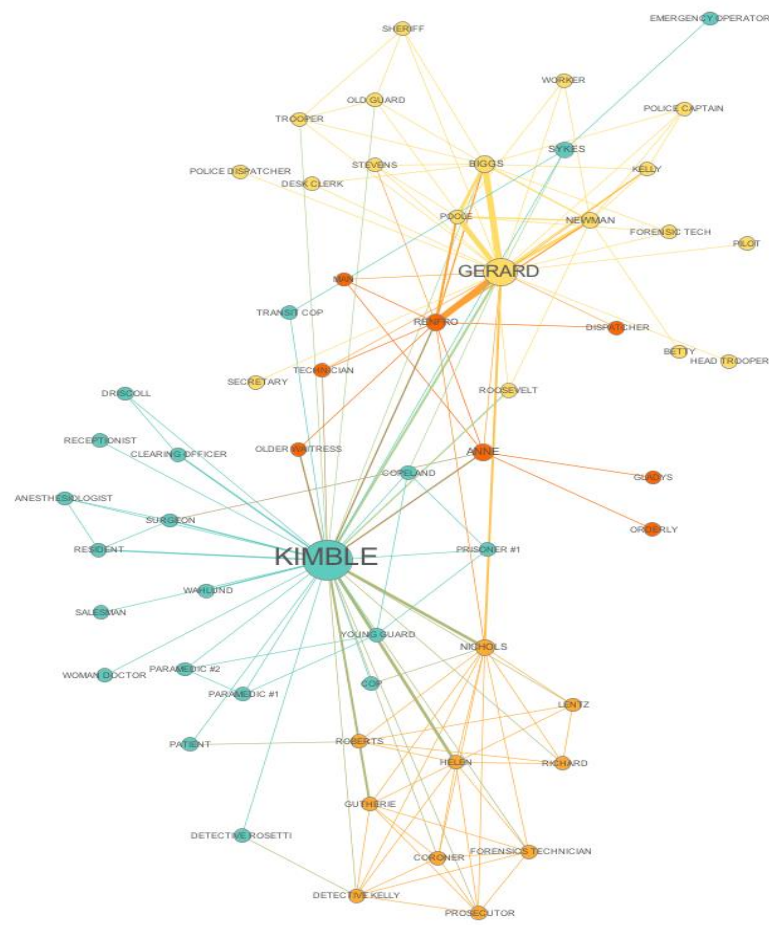
¿Quién es central en esta red?



¿Quién es más influyente?

MEDIDAS LOCALES DE CENTRALIDAD

The Fugitive (1993)



Ejemplo de Centralidad (1)

http://es.wikipedia.org/wiki/El_fugitivo_pel%C3%ADcula_de_1993



MEDIDAS LOCALES DE CENTRALIDAD

Ejemplo de Centralidad (2)

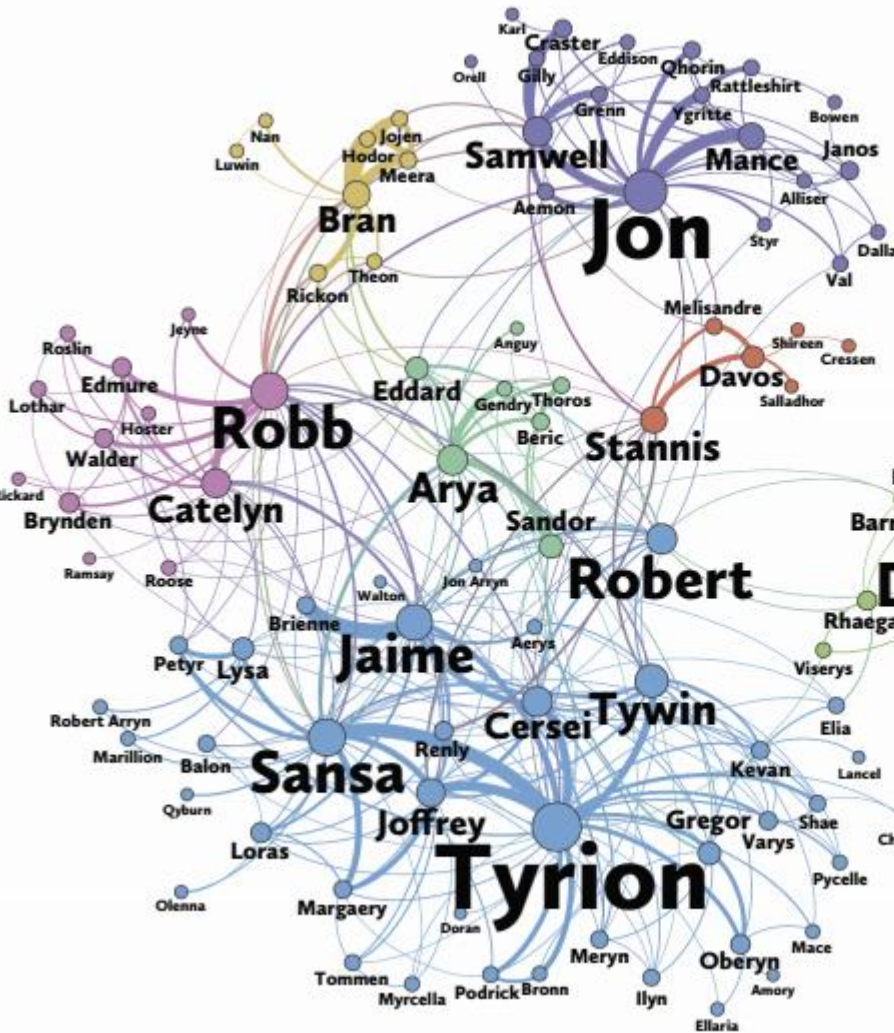


Figure 2. The social network generated from *A Storm of Swords*. The color of a vertex indicates its community. The size of a vertex corresponds to its PageRank value, and the size of its label corresponds to its betweenness centrality. An edge's thickness represents its weight.

QUARTZ

SCIENCE IS COMING

Mathematicians mapped out every "Game of Thrones" relationship to find the main character

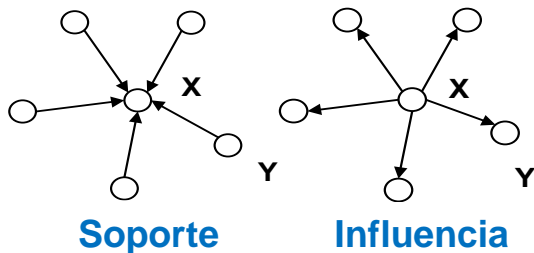


<https://qz.com/650796/mathematicians-mapped-out-every-game-of-thrones-relationship-to-find-the-main-character/>

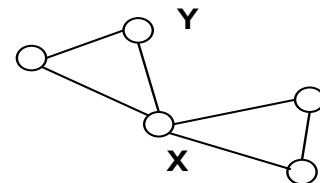
MEDIDAS LOCALES DE CENTRALIDAD

Existen varias medidas distintas de centralidad:

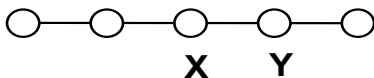
1. Grado:



2. Intermediación (betweenness):



3. Cercanía (closeness):



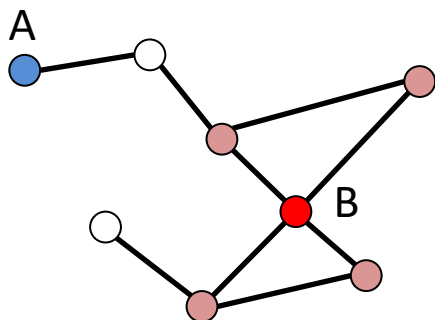
4. Excentricidad

5. Centralidad de vector propio

Es conveniente conocer bien las características de cada medida y usar varias

P.ej., las Centralidades de grado son medidas importantes pero no tienen en cuenta la estructura global de la red

No dirigida



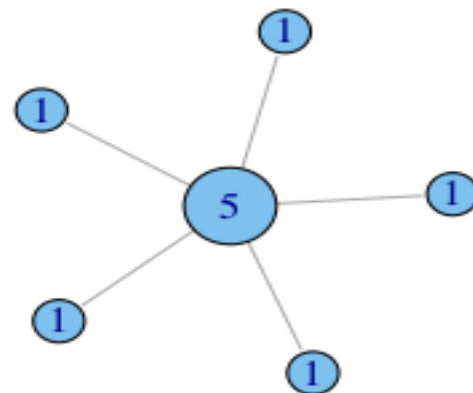
Centralidad de grado de un actor (C_D): número de enlaces que lo conectan con otros

$$C_D(A) = k_A = 1 \quad C_D(B) = k_B = 4$$

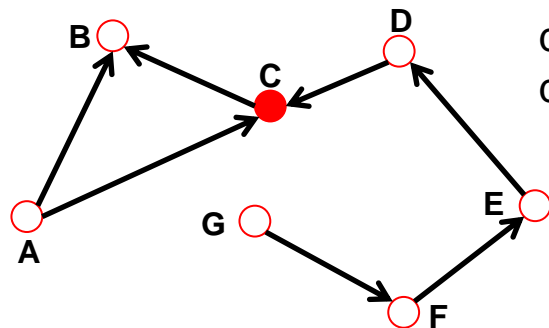
$C_D(i)$ se define en $\{0, g-1\}$, siendo g el número de nodos de la componente conexa

Interpretación: *Los actores con más amigos son más centrales*

Sólo mide la importancia con respecto a los vecinos más cercanos. Se asume que las conexiones de los amigos no importan, sólo importa lo que ellos pueden hacer directamente (p.ej. “ir a tomarse una cerveza contigo”, “ayudarte a hacer una práctica”, etc.)



Dirigida



En redes dirigidas, se define el **Prestigio de entrada** (*in-degree*), denominado **Soporte**, y el **Prestigio de salida** (*out-degree*), denominado **Influencia**:

$$P_D^{in}(C) = k_C^{in} = 2 \quad P_D^{out}(C) = k_C^{out} = 1$$

Ambos se definen en $\{0, g-1\}$

Interpretación Soporte: *Los actores que reciben muchos enlaces son prominentes*

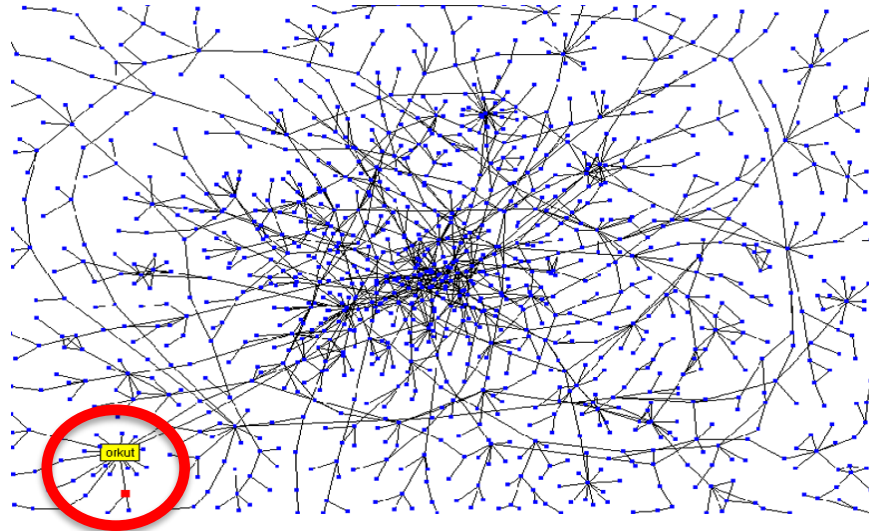
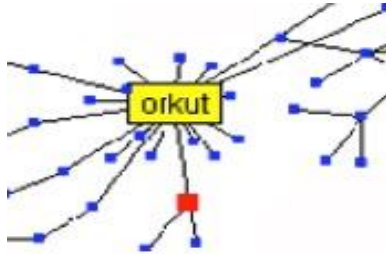
La idea básica es que muchos actores procuran tener enlaces directos a ellos, por lo que se puede considerar como una medida de importancia

Interpretación Influencia: *Los actores que tienen muchas conexiones directas con otros son influyentes*

Se entiende que pueden intercambiar o transferir información rápidamente a muchos otros (*argumento de la fortaleza de las conexiones débiles*)

MEDIDAS LOCALES DE CENTRALIDAD

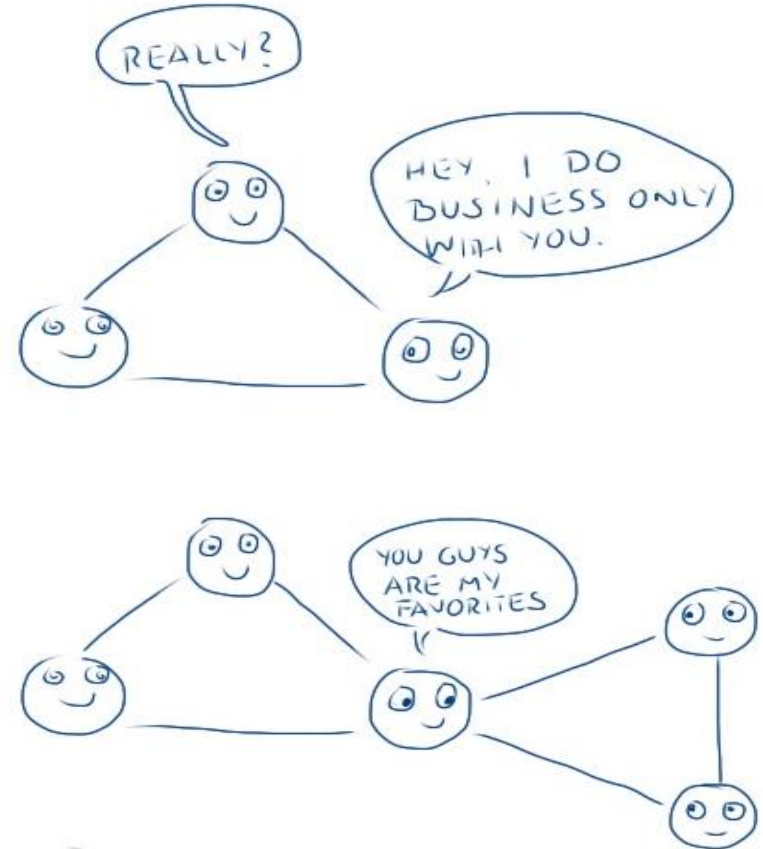
Stanford Social Web (ca. 1999)



Red de páginas web personales en Stanford

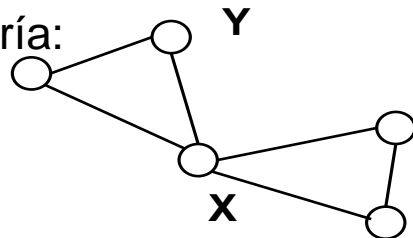
Comportamiento

El grado no captura las “corredurías” (*brokerage*)



La **intermediación** es una medida pensada para capturar la correduría:

$$C_B(i) = \sum_{j,k \in V(G)/i} g_{jk}(i)$$



donde $g_{jk}(i)$ es el número de caminos mínimos que conectan cualquier par de nodos $j \leftrightarrow k$ **y que incluyen al actor i** para el que calculamos la intermediación. $C_B(i)$ se define en $\{0, (g-1) \cdot (g-2)\}$ en redes dirigidas y en $\{0, (g-1) \cdot (g-2)/2\}$ en no dirigidas

Intuición: Ver al actor con una posición más favorable en la medida en que dicho actor esté situado entre los caminos geodésicos de todos los demás. En otras palabras, cuantos mas nodos necesiten pasar por mí para hacer sus conexiones indirectas por los caminos más cortos, más central seré yo

Es habitual considerar la medida normalizada (redes no dirigidas):

$$C'_B(i) = \frac{C_B(i)}{(g-1)(g-2)/2}$$

Número de pares de actores excluyendo el propio nodo i

Los actores con una intermediación alta ocupan roles críticos en la estructura de la red puesto que suelen ocupar una posición que les permite trabajar como **interfaces entre subgrupos de actores fuertemente unidos**

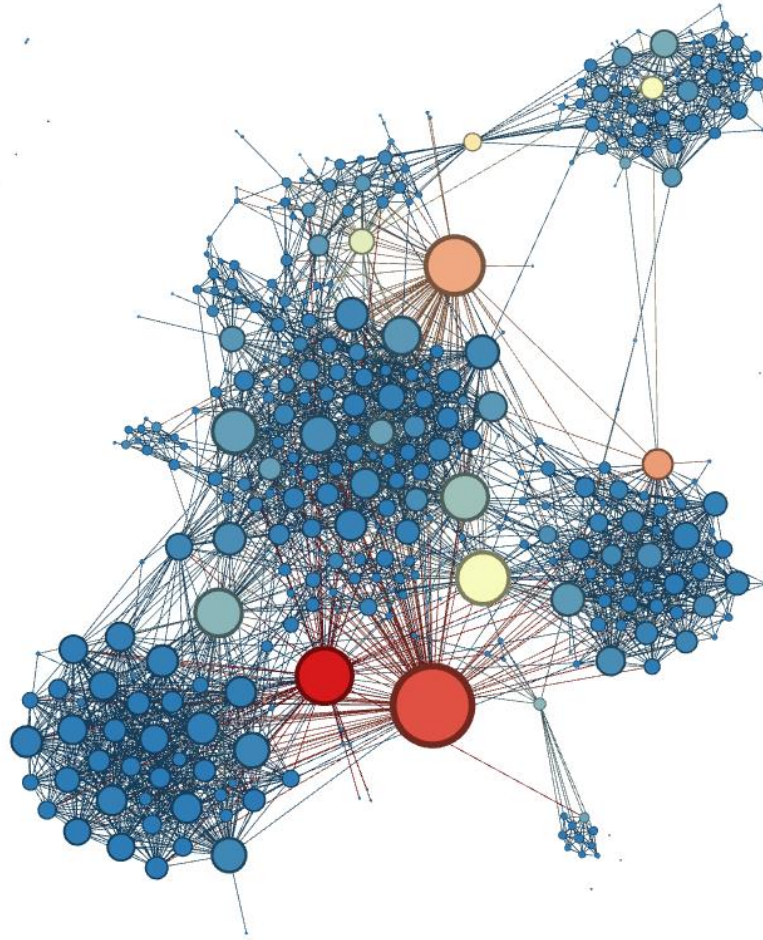
Son elementos vitales en la conexión entre distintas regiones de la red

En la perspectiva de las redes sociales, **las interacciones entre dos actores no adyacentes pueden depender de otros actores del conjunto, especialmente de aquellos situados en los caminos entre ambos**

Estos actores se denominan también **porteros** (*gatekeepers*) porque tienden a controlar el flujo de información entre comunidades

La intermediación también puede calcularse para las relaciones, midiendo el grado en que hacen posible otras conexiones. Determina **puentes locales**, enlaces con intermediaciones altas:

$$C'_B(e) = \sum_{j,k \in V(G)} g_{jk}(e)$$

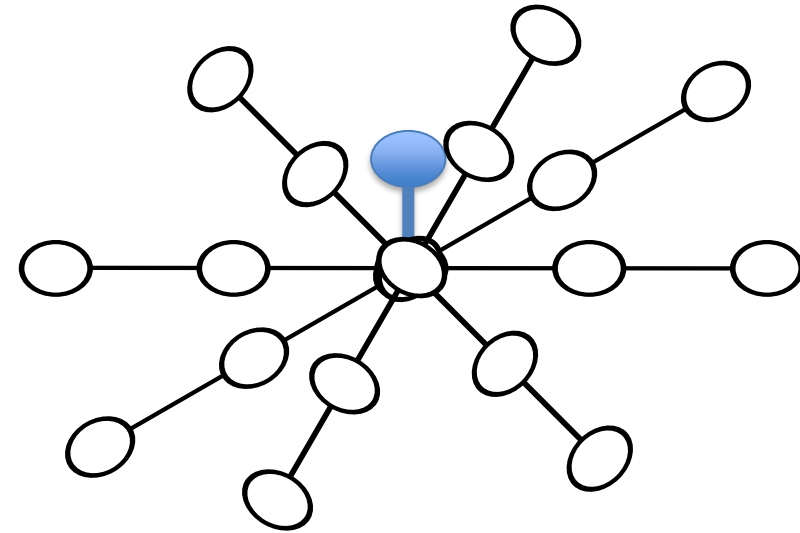


Red Personal de Contactos de Facebook de Oscar Córdón: el tamaño de los nodos indica el **grado** y el color la **intermediación** (más **azul**, menor valor; más **rojo**, mayor valor)

La **cercanía** es una tercera forma alternativa de medir la centralidad que se plantea el hecho de que puede no ser tan importante tener muchos amigos directos ni estar situado “entre” otros actores

En este caso, **se le da importancia a “estar en medio de las cosas”, no demasiado lejos del centro**, para lo cual no es necesario estar en una posición de correduría

Se enfatiza la distancia de un actor a otros en la red al concentrarse en la distancia geodésica de cada actor con todos los demás



La suma de estas distancias geodésicas (distancias de los caminos mínimos) para cada actor es la **lejanía** de dicho actor al resto

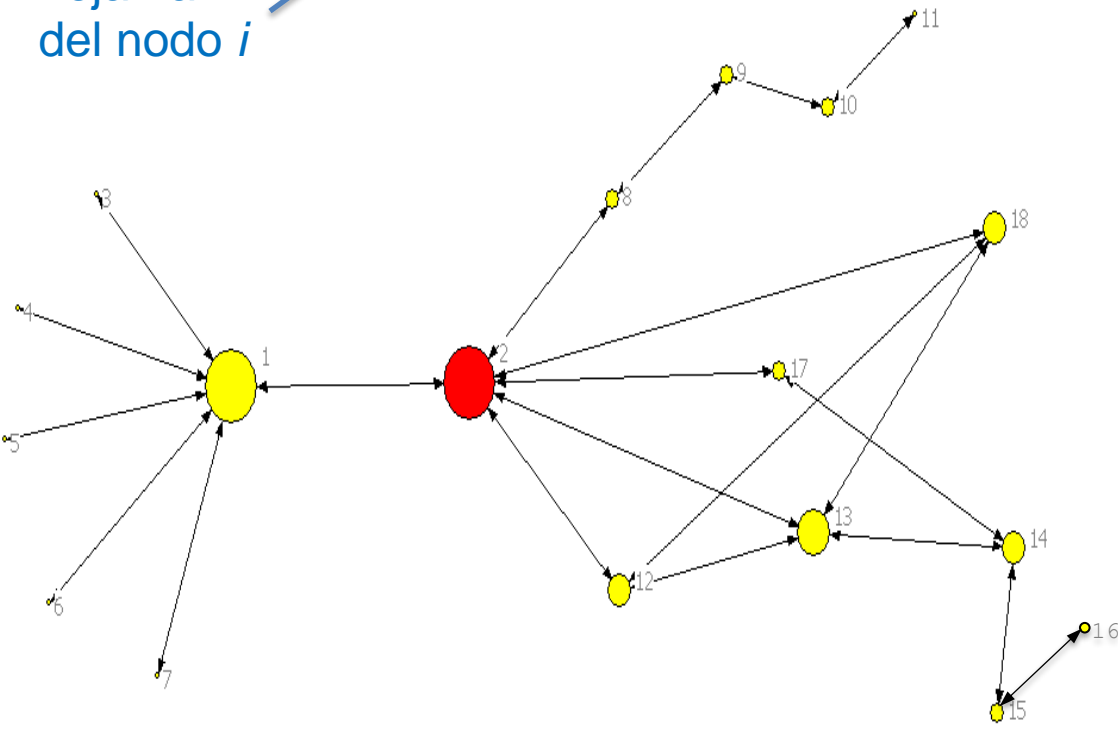
Lógicamente, cuanto más lejos se esté del centro peor. Por eso, **la inversa de esa suma (lejanía) es la medida de cercanía**

Fórmulas de la **Centralidad de cercanía** (sin normalizar, C_c , y normalizada, C'_c):

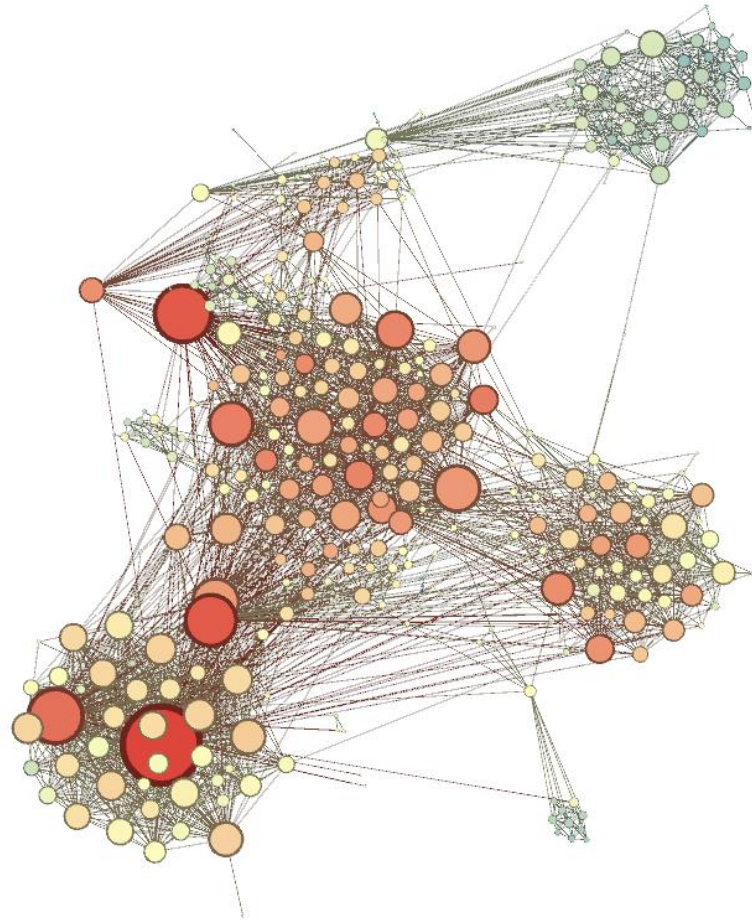
$$C_C(i) = \frac{1}{\sum_{j=1}^g d(i,j)}$$

Lejanía
del nodo i

$$C'_C(i) = \frac{C_C(i)}{\frac{1}{g-1}} = (g-1) \cdot C_C(i)$$



Actor	Lejanía	Cercanía	CercaníaN
2	34	0.02941	0.50000
1	40	0.02500	0.42500
13	42	0.02381	0.40476
17	44	0.02273	0.38636
8	44	0.02273	0.38636
12	45	0.02222	0.37778
18	45	0.02222	0.37778
14	52	0.01923	0.32692
6	56	0.01786	0.30357
5	56	0.01786	0.30357
7	56	0.01786	0.30357
3	56	0.01786	0.30357
4	56	0.01786	0.30357
9	56	0.01786	0.30357
15	66	0.01515	0.25758
10	70	0.01429	0.24286
16	82	0.01220	0.20732
11	86	0.01163	0.19767



Red Personal de Contactos de Facebook de Oscar Córdón: el tamaño de los nodos indica el **grado** y el color la **cercanía** (más **azul**, menor valor; más **rojo**, mayor valor)

La **excentricidad** de un actor de la red se define como **la distancia del camino mínimo más largo** entre él y cualquier otro actor de la red:

$$E(i) = \max_{j \in V(G)/i} d(i, j)$$

Los actores con un mayor valor de excentricidad se denominan **actores periféricos**, los de menor valor forman el **centro de la red**

Otra medida local de centralidad basada en distancias es la **Centralidad de excentricidad (C_E)**. Se define como la inversa de la excentricidad del nodo y como su valor máximo posible es 1, ya está directamente normalizada:

$$C_E(i) = \frac{1}{E(i)}$$

Los nodos con menor valor de excentricidad tienen mayor centralidad de excentricidad. De hecho, son los que componen el centro de la red

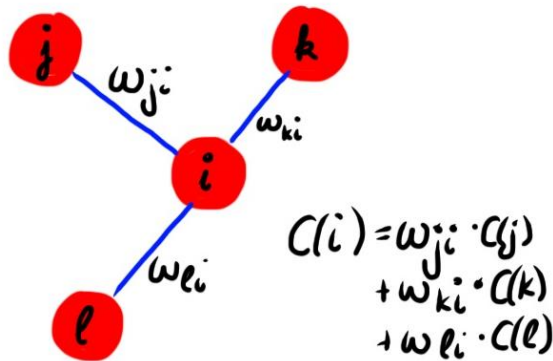
Esta medida se usa menos que las anteriores

Aún así, la excentricidad puede detectar “comportamientos extraños” en los actores de la red (algo excéntrico es algo raro, especial)

Se puede usar para detectar cuentas *fake* en redes sociales online (ej. X-Twitter) o para encontrar actores que tengan comportamientos muy diferentes al habitual en el sistema social estudiado

La **Centralidad de vector propio** se basa en que la centralidad de un nodo concreto depende de cómo de centrales sean sus vecinos (**prominencia**)

La idea básica es que el poder y el status de un actor (**ego**) se define recursivamente a partir del poder y el status de sus vecinos (**alters**)

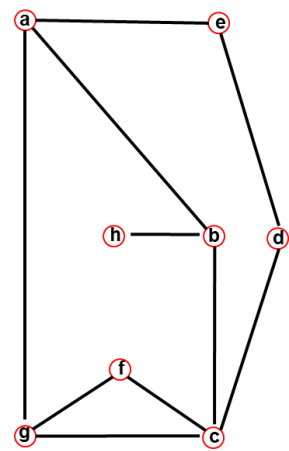


w_{ij} (a_{ij}) corresponde a la entrada de la matriz de adyacencia. Puede ser binaria $\{0,1\}$ o un peso numérico

La medida es válida para redes dirigidas (**Prestigio de rango**) y no dirigidas

Es una versión más elaborada de la Centralidad de grado al asumir que no todas las conexiones tienen la misma importancia. No se tiene en cuenta la cantidad sino la calidad de las mismas

En redes no ponderadas, la medida de Centralidad de vector propio, C_{VP} , se define como una **suma** de los valores de todos los actores que apunten a i :

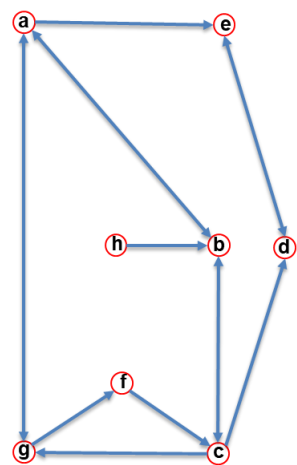


	a	b	c	d	e	f	g	h	
a	0	1	0	0	1	0	1	0	$C_{VP}(a) = 0.12$
b	1	0	1	0	0	0	0	1	$C_{VP}(b) = 0.14$
c	0	1	0	1	0	1	1	0	$C_{VP}(c) = 0.25$
d	0	0	1	0	1	0	0	0	$C_{VP}(d) = 0.21$
e	1	0	0	1	0	0	0	0	$C_{VP}(e) = 0.24$
f	0	0	1	0	0	0	1	0	$C_{VP}(f) = 0.25$
g	1	0	1	0	0	0	0	0	$C_{VP}(g) = 0.18$
h	0	1	0	0	0	0	0	0	$C_{VP}(h) = 0.05$

La centralidad de vector propio del nodo a sería:

$$C_{VP}(a) = w_{aa} \cdot C_{VP}(a) + w_{ba} \cdot C_{VP}(b) + \dots + w_{ha} \cdot C_{VP}(h) = 1 \cdot C_{VP}(b) + 1 \cdot C_{VP}(e) + 1 \cdot C_{VP}(g) = 0.14 + 0.24 + 0.18 = 0.56$$

En redes ponderadas, la medida de Centralidad de vector propio, C_{VP} , se define como una **suma con pesos** de los valores de todos los actores que apunten a i :



	a	b	c	d	e	f	g	h	
a	0	0.16	0	0	0.78	0	0.32	0	$C_{VP}(a) = 0.12$
b	0.40	0	0.44	0	0	0	0	0	$C_{VP}(b) = 0.14$
c	0	0.65	0	0.03	0	0	0.54	0	$C_{VP}(c) = 0.25$
d	0	0	0	0	1	0	0	0	$C_{VP}(d) = 0.21$
e	0	0	0	0.22	0	0	0	0	$C_{VP}(e) = 0.24$
f	0	0	0.59	0	0	0	0	0	$C_{VP}(f) = 0.25$
g	0.34	0	0	0	0	0.67	0	0	$C_{VP}(g) = 0.18$
h	0	0.98	0	0	0	0	0	0	$C_{VP}(h) = 0.05$

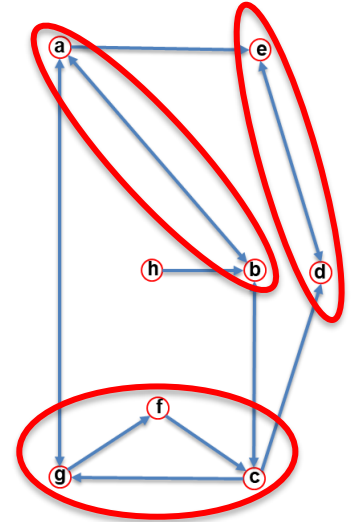
La centralidad de vector propio del nodo b sería:

$$C_{VP}(b) = w_{ab} \cdot C_{VP}(a) + w_{bb} \cdot C_{VP}(b) + \dots + w_{hb} \cdot C_{VP}(h) = 0.16 \cdot C_{VP}(a) + 0.65 \cdot C_{VP}(c) + 0.98 \cdot C_{VP}(h)$$
$$C_{VP}(b) = 0.16 \cdot 0.12 + 0.65 \cdot 0.25 + 0.98 \cdot 0.05 = 0.0192 + 0.1625 + 0.049 = 0.2307$$

La medida es muy potente pero su cálculo no es fácil

No se puede calcular en un solo paso como las anteriores al haber recursividad: para calcular la Centralidad de vector propio de un nodo A puede hacer falta la de otro B y para calcular la de B también puede ser necesario tener la de A

Incluso el propio nodo puede “aportarse” prestigio a sí mismo si existe un auto-enlace

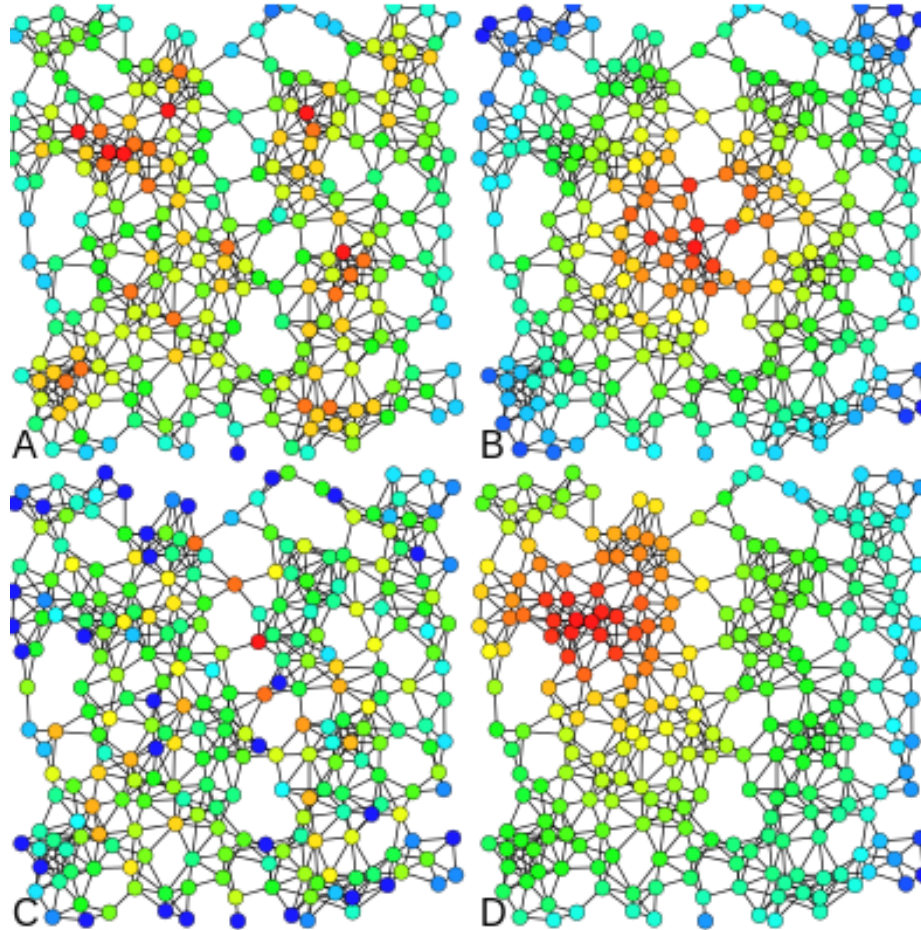


Existen distintos métodos para calcular esta medida de centralidad usando álgebra matricial y con algoritmos iterativos como el **Método de las Potencias**

https://es.wikipedia.org/wiki/Método_de_las_potencias

El algoritmo Pagerank que usa Google para su motor de búsqueda web está basado en esta medida de centralidad y en ese método

A) centralidad de grado; B) cercanía; C) intermediación; D) centralidad de vector propio



azul = menor valor

rojo = mayor valor

ALGUNAS APLICACIONES DEL ANÁLISIS DE REDES SOCIALES



Estudio realizado por Drew Conway, científico de datos, en 2011:

Collaboration with [Recorded Future](#), a Boston start-up that specializes in longitudinal entity extraction from the massive amount of open-source data generated daily. The question was:

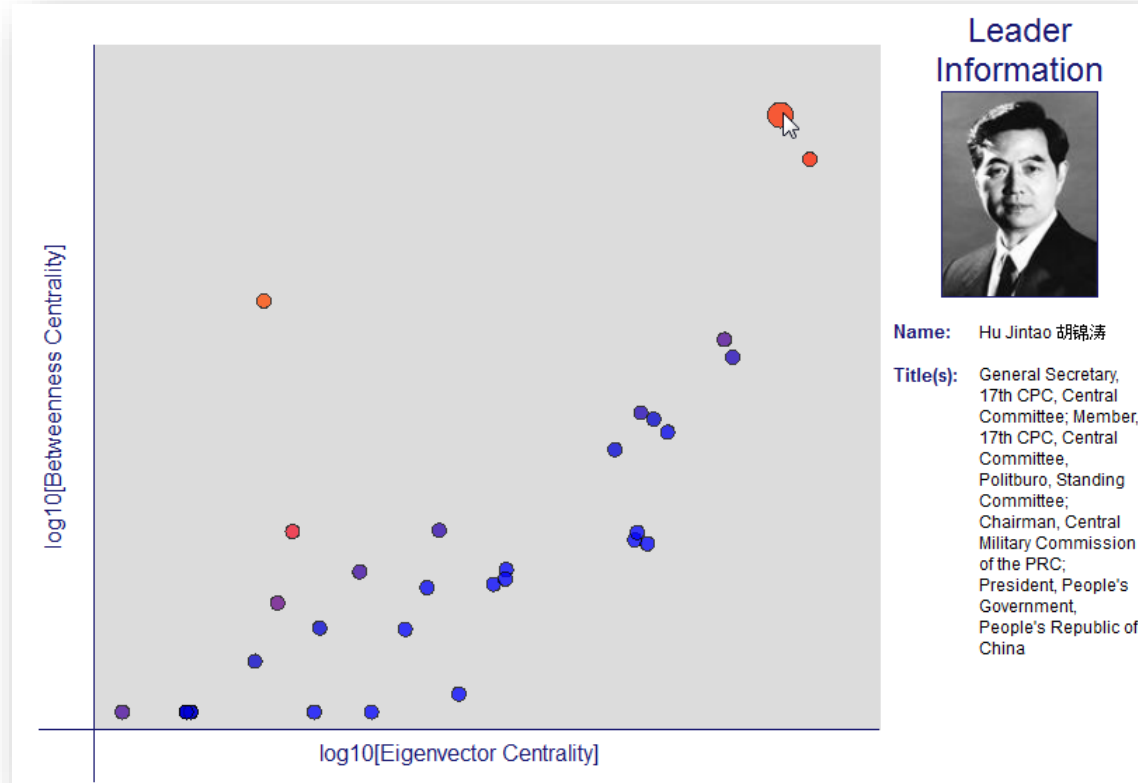
Who are the most central members of the China's leadership as we enter 2012?

For our analysis we focused on the China's leadership, as defined by the [CIA World Factbook](#), and extracted all of named entities in their data for 2011 (over 4 billion events) for which any of the 33 official Chinese leaders appear

The result is a [dataset with over 150,000 entities](#); including people, organizations, and places

To answer our questions, however, I used the **co-occurrence** of these **entities** in sentence fragments to build a large **network** of these entities

To visualize the results a **simple interactive scatter-plot of centrality measures** was used:

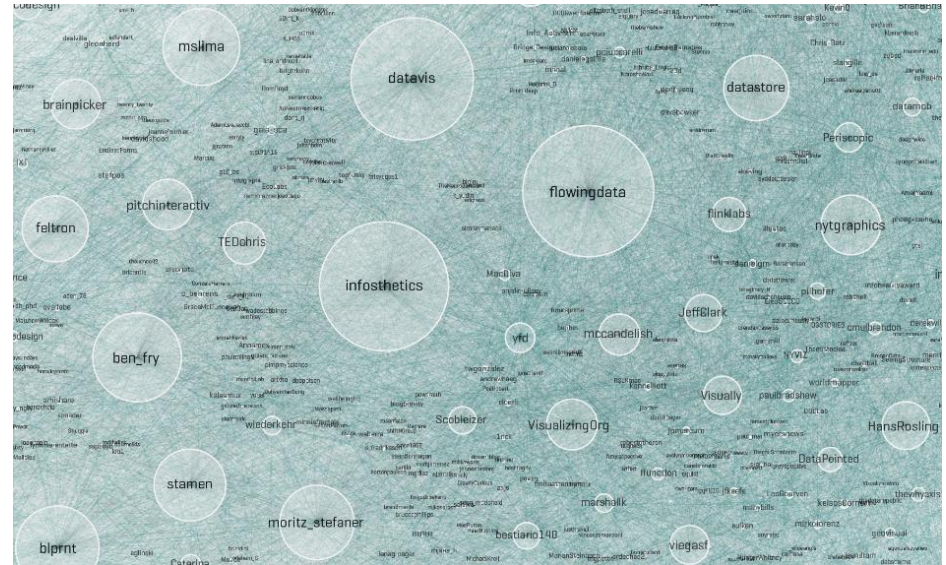


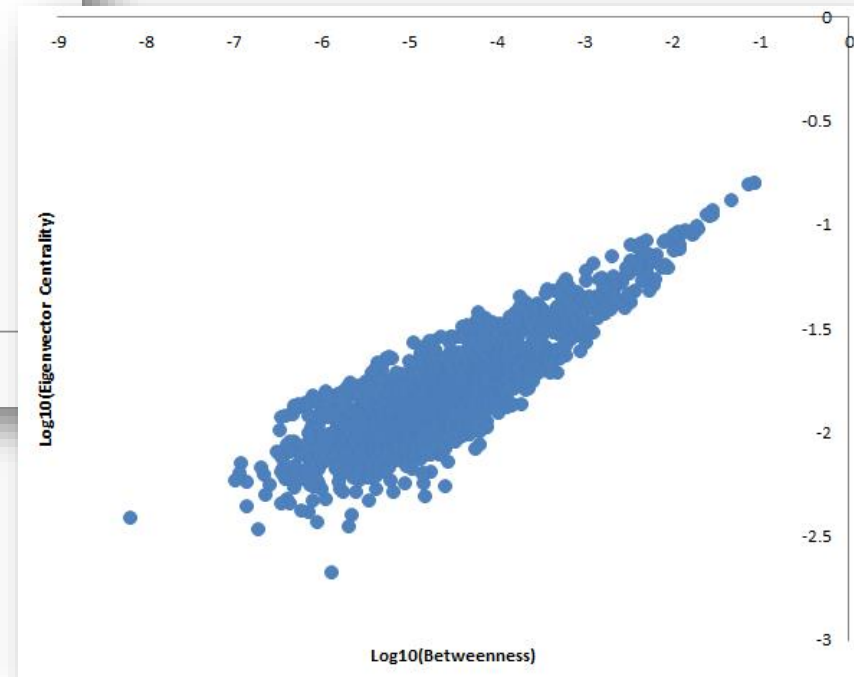
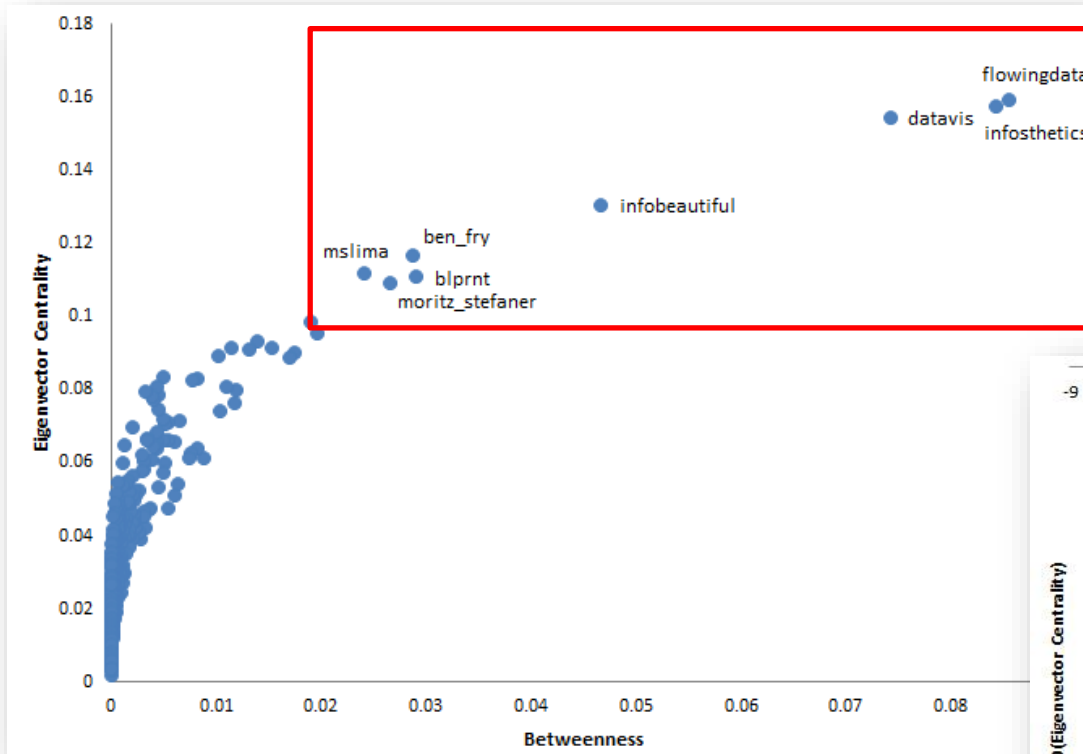
This is a great example of how an analysis can highlight the power of big data, lead to better informed follow-up questions, and convey the value of the business providing it

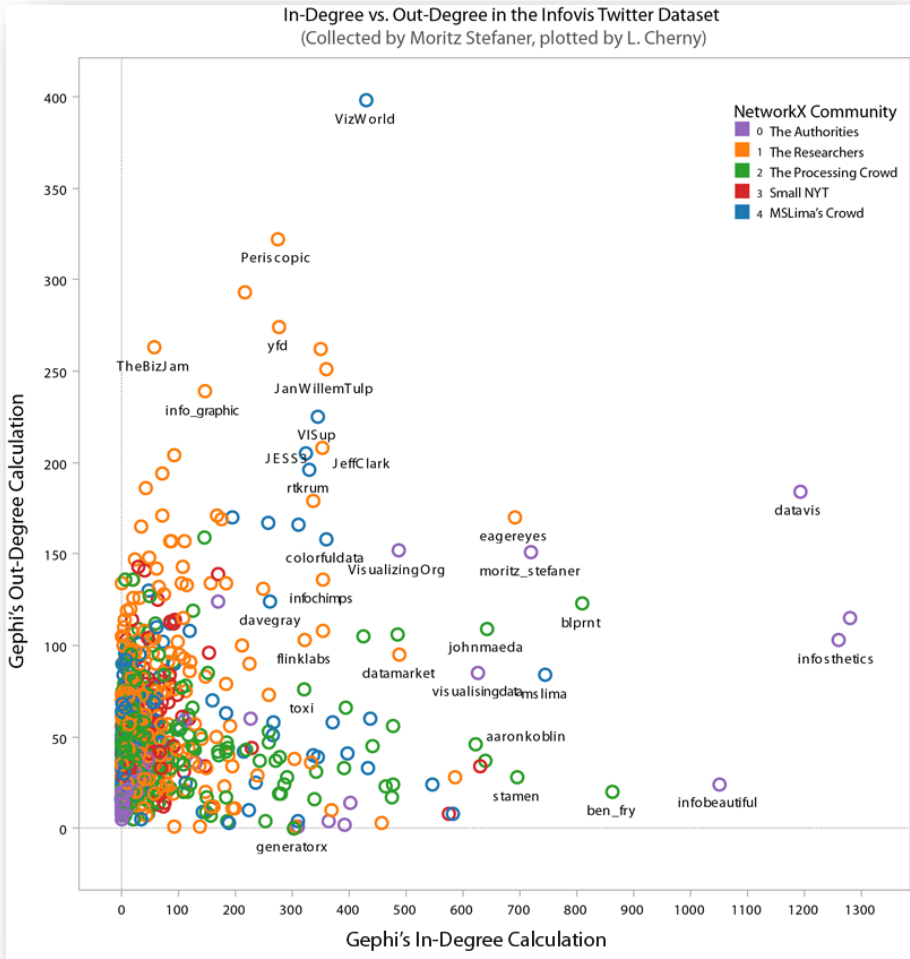
<http://blogger.ghostweather.com/2011/09/combining-through-infovis-twitter-network.html>

Se seleccionaron unas cuentas *semilla* y se extrajeron los seguidores y amigos mediante el API de Twitter. **Se filtraron aquellas cuentas con menos de 5 enlaces a las cuentas originales**

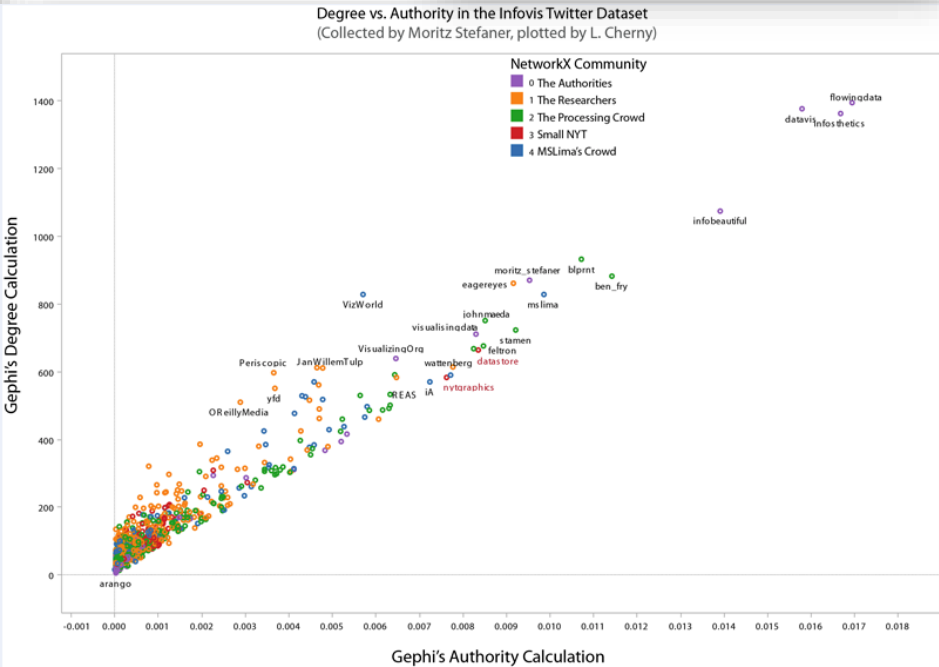
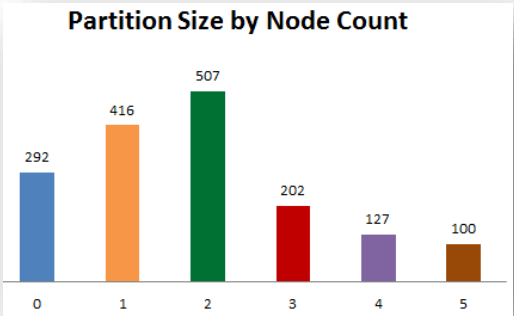
Se emplearon correlaciones entre medidas de centralidad para **determinar la influencia de las cuentas**

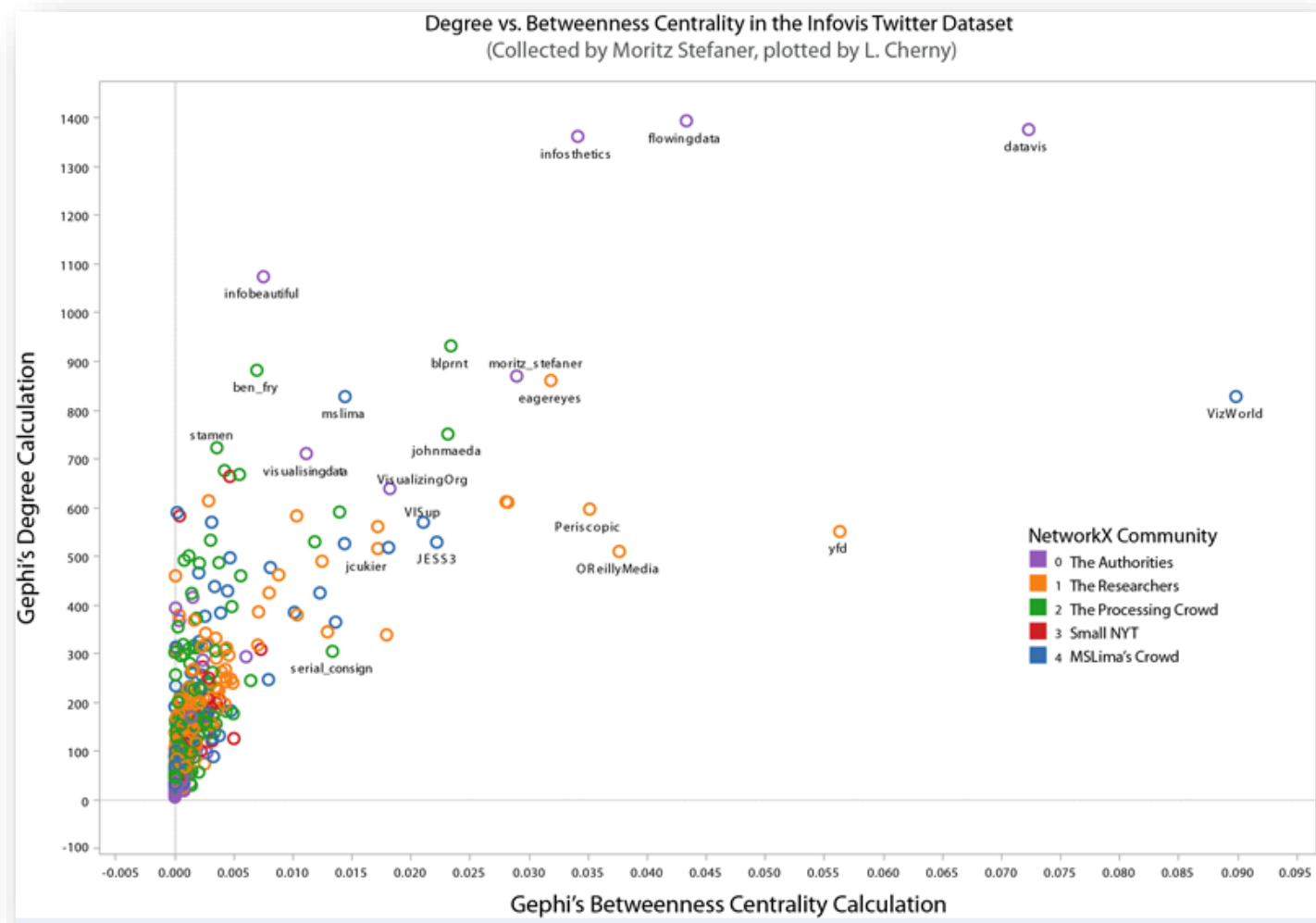






Detección de comunidades con el método de Lovaina





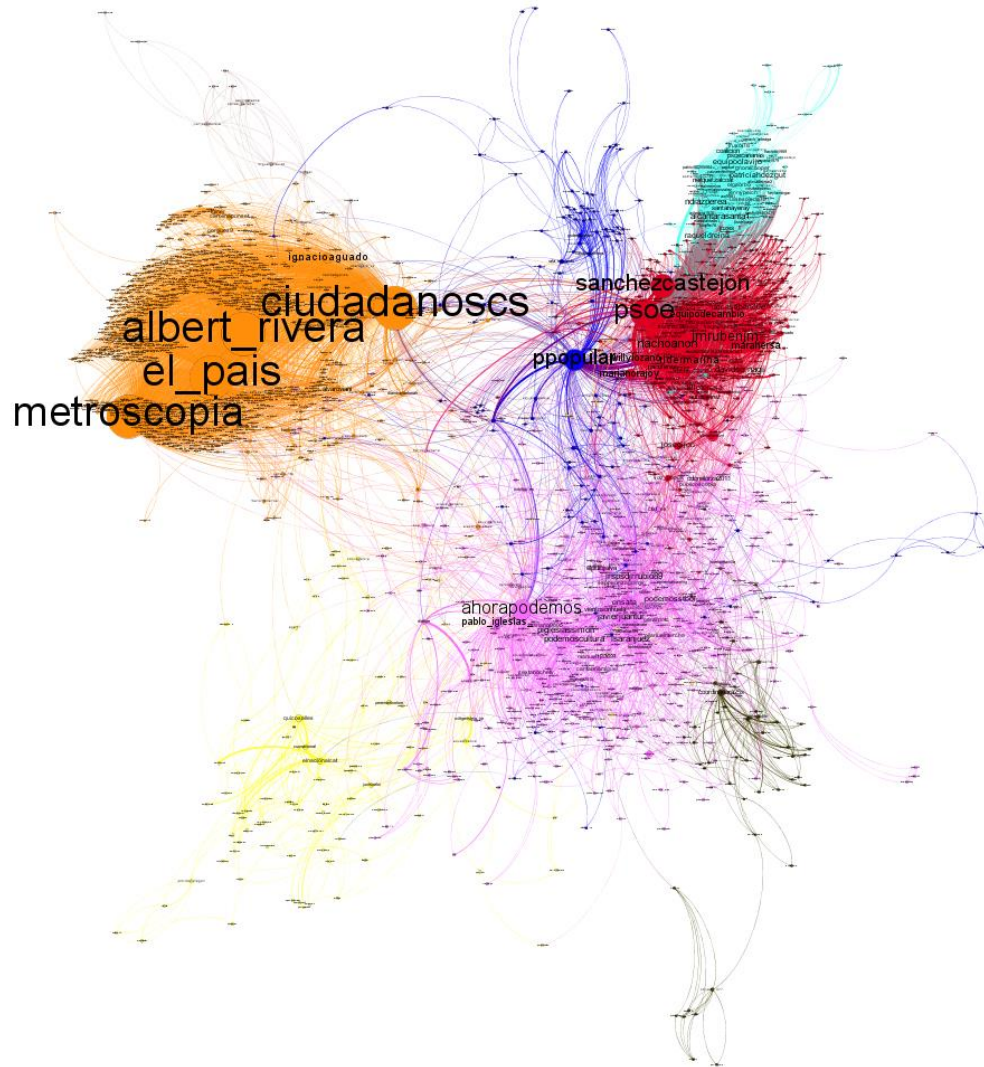
Red política descargada de Twitter, haciendo búsquedas a partir de cuentas semillas y hashtags relacionados

Los enlaces indican conversaciones, no seguidores. La red representa el estado en un momento concreto

Se filtra para obtener la parte más relevante de la red: 1807 nodos y 8639 enlaces

El tamaño de los nodos y las etiquetas indica el grado de la cuenta: a mayor tamaño, más importancia

Los colores representan una detección automática de comunidades



Usando medidas de centralidad, podemos evaluar la influencia de los distintos actores (cuentas) en el sistema social **en el momento concreto** (estos datos están adquiridos en 11N de 2015)

Es importante plantearse las preguntas que se desean responder y escoger la medida adecuada para resolverlas

Por ejemplo, dos preguntas podrían ser:

1. ¿Qué cuentas de Twitter lideran la discusión política antes de las elecciones?
2. ¿Qué cuentas hacen de puente en esa discusión, conectando distintas conversaciones, opciones políticas, etc?

1. ¿Qué cuentas de Twitter lideran la discusión política antes de las elecciones?

Para responder podemos usar las centralidades de grado y, sobre todo, de vector propio

Los 10 actores más relevantes por centralidad de grado y de vector propio:

Id	Label	Grado	Eigenvector Centrality
55	ciudadanoscs	737	0.972462
167	el_pais	680	1.0
8	albert_rivera	665	0.961907
238	metroscopia	632	0.934574
77	psoe	271	0.562495
16	sanchezcastejon	212	0.430676
12	ppopular	183	0.368587
1	ahorapodemos	161	0.247495
138	jmrubenjm	105	0.10175
104	nachoanon	94	0.129637

Id	Label	Grado	Eigenvector Centrality
167	el_pais	680	1.0
55	ciudadanoscs	737	0.972462
8	albert_rivera	665	0.961907
238	metroscopia	632	0.934574
77	psoe	271	0.562495
16	sanchezcastejon	212	0.430676
12	ppopular	183	0.368587
1	ahorapodemos	161	0.247495
625	ignacioaguado	73	0.189005
22	marianorajoy	68	0.156895

Podemos observar que no son exactamente los mismos ni tienen el mismo orden en cada medida. Salvo que el sistema social esté sesgado y dominado por unos pocos actores muy relevantes y con un ranking muy concreto, es lo habitual

1. ¿Qué cuentas de Twitter lideran la discusión política antes de las elecciones?

Id	Label	Grado	Grado de entrada	Grado de salida
55	ciudadanoscs	737	733	4
167	el_pais	680	680	0
8	albert_rivera	665	661	4
238	metroscopia	632	630	2
77	psoe	271	269	2
16	sanchezcastejon	212	212	0
12	ppopular	183	181	2
1	ahorapodemos	161	153	8
138	jmrubenjm	105	36	69
104	nachoanon	94	55	39
171	ridermarina	74	44	30
625	ignacioaguado	73	71	2
22	marianorajoy	68	68	0
443	equipodecambio	66	56	10
31	willylozano	66	58	8
105	pablo_iglesias_	65	64	1

Grado: Al ser una red dirigida, tenemos dos medidas de grado, la de entrada y la de salida, que indican el soporte y la influencia

Como es habitual en Twitter, las cuentas importantes tienen mucho grado de entrada pero poco de salida

En cuanto a partidos, los más relevantes de acuerdo a esta medida son las cuentas de Ciudadanos y de su entonces líder, Albert Ribera

Los siguen el PSOE y Pedro Sánchez y el PP, pero la cuenta de Mariano Rajoy es menos relevante

Podemos es el siguiente más relevante pero Pablo Iglesias está mucho más abajo

La segunda y la cuarta cuentas más relevantes para esta medida no corresponden a partidos, son un periódico (El País) y un instituto de sondeos políticos (Metroscopia)

Encontramos otro político de Ciudadanos, Ignacio Aguado, y varios militantes del PSOE junto con otra cuenta institucional del PSOE, @equipodecambio

1. ¿Qué cuentas de Twitter lideran la discusión política antes de las elecciones?

Id	Label	Eigenvector Centrality
167	el_pais	1.0
55	ciudadanoscs	0.972462
8	albert_rivera	0.961907
238	metroscopia	0.934574
77	psoe	0.562495
16	sanchezcastejon	0.430676
12	ppopular	0.368587
1	ahorapodemos	0.247495
625	ignacioaguado	0.189005
22	marianorajoy	0.156895
443	equipodecambio	0.15143
163	davidsernagil	0.150215
31	willylozano	0.135592
105	pablo_iglesias_	0.132353
104	nachoanon	0.129637
527	josegijon	0.128576
306	adrialastra	0.126312
171	ridermarina	0.118818
330	podemoscultura	0.117802

Vector Propio: Al contrario que en el Grado, la cuenta más relevante es la de El País

El político más relevante es Albert Rivera, en tercera posición, con mucha diferencia con el resto

En los partidos, Ciudadanos es el más relevante. De nuevo, esto queda demostrado tanto en el orden (segunda posición) como también en la diferencia de los valores de la medida

Le siguen PSOE (quinto), PP (séptimo) y Podemos (octavo), con diferencias significativas en la medida

Aparecen nuevas cuentas institucionales (@podemoscultura), de políticos (Adriana Lastra) y de militantes (@davidsernagil, @ridermarina, ...). Son los líderes de opinión del momento

Como hemos estudiado, **esta medida es más fiable para responder a nuestra pregunta porque refleja mejor el flujo global de información en la red** que el grado, mucho más local

2. ¿Qué cuentas hacen de puente conectando distintas conversaciones?

Id	Label	Betweenness Centrality	Grado	Eigenvector Centrality
138	jmrubenjm	0.011772	105	0.10175
72	isaranjuez	0.007964	54	0.03069
2	nalalop	0.007268	20	0.025144
197	esahora2015	0.007156	17	0.018691
32	ensata	0.007135	49	0.069391
405	blogsostenible	0.006066	10	0.009103
381	solabria	0.006006	9	0.006594
565	ceciliocastro	0.005909	9	0.016448
628	jlazaro	0.005542	17	0.006297
419	antgalan	0.005229	22	0.024141
114	bobestropajo	0.004801	25	0.034622
104	nachoanon	0.003157	94	0.129637
51	pilarluelmercha	0.002507	33	0.032122
499	marahersa	0.002376	65	0.085117
163	davidsermagil	0.002365	60	0.150215
137	pjsyddhh	0.002068	8	0.005436
17	podemosstboi	0.002068	47	0.050983
171	ridermarina	0.001888	74	0.118818
68	ctxt_es	0.001799	29	0.04428
762	edulorengarcia	0.001447	13	0.009856
31	willylozano	0.001445	66	0.135592
595	martidimauro	0.001374	7	0.004831
90	elnacionalcat	0.001299	28	0.022141
136	lareirinha	0.001278	14	0.002646
657	gatomaraunico	0.001243	7	0.004417
55	ciudadanoscs	0.001198	737	0.972462

Usaremos la **centralidad de intermediación** que determina qué actores de la red tienen una función de correduría. En este caso, son cuentas que hacen de puente entre distintas discusiones

Analizamos los 26 actores más relevantes para mostrar que la primera gran cuenta institucional de los partidos aparece en esa posición tan baja

Tiene sentido, las cuentas institucionales actúan más como cajas de resonancia e interactúan menos

Un poco más arriba encontramos la cuenta de Podemos en Sant Boi (@podemosstboi) y una cuenta de información de Cataluña (@elnacionacat). Es también coherente, siempre hay discusiones con Cataluña en periodo electoral

Los más relevantes son activistas políticos y cuentas que pedían el cambio de gobierno. Como vemos, **la mayoría tienen valores bajos en las otras medidas**

Referencias y Agradecimientos

Para elaborar las transparencias de este curso, he hecho uso de algunos materiales desarrollados por expertos en el área disponible en Internet:

- “Network Science Interactive Book Project” del Laszlo Barabasi Lab.
Northeastern University: <http://barabasilab.com/networksciencebook>
- K. Lerman. “Social Media. A Responsible User’s Guide”. University of South California: <http://www.isi.edu/integration/people/lerman/talks.html>
- Curso on-line “Social Network Analysis” de Lada Adamic, Coursera, Universidad de Michigan: <https://www.coursera.org/course/sna>
- L. Cherny. “Simplifying Social Network Visualizations”. Ghostweather Research & Design, LLC: <http://es.slideshare.net/arnicas/simplifying-social-network-diagrams?related=1>



Northeastern University
Center for Complex Network Research

