

Diapositive 1/18

Clusters, Grafos y Redes

Xavier Polanco

Montevideo, 10/2001clusters, grafos, redes1

Diapositive 2/18

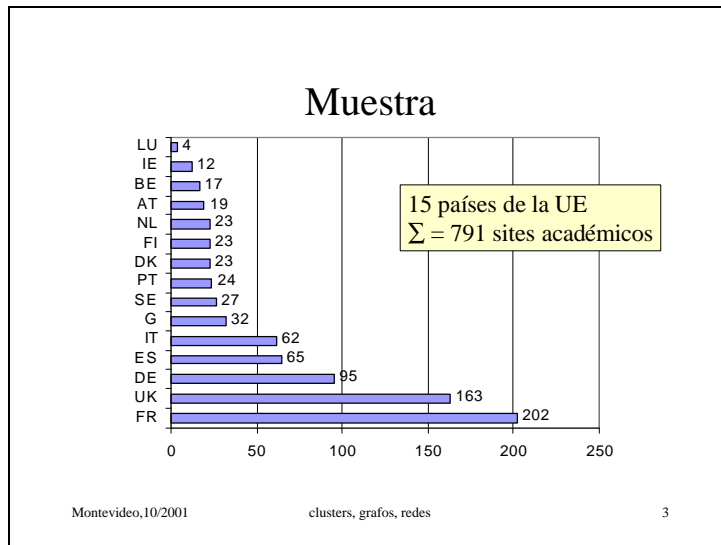
EICSTES

European Indicators, Cyberspace, and the Science-Technology-Economy System (IST-1999-20350)

*Clustering and Mapping European University Web Sites
Sample for Displaying Association and Visualizing Networks*
X. Polanco, M. A. Boudourides, D. Besagni, Ivana Roche
Crete, 18-22/06
Lisboa, 25-27/06
Nancy, 30/07: articulo
Montevideo, 15-18/10

Montevideo, 10/2001clusters, grafos, redes2

Diapositive 3/18



Diapositive 4

Cuestión

¿Cómo analizar Internet y el Web, y por qué?

¿Indicadores de qué tipo?

Montevideo,10/2001 clusters, grafos, redes 4

Diapositive 5

Sociedad post-industrial (Touraine, 1968; Bell,1973)
Sociedad post-moderna (Lyotard, 1979)
Sociedad del conocimiento (Böhme & Stehr, 1986)
Sociedad de la información
Sociedad de redes (network society) (Castells, 1996)

Inteligencia colectiva (P. Levy,1994: anthropologie du cyberspace)
Inteligencia conectada (de Kerckhove, 1997: connected Intelligence; *interactividad-hypertextualidad-conexidad o conectividad*)

Montevideo,10/2001 clusters, grafos, redes 5

Diapositive 6/18

Web = Grafo, Red

“The Web is an example of a social network”
(Chakrabarti, 2000)

“The Web is a directed network, each document can be characterized by the number of incoming, k_{in} , and outgoing, k_{out} , links”
(Barabasi, 2001)

“Consider the directed graph whose nodes correspond to static pages on the Web, and whose arcs correspond to hyperlinks between these pages”
(Broder et al, 2001)

Montevideo,10/2001 clusters, grafos, redes 6

Diapositive 7/18

Red social & comunidad virtual

“When a computer network connects people or organizations, it is a social network”

“Examining the computer-supported social networks in areas as virtual communities”
(Garton, Haythornthwaite, Wellman, 1997)

- CCM: computer-mediated communication
- Information exchange via computer networks

Montevideo,10/2001 clusters, grafos, redes 7

Diapositive 8/18

Análisis de Internet & Web

Watts & Strogatz, 1998, collective dynamics of ‘small-networks’ (*Nature*)

Lawrence & Gilles, 1999: accessibility of information on the web (*Nature*)

Albert et al, 1999, diameter of the WWW (*Nature*)

Huberman & Adamic, 1999, growth dynamics of the WWW (*Nature*)

Albert et al, 2000: error and attack tolerance of complex networks (*Phys.Rev.Lett.*)

Cohen et al, 2000: resilience of the internet to random breakdowns (*Phys.Rev.Lett.*)

Callaway et al, 2000: network robustness and fragility (*Phys.Rev.Lett.*)

Broder et al, 2000: graph structure in the web (*Comp.Netw.*)

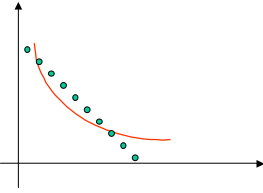
Barabasi, 2001, the physics of the web (*physicsweb.org*)

Montevideo,10/2001 clusters, grafos, redes 8

Diapositive 9/18

Leyes bibliométricas

Lotka (1926) (autores, artículos)
Bradford (1934) (revistas, artículos)
Zipf (1949) (palabras, textos)



Montevideo,10/2001clusters, grafos, redes9

Diapositive 10/18

Power-law

POWER LAW: $P(k) = ck^{-\alpha}$

“the power-law distribution implies that there is an abundance of nodes with only a few links, and a small – but significant – minority that have a very large number of links”
(Faloutsos, 1999; Broder et al, 2000; Barabasi, 2001)

“the majority of documents on the Web have only
Barabasi, 2000)

Montevideo,10/2001clusters, grafos, redes10

Diapositive 11/18

Proposición

Clasificación (Clustering)

↓

Clusters

Análisis de redes sociales

↓

Redes (Networks)

Clusters ↔ Redes (Networks)

Teoría de Grafos

Montevideo,10/2001clusters, grafos, redes11

Diapositive 12/18

Datos

$N = 791$
 $d_{ij}(j_i) = 5.308.204 = 91\%$
 $d_{ii} = 511.470 = 9\%$
 $\Sigma = 5.819.674$
 $Págs = 12.595.809$

Montevideo,10/2001
clusters, grafos, redes
12

Diapositive 13/18

Asociaciones \neq Links

$N' < N$
 $N' = 242 = 30\%$

$C(i,j) = 4$ Asociación

$E(i,j) = C(i,j)^2 / C(i) \times C(j)$

$E(i,j) = [0, 1]$

Montevideo,10/2001
clusters, grafos, redes
13

Diapositive 14/18

Clustering

Parámetros del algoritmo:

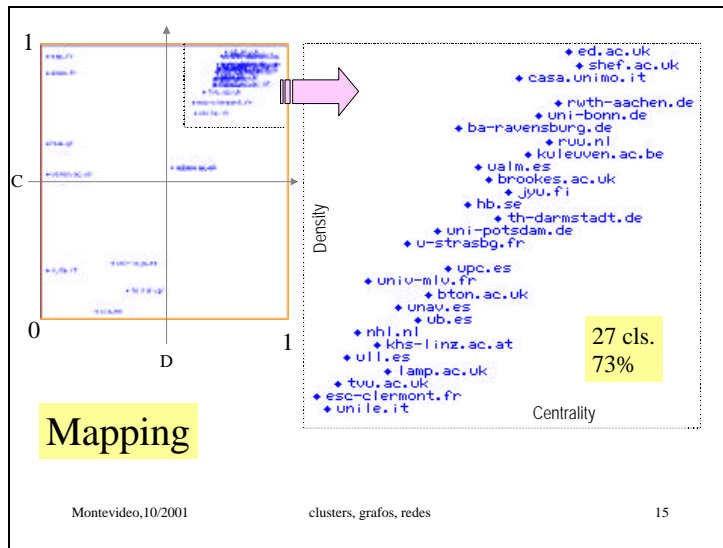
Núm. min. de items	4
Núm. max.	10
Núm. min. de as.int.	10
Núm. max.	20
Núm. max. de as.ext.	10

Resultado:

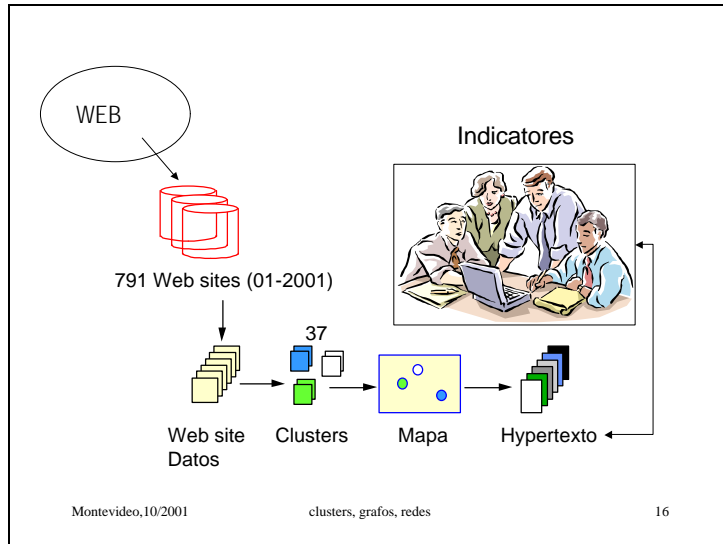
37 clusters	
242 sites componentes	30%
770 sites clasificados	97%

Montevideo,10/2001
clusters, grafos, redes
14

Diapositive 15/18



Diapositive 16/18



Diapositive 17/18

Indicadores

- Índice de Densidad (densité, density)
- Índice de Centralidad (centralité, centrality)
- Racio estructural C/D
- Índice de Transformación (transformation)

$$T = \frac{Cl(t) \cap Cl(t+1)}{Cl(t) + Cl(t+1)}$$

Montevideo,10/2001 clusters, grafos, redes 17

