

# Análisis de Redes

George G Vega<sup>1</sup>

Superintendencia de Pensiones

27 de junio, 2014

---

<sup>1</sup><mailto:gvegayon@caltech.edu>

# Contenidos

- 1 Resiliencia
- 2 Homofilia (assortative mixing)
- 3 Centralidad
- 4 Importar datos a Gephi
- 5 Referencias

# Resiliencia de un grafo

*Un grafo  $G$  (de una cierta clase) posee la propiedad  $P$  [...] Qué tan fuertemente poseida se encuentra dicha propiedad? [...] Definimos **resiliencia** del grafo  $G$  con respecto a  $P$  como la medida que indica que tanto debe cambiar  $G$  para destruir  $P$ .*

*[...] (**Resiliencia**) Sea  $P$  una propiedad monotónicamente creciente. La resiliencia global de  $G$  con respecto a  $P$  es el número mínimo  $r$  tal que al eliminar  $r$  arcos/vertices de  $G$  resulta en la pérdida de la propiedad  $P$*

Fuente: Adaptado de [6]

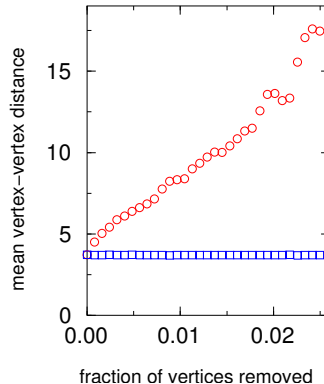


FIG. 7 Mean vertex-vertex distance on a graph representation of the Internet at the autonomous system level, as vertices are removed one by one. If vertices are removed in random order (squares), distance increases only very slightly, but if they are removed in order of their degrees, starting with the highest degree vertices (circles), then distance increases sharply. After Albert *et al.* [15].

		women			
		black	hispanic	white	other
men	black	506	32	69	26
	hispanic	23	308	114	38
	white	26	46	599	68
	other	10	14	47	32

TABLE III Couples in the study of Catania *et al.* [85] tabulated by race of either partner. After Morris [302].

A classic example of assortative mixing in social networks is mixing by race. Table III for example reproduces results from a study of 1958 couples in the city of San Francisco, California. Among other things, the study

Fuente: [5]

# Homofilia (assortative mixing)

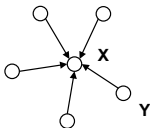
	network	$n$	$r$
real-world networks	physics coauthorship <sup>a</sup>	52 909	0.363
	biology coauthorship <sup>a</sup>	1 520 251	0.127
	mathematics coauthorship <sup>b</sup>	253 339	0.120
	film actor collaborations <sup>c</sup>	449 913	0.208
	company directors <sup>d</sup>	7 673	0.276
	Internet <sup>e</sup>	10 697	-0.189
	World-Wide Web <sup>f</sup>	269 504	-0.065
	protein interactions <sup>g</sup>	2 115	-0.156
	neural network <sup>h</sup>	307	-0.163
models	food web <sup>i</sup>	92	-0.276
	random graph <sup>u</sup>		0
	Callaway <i>et al.</i> <sup>v</sup>		$\delta/(1 + 2\delta)$
	Barabási and Albert <sup>w</sup>		0

TABLE I: Size  $n$  and assortativity coefficient  $r$  for a number of different networks: collaboration networks of (a) scientists in physics and biology [16], (b) mathematicians [17], (c) film actors [4], and (d) businesspeople [18]; (e) connections between autonomous systems on the Internet [19]; (f) undirected hyperlinks between Web pages in a single domain [6]; (g) protein-protein interaction network in yeast [20]; (h) undirected (and unweighted) synaptic connections in the neural network of the nematode *C. Elegans* [4]; (i) undirected trophic relations in the food web of Little Rock Lake, Wisconsin [21]. The last three lines give analytic results for model networks in the limit of large network size: (u) the random

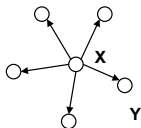
# Centralidad

## Comparación de tipos de centralidad

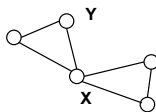
In each of the following networks, X has higher centrality than Y according to a particular measure



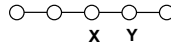
indegree



outdegree



betweenness



closeness

Fuente: Extraído de *The structure and function of complex networks*, Newman (2003)

[5]

# Centralidad

## Intermediación (betweenness)

*[betweenness centrality] Se define como la porción de veces que el nodo  $i$  necesita al nodo  $j$  (sobre el cual se está midiendo la centralidad) para alcanzar al nodo  $k$ . Específicamente, si  $g_{ij}$  es el número de rutas de  $i$  a  $j$ , y  $g_{ikj}$  es el número de geodésicas que pasa por el nodo  $k$ , entonces la centralidad de intermediación está dada por:*

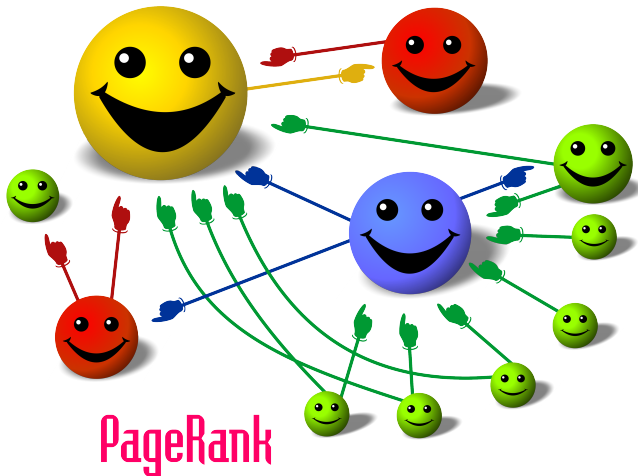
$$C_B = \sum_i \sum_j \frac{g_{ikj}}{g_{ij}}, i \neq j \neq k \quad (1)$$

*En términos sencillos, [...] básicamente cuenta el número de geodésicas que pasan a través del nodo  $k$ . [2]*



# Centralidad

## PageRank



Caricatura de PageRank

# Importad datos a Gephi

Existen varias formas de importar datos a Gephi:

- **GEXF** Formato XML, soporta atributos dinámicos y estáticos + jerarquía <http://gexf.net>.
- **GDF** Texto plano  
[http://guess.wikispot.org/The\\_GUESS\\_.gdf\\_format](http://guess.wikispot.org/The_GUESS_.gdf_format).
- **Pajek NET** Formato de pajek, texto plano <http://netwiki.amath.unc.edu/DataFormats/PajekNetAndPajFormats>
- **UCINET**  
<http://www.analytictech.com/networks/dataentry.htm>
- **Archivos CSV** los revisaremos en breve...

# Importar datos a Gephi

## Pasos a seguir

Una manera de importar datos a Gephi es a través de archivos delimitados por comas (CSV). Para hacerlo son necesarios dos archivos, uno con vértices y otro con arcos.

Pasos según [3]

- ❶ Crear un archivo de **vértices** cuyo contenido es:
  - ❶ Identificador.
  - ❷ Número de identificación
  - ❸ Etiqueta
- ❷ Crear un archivo de **arcos** cuyo contenido es:
  - ❶ Origen
  - ❷ Destino
  - ❸ Tipo (directed/undirected)
  - ❹ Número de identificación
  - ❺ Etiqueta
- ❸ Guardar ambos archivos en formato .csv
- ❹ Abrir Gephi y dirigirse al *laboratorio de datos*, luego clicar en *Importar archivo de nodos* y *Importar archivo de arcos*.

# Referencias I



Lada Adamic.

Social network analysis, 2014.

Accessed 18 jun, 2014.



Stephen P Borgatti.

Centrality and network flow.

*Social networks*, 27(1):55–71, 2005.



K. Cherven.

*Network Graph Analysis and Visualization with Gephi.*

Community experience distilled. Packt Publishing, 2013.



Mark EJ Newman.

Assortative mixing in networks.

*Physical review letters*, 89(20):208701, 2002.

## Referencias II



Mark EJ Newman.

The structure and function of complex networks.

*SIAM review*, 45(2):167–256, 2003.



Benny Sudakov and Van H Vu.

Local resilience of graphs.

*Random Structures & Algorithms*, 33(4):409–433, 2008.