

Aprendizaje no supervisado

VC10: Análisis de grafos, PageRank y otros

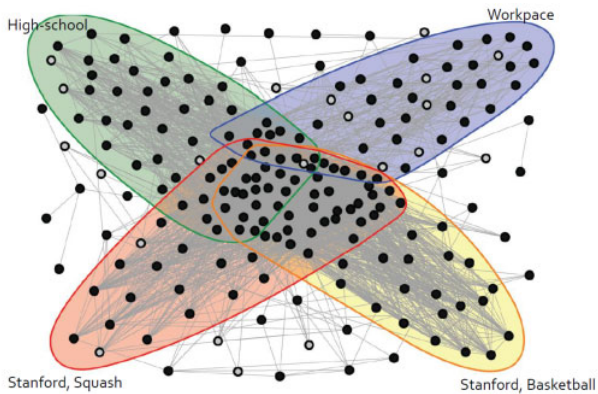
Félix José Fuentes Hurtado

felixjose.fuentes@campusviu.es

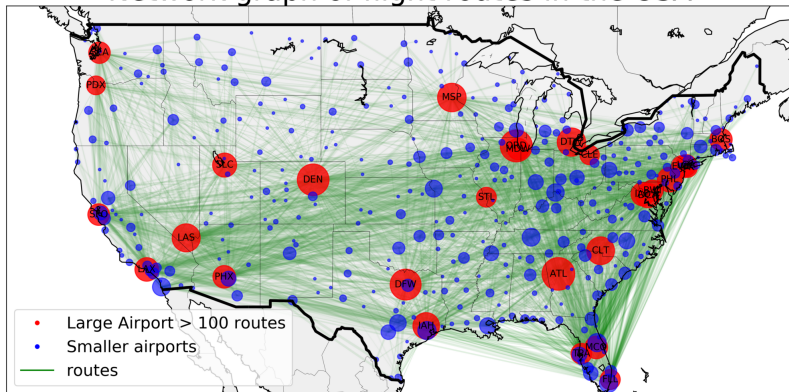
Universidad Internacional de Valencia

Agrupamiento

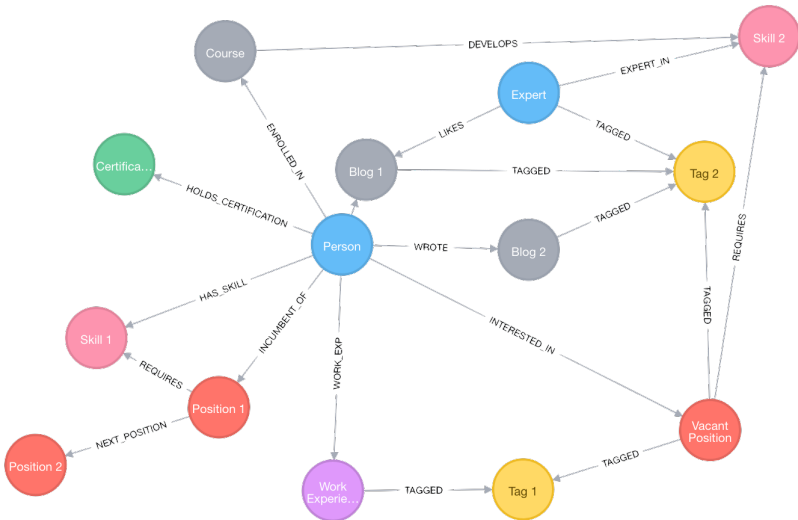
La Web como grafo



Network graph of flight routes in the USA



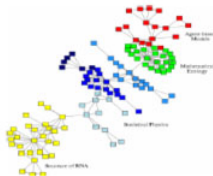
La Web como grafo



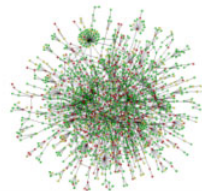
La Web como grafo



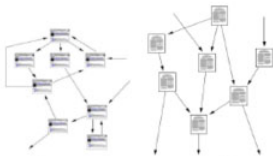
Online social networks



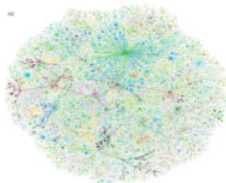
Collaboration networks



Systems biology networks



Web graphs & citation networks



Computer & Internet networks



Communication networks



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item

Donald Trump

Donald John Trump (born June 14, 1946) is an American businessman, television producer, and politician who is the **Republican Party** nominee for **President of the United States** in the 2016 election. He is the chairman and president of **The Trump Organization**, which is the principal holding company for his real estate ventures and other business interests. During his career, Trump has built office towers, hotels, casinos, golf courses, an urban development project in Manhattan, and other branded facilities worldwide.

Trump was born and raised in New York City and received a bachelor's degree in economics from the Wharton School of the University of Pennsylvania in 1968. In 1971 he was given control of his father Fred Trump's real estate and construction firm and later renamed it The Trump Organization, rising to public prominence shortly thereafter. Trump has appeared at the Miss USA pageants, which he owned from 1996 to 2015, and has made cameo appearances in films and television series. He sought the Reform Party presidential nomination in 2000, but withdrew before voting began. He hosted and co-produced *The Apprentice*, a reality television series on NBC, from 2004 to 2015. As of 2016, he was listed by *Forbes* as the 324th wealthiest person in the world, and 156th in the United States.

In June 2015, Trump announced his candidacy for president as a Republican and quickly emerged as the front-runner for his party's nomination. In May 2016, his remaining Republican rivals suspended their campaigns, and in July he was formally nominated for president at the 2016 Republican National Convention. Trump's campaign has received unprecedented media coverage and international attention. Many of his statements in interviews, on Twitter, and at campaign rallies have been controversial or false. Several rallies during the primaries were accompanied by protests or riots. On October 7, a 2005 audio recording surfaced in which Trump bragged about forcibly kissing and groping women; at least fifteen women accused him of similar conduct shortly thereafter.^{[R][R]} He apologized for the 2005 comments and denied the allegations.

Trump's platform includes renegotiation of U.S.–China trade deals, opposition to particular trade agreements such as NAFTA and the Trans-Pacific Partnership, stronger enforcement of immigration laws together with building a wall along the United States–Mexico border, and US isolationism, such, support and reinforcement of the Paris Agreement.

339. ^a Linick, Jack (July 7, 2015). "More People Are Running for Presidential Nomination Than Ever". *Time*. Retrieved February 14, 2016.
340. ^a Howell, Kellan (March 9, 2016). "Donald Trump helps GOP presidential debates break TV ratings records". *Time*. Retrieved October 8, 2016.
341. ^a "Donald Trump, Ted Cruz Angling For One-On-One Republican Race". *Fortune*. March 6, 2016.
342. ^a Bump, Philip (March 23, 2016). "Why Donald Trump is poised to win the nomination and lose the general election, in one poll". *The Washington Post*.
343. ^a Nussbaum, Matthew (May 3, 2016). "RNC Chairman: Trump is our nominee". *Politico.com*. Retrieved May 4, 2016.
344. ^a Bump, Philip. "Trump got the most GOP votes ever — both for and against him — and other fun facts". *The Washington Post*. Retrieved July 12, 2016.
345. ^a Berenson, Tessa (May 5, 2016). "Donald Trump Tells West Virginia Primary Voters to Stay Home". *Time*.
346. ^a "Fuller picture emerges of man arrested at Trump rally". Associated Press.

External links

- Official website
- Donald Trump at the Internet Movie Database

Donald Trump



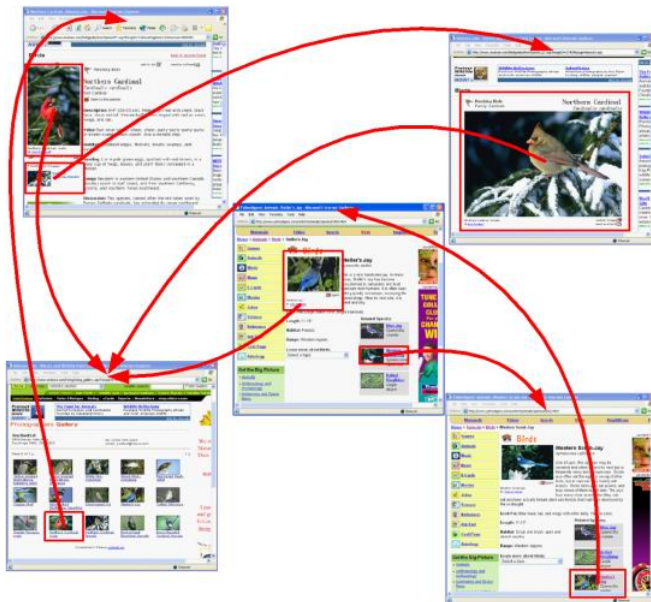
Trump in August 2015

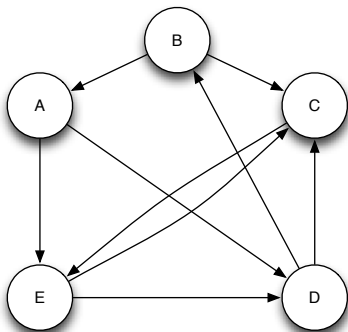
Born	Donald John Trump <div>June 14, 1946 (age 70)</div> <div>Queens, New York City, U.S.</div>
Residence	Manhattan, New York City
Alma mater	Fordham University
Occupation	Businessman · television producer · politician
Net worth	US\$3.7 billion (2016) ^[R]

708. ^a "Algemeiner Honors Joan Rivers, Donald Trump, Yuli Edelstein at Second Annual 'Jewish 100' Gala". *Algemeiner Journal*. Brooklyn, NY. February 5, 2015.
707. ^a Hasocut, Henry (March 27, 2015). "2015 New Jersey Boxing Hall of Fame Inductees". *New Jersey Boxing Hall of Fame*.
708. ^a "MC-LEF Events". Marine Corps–Law Enforcement Foundation. 2015. Archived from the original on August 19, 2015. "Donald Trump received our Commandant's Leadership Award."
709. ^a Madan, Monique (March 4, 2015). "Donald Trump gets his key to Dorale". *The Miami Herald*. Miami. Archived from the original on July 8, 2015. Retrieved August 19, 2015.
710. ^a Hidalgo, Daniel (August 5, 2015). "Doral lets Donald Trump keep key to city; also gives initial OK to four new developments". *The Miami Herald*. Miami. Archived from the original on August 19, 2015. Retrieved August 19, 2015.
711. ^a Elmers, Renee (April 21, 2016). "Donald Trump: 'The rule breaker'". *Time*.

Library resources about
Donald Trump

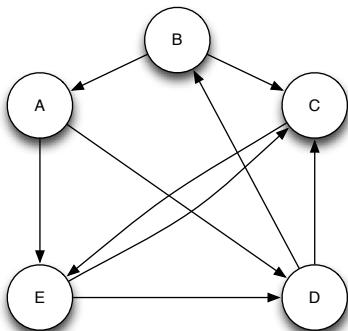
La Web como grafo



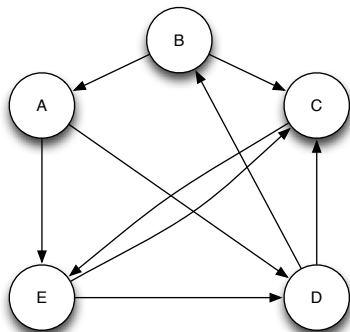


Matriz de adyacencia

$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

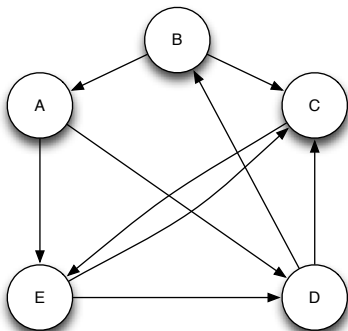


In-degree
Out-degree



Matriz de transición

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$



Descripción / clustering /
clasificación de grafos

Subgrafo

A priori algorithm

Definición probabilista

El PageRank de una página web es la probabilidad de que un internauta acabe en dicha web tras navegar aleatoriamente por Internet desde un punto de inicio al azar.

Random Walk

Otra definiciones

- ▶ Álgebra
- ▶ Sistema dinámico

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = v \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$a_{11} \cdot e_1 + a_{12} \cdot e_2 + a_{13} \cdot e_3 + a_{14} \cdot e_4 + a_{15} \cdot e_5 = v \cdot e_1$$

$$a_{21} \cdot e_1 + a_{22} \cdot e_2 + a_{23} \cdot e_3 + a_{24} \cdot e_4 + a_{25} \cdot e_5 = v \cdot e_2$$

$$a_{31} \cdot e_1 + a_{32} \cdot e_2 + a_{33} \cdot e_3 + a_{34} \cdot e_4 + a_{35} \cdot e_5 = v \cdot e_3$$

$$a_{41} \cdot e_1 + a_{42} \cdot e_2 + a_{43} \cdot e_3 + a_{44} \cdot e_4 + a_{45} \cdot e_5 = v \cdot e_4$$

$$a_{51} \cdot e_1 + a_{52} \cdot e_2 + a_{53} \cdot e_3 + a_{54} \cdot e_4 + a_{55} \cdot e_5 = v \cdot e_5$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = v \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$a_{11} \cdot e_1 + a_{12} \cdot e_2 + a_{13} \cdot e_3 + a_{14} \cdot e_4 + a_{15} \cdot e_5 = v \cdot e_1$$

$$a_{21} \cdot e_1 + a_{22} \cdot e_2 + a_{23} \cdot e_3 + a_{24} \cdot e_4 + a_{25} \cdot e_5 = v \cdot e_2$$

$$a_{31} \cdot e_1 + a_{32} \cdot e_2 + a_{33} \cdot e_3 + a_{34} \cdot e_4 + a_{35} \cdot e_5 = v \cdot e_3$$

$$a_{41} \cdot e_1 + a_{42} \cdot e_2 + a_{43} \cdot e_3 + a_{44} \cdot e_4 + a_{45} \cdot e_5 = v \cdot e_4$$

$$a_{51} \cdot e_1 + a_{52} \cdot e_2 + a_{53} \cdot e_3 + a_{54} \cdot e_4 + a_{55} \cdot e_5 = v \cdot e_5$$

Cualquier matriz estocástica izquierda (columna) tiene 1 como valor propio

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = a_{11} \cdot e_1 + \frac{1}{2} \cdot e_2 + a_{13} \cdot e_3 + a_{14} \cdot e_4 + a_{15} \cdot e_5$$

$$e_2 = a_{21} \cdot e_1 + a_{22} \cdot e_2 + a_{23} \cdot e_3 + \frac{1}{2} \cdot e_4 + a_{25} \cdot e_5$$

$$e_3 = a_{31} \cdot e_1 + \frac{1}{2} \cdot e_2 + a_{33} \cdot e_3 + \frac{1}{2} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_4 = \frac{1}{2} \cdot e_1 + a_{42} \cdot e_2 + a_{43} \cdot e_3 + a_{44} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_5 = \frac{1}{2} \cdot e_1 + a_{52} \cdot e_2 + 1 \cdot e_3 + a_{54} \cdot e_4 + a_{55} \cdot e_5$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{1}{4} \cdot e_4 + \frac{1}{2} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_4 = \frac{1}{2} \cdot e_1 + \frac{1}{2} \cdot e_5$$

$$e_5 = \frac{1}{2} \cdot e_1 + e_3$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{3}{4} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_4 = \frac{1}{2} \cdot e_1 + \frac{1}{2} \cdot e_5$$

$$e_5 = \frac{1}{2} \cdot e_1 + e_3$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{3}{4} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_4 = \frac{1}{8} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$e_5 = \frac{1}{8} \cdot e_4 + e_3$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{3}{4} \cdot e_4 + \frac{1}{2} \cdot e_5$$

$$\frac{7}{8} \cdot e_4 = \frac{1}{2} \cdot e_5$$

$$e_5 = \frac{1}{8} \cdot e_4 + e_3$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{3}{4} \cdot e_4 + \frac{7}{8} \cdot e_4$$

$$e_4 = e_4$$

$$e_5 = \frac{7}{4} \cdot e_4$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{13}{8} \cdot e_4$$

$$e_4 = e_4$$

$$e_5 = \frac{7}{4} \cdot e_4$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \cdot e_4 \\ \frac{1}{2} \cdot e_4 \\ \frac{13}{8} \cdot e_4 \\ e_4 \\ \frac{7}{4} \cdot e_4 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{13}{8} \cdot e_4$$

$$e_4 = e_4$$

$$e_5 = \frac{7}{4} \cdot e_4$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \frac{e_4}{8} \cdot \begin{bmatrix} 2 \\ 4 \\ 13 \\ 8 \\ 14 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{13}{8} \cdot e_4$$

$$e_4 = e_4$$

$$e_5 = \frac{7}{4} \cdot e_4$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 4 \\ 13 \\ 8 \\ 14 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 13 \\ 8 \\ 14 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{13}{8} \cdot e_4$$

$$e_4 = 8$$

$$e_5 = \frac{7}{4} \cdot e_4$$

$$\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,195 \\ 0,341 \end{bmatrix} = \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,195 \\ 0,341 \end{bmatrix}$$

$$e_1 = \frac{1}{4} \cdot e_4$$

$$e_2 = \frac{1}{2} \cdot e_4$$

$$e_3 = \frac{13}{8} \cdot e_4$$

$$e_4 = \frac{8}{41}$$

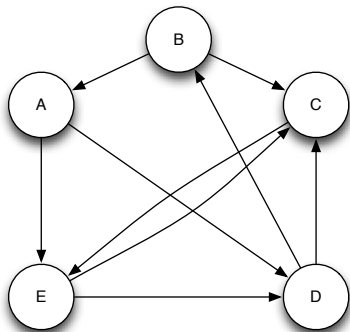
$$e_5 = \frac{7}{4} \cdot e_4$$

PageRank

Sistema dinámico

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$



$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Inicio aleatorio

$$F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

PageRank

Sistema dinámico

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A \cdot F = \begin{bmatrix} \frac{1}{10} \\ \frac{1}{10} \\ \frac{3}{10} \\ \frac{2}{10} \\ \frac{3}{10} \end{bmatrix}$$

PageRank

Sistema dinámico

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^2 \cdot F = \begin{bmatrix} \frac{1}{20} \\ \frac{1}{10} \\ \frac{3}{10} \\ \frac{2}{10} \\ \frac{7}{20} \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} \frac{1}{8} & 0 & 0 & 0 & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{3}{8} \\ 0 & \frac{3}{8} & 0 & \frac{3}{8} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{3}{8} & \frac{1}{4} \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^3 \cdot F = \begin{bmatrix} \frac{1}{20} \\ \frac{1}{10} \\ \frac{13}{40} \\ \frac{1}{5} \\ \frac{13}{40} \end{bmatrix}$$

PageRank

Sistema dinámico

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^* = \begin{bmatrix} \dots \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^* \cdot F = \begin{bmatrix} \dots \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^6 = \begin{bmatrix} 0,08 & 0 & 0,06 & 0,05 & 0,05 \\ 0,09 & 0,14 & 0,09 & 0,06 & 0,11 \\ 0,34 & 0,28 & 0,38 & 0,30 & 0,28 \\ 0,17 & 0,19 & 0,22 & 0,23 & 0,16 \\ 0,31 & 0,39 & 0,25 & 0,36 & 0,41 \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^6 \cdot F = \begin{bmatrix} 0,047 \\ 0,100 \\ 0,316 \\ 0,194 \\ 0,344 \end{bmatrix}$$

$$E = \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,195 \\ 0,341 \end{bmatrix}$$

PageRank

Sistema dinámico

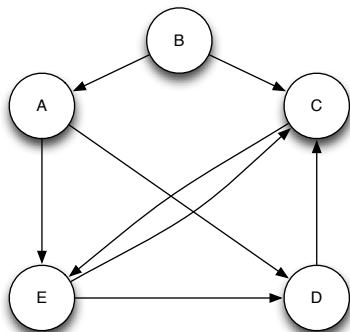
$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^* = \begin{bmatrix} \dots \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^* \cdot F = \begin{bmatrix} \dots \end{bmatrix} \quad E = \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,195 \\ 0,341 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix} \quad A^{10} = \begin{bmatrix} 0,05 & 0,04 & 0,06 & 0,04 & 0,04 \\ 0,09 & 0,11 & 0,09 & 0,10 & 0,10 \\ 0,32 & 0,31 & 0,33 & 0,32 & 0,31 \\ 0,20 & 0,18 & 0,20 & 0,20 & 0,19 \\ 0,34 & 0,36 & 0,32 & 0,36 & 0,36 \end{bmatrix} \quad F = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$$

$$A^{10} \cdot F = \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,194 \\ 0,342 \end{bmatrix}$$

$$E = \begin{bmatrix} 0,049 \\ 0,098 \\ 0,317 \\ 0,195 \\ 0,341 \end{bmatrix}$$



Matriz de adyacencia

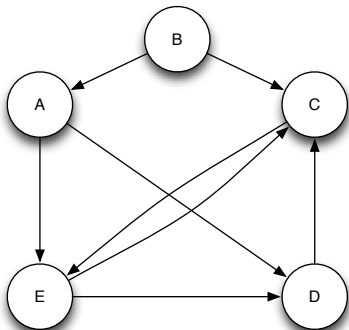
$$D = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Matriz de transición

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$

PageRank

Corrección



PageRank

$$\begin{bmatrix} 0 \\ 0 \\ 0,4 \\ 0,2 \\ 0,4 \end{bmatrix}$$

$$\text{PageRank} = A^k \cdot F$$

$$\text{PageRank} = A \cdot \text{PageRank}$$

$$\text{PageRank} = M^k \cdot F$$

$$M = (1 - p) \cdot A + p \cdot B$$

$$p \in [0, 1]$$

$$B = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

$$\text{PageRank} = ((1 - p) \cdot A + p \cdot B) \cdot \text{PageRank}$$

$$p \in [0, 1]$$

$$\frac{1}{n} \cdot e_1 + \frac{1}{n} \cdot e_2 + \cdots + \frac{1}{n} \cdot e_i + \cdots + \frac{1}{n} \cdot e_n = \frac{1}{n}, \sum_{i=1}^n e_i = 1$$

$$B = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

$$\text{PageRank} = (1 - p) \cdot A \cdot \text{PageRank} + p \cdot b$$

$$p \in [0, 1]$$

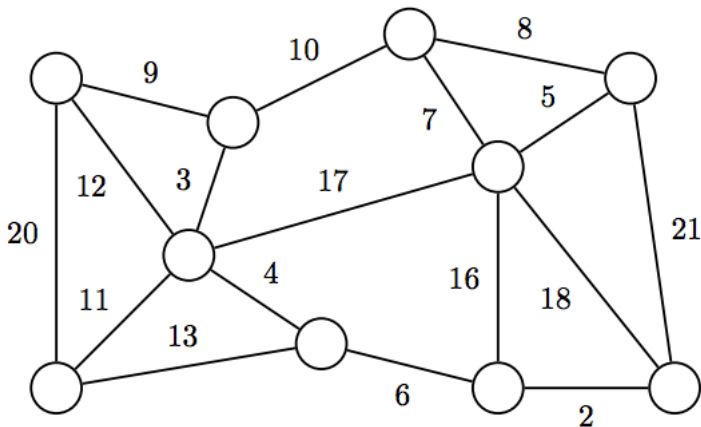
$$b = \begin{bmatrix} \frac{1}{n} \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 & 0 & 0 \end{bmatrix}$$

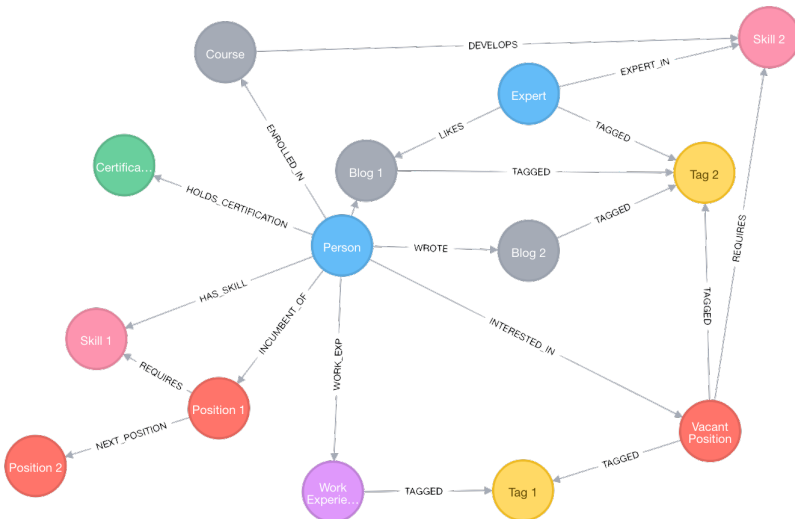
$$p = 0,15$$

$$M = \begin{bmatrix} 0,030 & 0,455 & 0,03 & 0,03 & 0,030 \\ 0,030 & 0,030 & 0,03 & 0,03 & 0,030 \\ 0,030 & 0,455 & 0,03 & 0,88 & 0,455 \\ 0,455 & 0,030 & 0,03 & 0,03 & 0,455 \\ 0,455 & 0,030 & 0,88 & 0,03 & 0,030 \end{bmatrix}$$

$$E = \begin{bmatrix} 0,043 \\ 0,030 \\ 0,366 \\ 0,201 \\ 0,360 \end{bmatrix}$$



Grafos



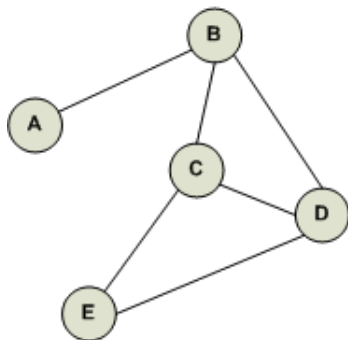


Fig 1. Undirected Graph

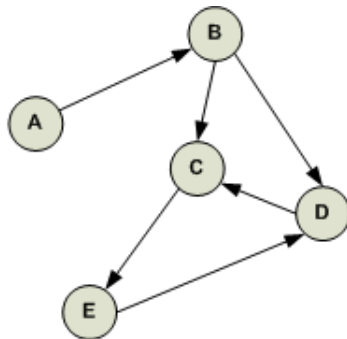
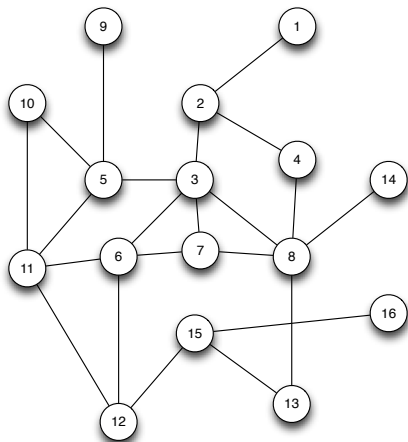


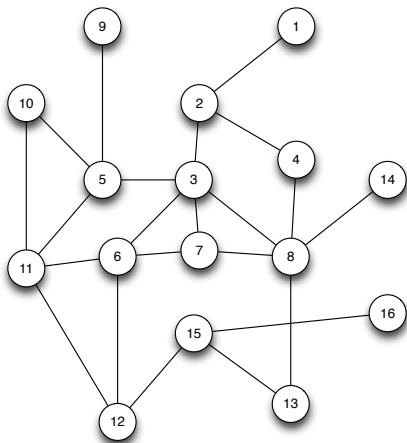
Fig 2. Directed Graph

Personalized PageRank (PPR)



$$\text{rank} = M' \cdot \text{rank}$$

Personalized PageRank (PPR)

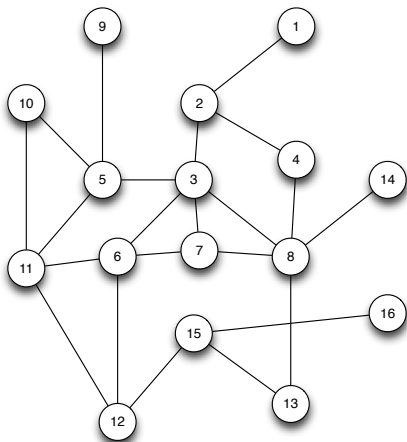


$$\text{rank} = (1 - p) \cdot M \cdot \text{rank} + p v$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁰
1	100
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0

Personalized PageRank (PPR)

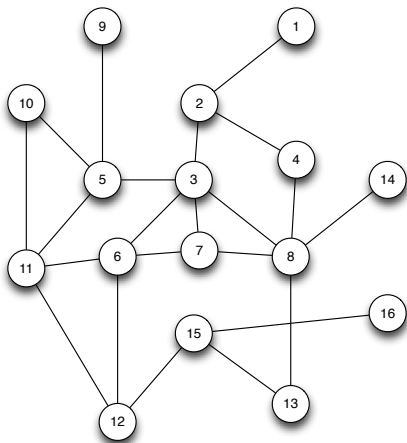


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ¹
1	15
2	85
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0

Personalized PageRank (PPR)

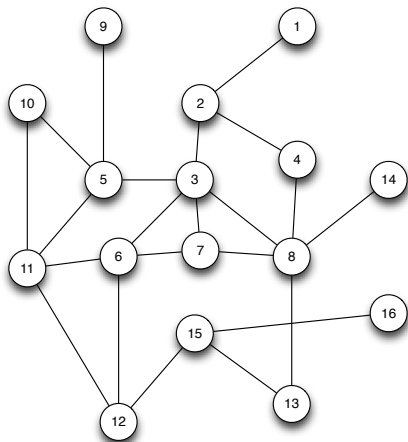


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ²
1	39
2	13
3	24
4	24
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0

Personalized PageRank (PPR)

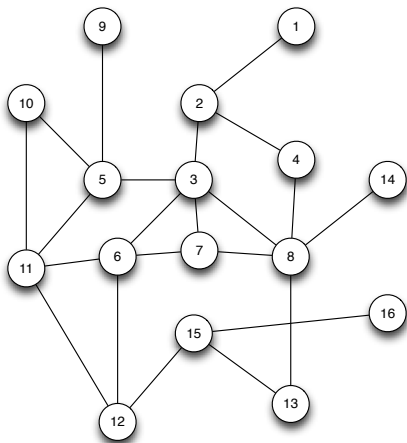


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ³
1	19
2	48
3	4
4	4
5	4
6	4
7	4
8	14
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0

Personalized PageRank (PPR)

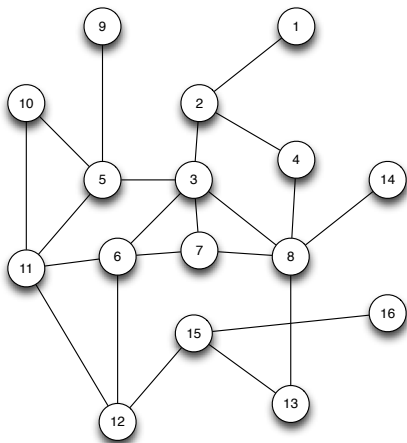


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁴
1	28
2	18
3	19
4	16
5	1
6	2
7	4
8	3
9	1
10	1
11	2
12	1
13	2
14	2
15	0
16	0

Personalized PageRank (PPR)

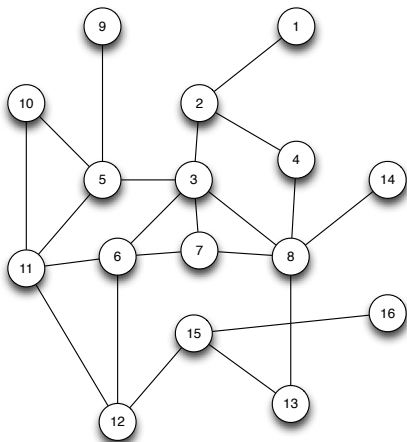


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁵
1	20
2	34
3	7
4	6
5	5
6	5
7	4
8	14
9	0
10	1
11	1
12	1
13	1
14	1
15	1
16	0

Personalized PageRank (PPR)

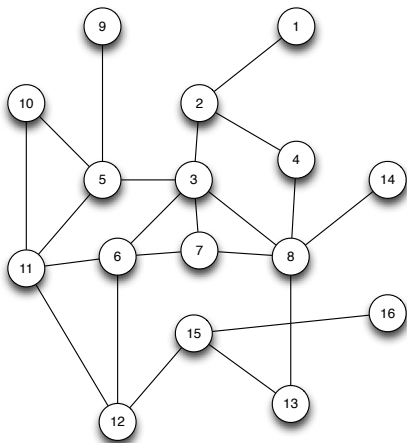


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁶
1	25
2	21
3	15
4	12
5	2
6	3
7	5
8	6
9	1
10	1
11	2
12	2
13	3
14	2
15	0
16	0

Personalized PageRank (PPR)

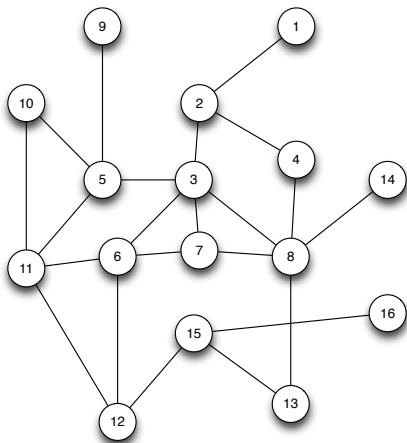


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁷
1	21
2	29
3	9
4	7
5	4
6	5
7	4
8	12
9	0
10	1
11	2
12	1
13	1
14	1
15	2
16	0

Personalized PageRank (PPR)

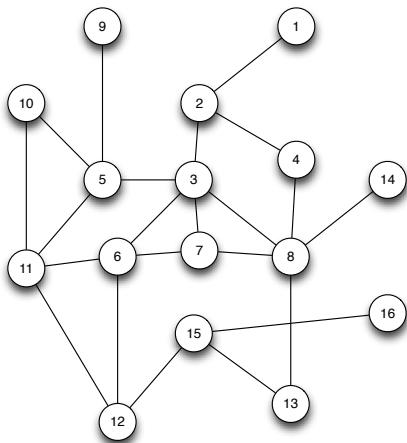


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁸
1	23
2	22
3	13
4	10
5	3
6	4
7	5
8	7
9	1
10	1
11	3
12	2
13	3
14	2
15	1
16	1

Personalized PageRank (PPR)

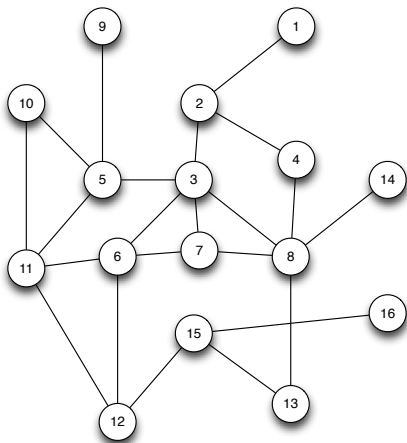


$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ⁹
1	21
2	26
3	10
4	7
5	4
6	5
7	4
8	11
9	1
10	1
11	2
12	2
13	1
14	1
15	2
16	0

Personalized PageRank (PPR)



$$rank = (1 - p) \cdot M \cdot rank + pv$$

$$v_i = \frac{1}{|\mathbf{T}|} \mathbf{1}_{[i \in \mathbf{T}]}$$

Nodo	Rank ¹⁰
1	22
2	23
3	12
4	9
5	3
6	4
7	5
8	8
9	1
10	1
11	3
12	2
13	2
14	2
15	1
16	1

Hubs and authorities

Authorities: Webs con el mejor contenido relevante

Hubs: Webs que apuntan eficientemente al contenido relevante

Procedimiento

Dada una consulta Q :

- ▶ Coger el subgrafo R_Q formado por las N ($= 200$) webs que más referencias a Q contengan
- ▶ Obtener S_Q , el subgrafo que amplía R_Q con webs apuntadas o que apuntan a las webs de R_Q
- ▶ Calcular un peso a_i y h_i para cada web (w_i) de S_Q :

$$a_i = \sum_{j \in IN(w_i)} h_j$$

$$h_i = \sum_{j \in OUT(w_i)} a_j$$

donde $IN(w_i)$ agrupa todas las webs que apuntan a w_i y $OUT(w_i)$ agrupa todas las webs apuntadas por w_i .

Procedimiento

Dada una consulta Q :

- ▶ Inicializar el vector de pesos \mathbf{h} todo a 1.
- ▶ Actualizar los valores iterativamente:

$$\mathbf{h} = \mathbf{D} \cdot \mathbf{a}$$

$$\mathbf{a} = \mathbf{D}^t \cdot \mathbf{h}$$

Aunque también se pueden ver como vectores autocomputados:

$$\mathbf{h}^k = (\mathbf{D} \cdot \mathbf{D}^t) \cdot \mathbf{h}^{k-1}$$

$$\mathbf{a}^k = (\mathbf{D}^t \cdot \mathbf{D}) \cdot \mathbf{a}^{k-1}$$

Inferir nuevas relaciones de *cierto* tipo

PRA

Algoritmo en tres fases

- ▶ Buscar caminos (y seleccionar)
- ▶ Rellenar matriz con probabilidades
- ▶ Aprender clasificador

Consideraciones

- ▶ Datos iniciales: Pares reales del tipo de relación en cuestión
- ▶ ¿Pares negativos?
- ▶ Random walk with restart (RWR ó PPR)
- ▶ Con la matriz rellena, se puede aprender cualquier clasificador

Inferir nuevas relaciones de *cierto* tipo

PRA

Primera fase

- ▶ Lanzar múltiples random walks desde un componente de uno de los pares de nodos de aprendizaje
- ▶ Guardar todos los caminos que haya llevado (en algún momento del proceso) a pasar por el otro nodo del par

Ejemplo: Relación “juega_a”, con pares reales tales como (Real Sociedad, Fútbol masculino) ó (Espanyol, Fútbol masculino)
Algunos caminos encontrados:

compite_en + competición_de
compite_en + ¬compite_en + compite_en + competición_de
juega_contra + compite_en + competición_de
equipo_de + ¬equipo_de + compite_en + competición_de

Inferir nuevas relaciones de *cierto* tipo

PRA

Segunda fase

- ▶ Para cada camino detectado –columna–, rellenar cada celda con la probabilidad de llegar al *target* (t_i) desde el *origin* (o_i), dado un par $-r_i = (o_i, t_i)$, fila– usando el camino en cuestión
- ▶ Generar los negativos y completar igualmente el dataset

	camino_1	camino_2	...	camino_m	C
(r_1)	$p_{c_1}(o_1, t_1)$	$p_{c_2}(o_1, t_1)$...	$p_{c_m}(o_1, t_1)$	+
(r_2)	$p_{c_1}(o_2, t_2)$	$p_{c_2}(o_2, t_2)$...	$p_{c_m}(o_2, t_2)$	+
...
(r_{n+1})	$p_{c_1}(o_{n+1}, t_{n+1})$	$p_{c_2}(o_{n+1}, t_{n+1})$...	$p_{c_m}(o_{n+1}, t_{n+1})$	–
(r_{n+2})	$p_{c_1}(o_{n+2}, t_{n+2})$	$p_{c_2}(o_{n+2}, t_{n+2})$...	$p_{c_m}(o_{n+2}, t_{n+2})$	–
...

Inferir nuevas relaciones de *cierto* tipo

PRA

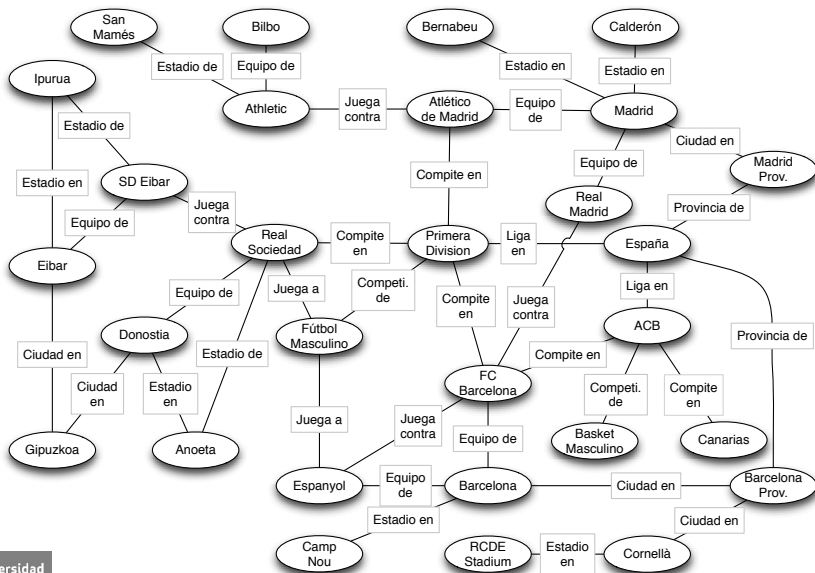
Tercera fase

- ▶ Aprender un clasificador (regresión logística)
- ▶ Usar para predecir (inferir) nuevos pares de la relación

	camino_1	camino_2	...	camino_m	C
(r_1)	$p_{c_1}(o_1, t_1)$	$p_{c_2}(o_1, t_1)$...	$p_{c_m}(o_1, t_1)$	+
(r_2)	$p_{c_1}(o_2, t_2)$	$p_{c_2}(o_2, t_2)$...	$p_{c_m}(o_2, t_2)$	+
...
(r_{n+1})	$p_{c_1}(o_{n+1}, t_{n+1})$	$p_{c_2}(o_{n+1}, t_{n+1})$...	$p_{c_m}(o_{n+1}, t_{n+1})$	-
(r_{n+2})	$p_{c_1}(o_{n+2}, t_{n+2})$	$p_{c_2}(o_{n+2}, t_{n+2})$...	$p_{c_m}(o_{n+2}, t_{n+2})$	-
...

Inferir nuevas relaciones de *cierto* tipo

PRA



R. Agrawal, R. Srikant. Fast algorithms for mining association rules. Proc. 20th International Conference on Very Large Data Bases (VLDB), 487-499, 1994.

L. Page, S. Brin, R. Motwani, T. Winograd. The pagerank citation ranking: Bringing order to the web. Tech. Report 1999-66, Stanford InfoLab, 1999.

T. H. Haveliwala. Topic-sensitive pagerank. Proc. 11th International World Wide Web Conference, 517-526, 2002.

J. M. Kleinberg. Authoritative sources in a hyperlinked environment. Journal for the Association of Computing Machinery, 46(19):604-632, 1999.

N. Lao, W. W. Cohen. Relational retrieval using a combination of path-constrained random walks. Machine Learning, 81(1):53-67,

Aprendizaje no supervisado

VC10: Análisis de grafos, PageRank y otros

Félix José Fuentes Hurtado

felixjose.fuentes@campusviu.es

Universidad Internacional de Valencia