



## Mapping world-class universities on the web

Jose Luis Ortega \*, Isidro F. Aguillo

*Cybermetrics Lab, IEDCYT-CSIC, Joaquín Costa, 22, 28002 Madrid, Spain*

### ARTICLE INFO

#### Article history:

Received 13 June 2008

Received in revised form 15 October 2008

Accepted 16 October 2008

Available online 29 November 2008

#### Keywords:

Information visualization

Social network analysis

Webometrics

University web

### ABSTRACT

A visual display of the most important universities in the world is the aim of this paper. It shows the topological characteristics and describes the web relationships among universities of different countries and continents. The first 1000 higher education institutions from the Ranking Web of World Universities were selected and their link relationships were obtained from Yahoo! Search. Network graphs and geographical maps were built from the search engine data. Social network analysis techniques were used to analyse and describe the structural properties of the whole of the network and its nodes. The results show that the world-class university network is constituted from national sub-networks that merge in a central core where the principal universities of each country pull their networks toward international link relationships. The United States dominates the world network, and within Europe the British and the German sub-networks stand out.

© 2008 Elsevier Ltd. All rights reserved.

### 1. Introduction

The World Wide Web has become a key medium for promoting and developing the academic, scientific and educational competences of a university. E-learning programs and open access initiatives allow knowledge of these institutions to spread beyond physical boundaries. The Web can hence be used as a way to attract students, scholars and funding from other places, spreading the prestige of these educational institutions all over the world. This has provoked competition between universities to achieve an advantageous visibility on the Web and to improve their position in search engine results.

Web performance has been analysed from different points of view. Web data have been used as an indicator of the educational and scientific activity developed on the Web, relating web indicators with academic outputs (Smith, 2008; Thelwall, 2002a; Thelwall & Harries, 2003, 2004) or bibliometric indicators (Aguillo, Granadino, Ortega, & Prieto, 2006). Visualization of Information (Chen, 2003) has also been a suitable tool for mapping university linkages and showing visual relationships according to several variables. The first attempts used multivariate analysis to plot and group universities (Polanco, Boudourides, Besagni, & Roche, 2001; Vaughan, 2006). Now, Network Analysis offers additional structural and visual possibilities. Heimeriks and Van Den Besselaar (2006) used these analysis techniques to detect four geographical zones in the European Union (EU) academic web space: Scandinavia, UK, Germany and South Europe. Similar results were obtained by Ortega, Aguillo, Cothey, and Scharnhorst (2008), finding that European universities are grouped in local or national sub-networks which are connected with other sub-networks for linguistic or geographical reasons (Thelwall, 2002b; Thelwall, Tang, & Price, 2003). Lately, Thelwall and Zuccala (2008) have studied the link relationship between universities and national web spaces in Europe, describing the European web relationships at the country level.

All these studies were focused on countries such as Spain (Ortega & Aguillo, 2007; Thelwall & Aguillo, 2003) Canada (Vaughan, 2006; Vaughan & Thelwall, 2005) or regions such as the EU (Heimeriks et al., 2006; Ortega et al., 2008; Thelwall & Zuccala, 2008) or Scandinavia (Ortega & Aguillo, 2008a). However, mapping universities at a global level and with a large and consistent population has not been attempted.

\* Corresponding author. Tel.: +34 916022890; fax: +34 916022971.

E-mail addresses: [jortega@orgc.csic.es](mailto:jortega@orgc.csic.es) (J.L. Ortega), [isidro@cindoc.csic.es](mailto:isidro@cindoc.csic.es) (I.F. Aguillo).

## 2. Objectives

The purpose of this paper is to present a visual display of the 1000 most important universities in the World according to the Ranking Web of World Universities ([www.webometrics.info](http://www.webometrics.info)). This map intends to show the topological characteristics of this network and to describe the relationships among universities of different countries and continents. We also present, through network analysis techniques, the most important universities in the network structure, the gateway universities that connect different web spaces or sub-networks and the network core.

## 3. Methods

### 3.1. Data extraction

We have selected the first 1000 higher education institutions from the Ranking Web of World Universities. This ranking orders the universities according to four main web characteristics from their institutional web domain. The volume of contents is measured by the number of pages freely accessible, their visibility by the number of incoming links. The number of rich files is used as an indicator because rich files are a format to spread scientific and technical data and results. And total number of document indexed in Google Scholar is an indicator of the scientific publications on the Web. Each web domain is ranked by the linear aggregation of these indicators, building the webometrics ranking (WR)

$$WR = 4(\text{visibility rank}) + 2(\text{site rank}) + 1(\text{Rich file rank}) + 1(\text{Google Scholar rank})$$

More information about the methodology of Ranking Web of World Universities is available in About the Ranking ([www.webometrics.info/about\\_rank.html](http://www.webometrics.info/about_rank.html)) and in Methodology ([www.webometrics.info/methodology.html](http://www.webometrics.info/methodology.html)).

Together, these indicators make it possible to describe the performance of these academic institutions on the Web, being a complement to other educational and scientific rankings. The main search engines (Google, Yahoo! Search, Live Search and Exalead) are used to implement this ranking (Aguillo, Ortega, & Fernandez, 2008).

Thousand institutions were selected because a digital divide was perceived between North American universities and the rest of the World. If we observe the top 200 list, we detect 59.5% North American universities and 40.36% in the top 500 list (Aguillo et al., 2008). Hence, we have decided to take a wide sample that represents all the continents.

An asymmetric link matrix between this set of universities was built. So, the rows show the inlinks received by each web domain and the columns the number of outlinks that a web domain makes. This asymmetric matrix is necessary in order to build directed networks where the mean of inlinks and outlinks have to be specify. So, links from the university domain A to B are not the same to links from university domain B to A.

Data extraction was made in February 2008 from Yahoo! Search. It allows several search operators in a single search string and the web coverage is rather wide. The following queries were used to obtain links from the university domain (A) to the university domain (B) and vice versa:

site: {university domain (A)} linkdomain:{university domain (B)} and to obtain the total number of pages indexed in the university domain (A):

site: {university domain (A)}

An SQL routine was used to submit the 1 001 000 needed queries to build the link matrix.

### 3.2. Geographical map

We built a geographical map in order to show the distribution of pages and link flows at the level of countries. A base map was downloaded from Blue Marble Geographics web site ([www.bluemarblegeo.com](http://www.bluemarblegeo.com)) in ESRI shapefile format. We used the geographical information system (GIS) software MapViewer 6 to build the final map. This map has two layers: a hutch map which represents the number of web pages by country and a flow map which shows the links between countries. Quantitative data such as number of web pages and in- and outlinks were added through a spreadsheet and assigned to each map zone (country). The classification method used in both layers was Jenks' natural breaks (Jenks, 1963). This method determines the best arrangement of values into classes by iteratively comparing sums of the squared difference between observed values within each class and class means. This method improves the visualization and the interpretation of the results, because it creates more significant differences between classes.

### 3.3. Network graph

A network graph was built with the links between the 1000 university web domains. Several variables have been used in order to add information about the network configuration. Nodes size shows the volume of web pages that each university publishes on the Web, colours represent the nationality of each higher education organization and arc width shows the frequency of in- and outlinks between two university domains.

**Table 1**  
Universities distribution by country (15 first).

Country	Universities	%
United States	369	36.9
United Kingdom	68	6.8
Germany	66	6.6
France	50	5.0
Spain	41	4.1
Canada	39	3.9
Japan	35	3.5
Italy	34	3.4
Australia	30	3.0
China	17	1.7
Taiwan	17	1.7
Sweden	15	1.5
Brazil	14	1.4
The Netherlands	13	1.3
Finland	12	1.2
Rest of the World	180	18.0
Total	1000	100

The software used to visualise the network was Pajek 1.02. Links matrix was converted to a network file (.net) and the nominal variable (countries) was turned in cluster file (.clu) and in vector file (.vec) the discrete one (number of web pages). These three files were merged to present the final graph output. We selected a cut-off of minimum 50 links to improve the network visualization. Also we used the Fruchterman-Reingold algorithm to lay out the network because it is the fast for large networks (de Nooy, Mrvar, & Batagelj, 2005).

Several social network indicators were used to describe the network topology and the main characteristics of the nodes:

- **K-Core:** a sub-network in which each node has at least degree  $k$ . K-Cores detect groups with a strong link density. In free-scale networks, such as the Web, the core with the highest degree is the central core of the network, detecting the set of nodes the network rests on (Seidman, 1983).
- **Degree:** the number of lines connecting a node. This can be normalized (nDegree) by the total number of nodes in the network. In a directed network such as the Web we can count only the incoming links (InDegree) or the outgoing links (OutDegree). In Webometrics, InDegree has been characterised as the visibility of a web domain (Cothey, 2005; Kretschmer & Kretschmer, 2006), while OutDegree as the property to generate web traffic.
- **Betweenness:** the capacity of one node to help connect those nodes that are not directly connected to each other. Its normalization is the percentage over the total number of nodes in the network. From a webometric point of view, this measure allows us to detect hubs or gateways that connect different web sub-networks (Faba-Pérez, Zapico-Alonso, Guerrero-Bote & Moya-Anegón, 2005; Ingwersen, 1998).

## 4. Results

### 4.1. Descriptive analysis

Prior to the link analysis we made a frequency distribution by country of the 1000 selected universities.

Table 1 shows the number of universities by country, listing only the first 15 countries. The United States (US) universities are 36.9% of the entire sample, followed by the United Kingdom (UK) (6.8%) and Germany (6.6%). This distribution is also observed in the Top 200 of the ranking which suggests that there is a digital divide in favour of US universities. The low performance of poorer countries like Russia (0.6%) and India (0.4%) is also clear.

### 4.2. Geographical map<sup>1</sup>

Fig. 1 shows the geographical distribution of web pages by country and the incoming and outgoing links among these countries. Two regions stand out for their large amount of web pages: North America (USA and Canada) and the European Union (EU) zone. The USA is the country with most web pages (50.57%), holding half of the world's academic web pages indexed in Yahoo! Search. It is followed by Germany (7.14%) and the UK (4.28%) in the EU. Besides these zones, notice the web development of Japan (2.35%), Australia (2.35%) and China (2.33%) in the East and Brazil (.94%) in South America. In contrast, two zones have no universities in the sample: Africa (with the exception of South Africa) and the Middle East (with the exception of Israel and Saudi Arabia).

<sup>1</sup> Full colour pictures can be seen in the e-print version of this journal or in: [http://internetlab.cindoc.csic.es/cv/11/world\\_map/](http://internetlab.cindoc.csic.es/cv/11/world_map/).

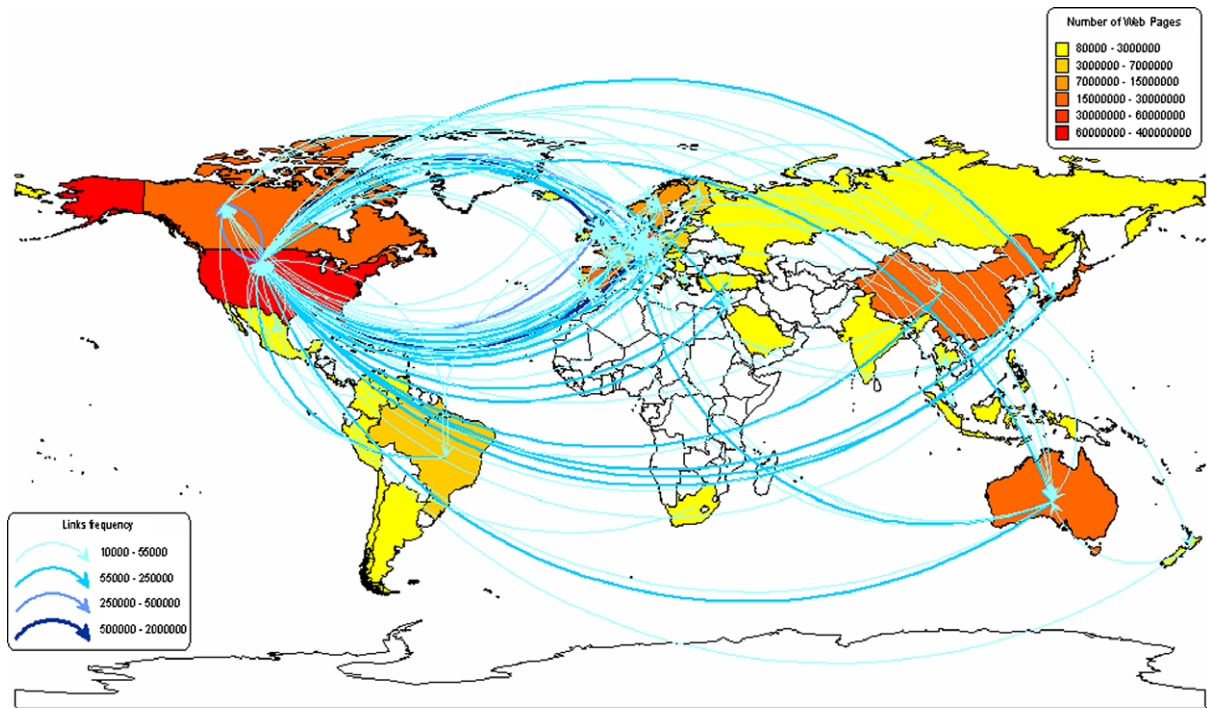


Fig. 1. Geographical map of the distribution of pages by country and their link flows.

From the US position, the upper loops show the outgoing links and the lower loops the incoming ones. The most important link flows are between North American countries and EU countries, while second are links between East Asian and Oceanic countries and the US.

#### 4.3. Network graph

The World class network (Fig. 2) shows clustering because its clustering coefficient ( $C = 527.25$ ) is considerably higher than the same for a random network ( $C = 35.14$ ). And its average path length ( $l = 2.26$ ) is also rather low (Watts & Strogatz, 1998). Visually, small-world properties can be seen through the traversal links that run across the network, connecting distant clusters (Fig. 2). The in and out degree frequency distributions follow a power law trend ( $\gamma_{in} = .81$ ;  $\gamma_{out} = .73$ ), suggesting scale-free properties as well (Barabasi, Albert, & Jeong, 2000).

Fig. 2 shows the graph of the 1000 higher education institutions. First, each university is linked with the universities of its own country. Thus, we can visually detect homogeneous national groups such as Germany (red), the UK (light green) or Japan (orange).<sup>2</sup> However, we can also see that there are countries that do not constitute a compact group such as France (dark blue), Canada (white) and other countries with a small set of universities such as the Netherlands (dark red). This may be because some countries are included in other larger national sub-networks, indeed Canada is related to the US and the Netherlands with the UK. This describes a cumulative process in which each national sub-network is aggregated to other ones like an accretion model.

The graph also shows linguistic (Thelwall et al., 2003) and geographical relationships (Thelwall, 2002a, 2002b). The European countries are located on the top side of the picture, while the bottom side is mainly taken up by Asian and American ones. It shows, for example, that Spanish (purple) universities are between the European and the Latin-American ones. Observe that size is related to link attraction, because the large universities are located in the core of the network. Nevertheless, some countries, specifically Asian ones (China, Japan and Taiwan), have large universities that are far from the core. This may be caused by low development of English pages by these countries (Vaughan & Thelwall, 2004).

The main core of the World network was detected with the  $k$ -cores method. The central core is 116 nodes with degree 93. This highly connected cluster has 98 American universities. The rest are from Canada (7) and Europe (11). Fig. 3 shows in detail this central core, highlighting universities like Harvard, Stanford or Massachusetts Institute of Technology (MIT) which

<sup>2</sup> For interpretation of color in Fig. 2, the reader is referred to the web version of this article.

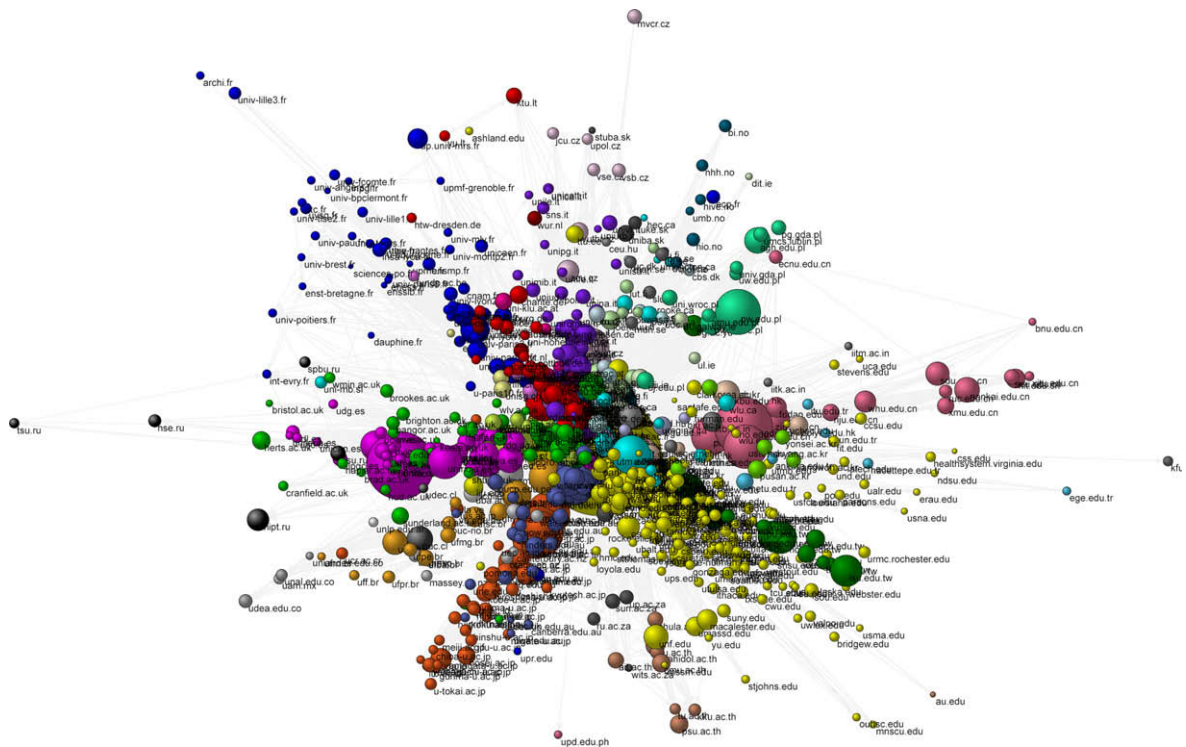


Fig. 2. Network graph of the World class universities on the web ( $N = 1000$  arcs  $\geq 50$  links).

are located in the centre of the graph and attract a huge amount of links from the entire network. Next, the important European universities in the core of the network bring closer their national networks, as with Cambridge and the British network, Trier and the German one and the Swiss Federal Institute of Technology Zurich (ETHZ) and Switzerland. This principal universities act as gateways that connect their sub-networks to the network core, causing a high density of the network. However, there is no presence of Asian, African and Latin-American universities, with the exception of the Israeli, Turkish and some Asiatic ones which are located around the United States sub-network. This may cause that countries do not linguistically and geographically integrated with other countries are mainly connected to the USA ones which have a great weight on the network.

We also calculated the in- and out-degree of each university according to all the network and ranked it. Hence Table 2 shows the top 10 universities by InDegree. United States universities are the most interconnected in the network. MIT (78.1) and the universities of Berkeley (73.5) and Stanford (73.1) are the web domains most linked in the network (Table 2). In contrast, Table 3 shows the top 10 universities by OutDegree. These are the universities that keep the network connected, making outgoing links. This table is also dominated by US universities, particularly the universities of Wisconsin-Madison (47), Stanford (41.8) and Florida (41.2) (Table 3). Notice that both tables only include US universities and the first European universities in the indegree rank are Cambridge in 18th and Leeds in 19th. In the outdegree, the first are ETHZ in 15th and the University of Amsterdam in 22nd.

As above, the World network is the aggregated union of national sub-networks. The betweenness centrality index detects the gateway universities that tend to connect these national sub-networks with the remaining ones. Table 4 shows the principal universities in each country according to their betweenness centrality. For example, the Japanese university with highest betweenness centrality in all the network is the University of Tokyo and the Taiwanese one is the National Taiwan University. Because this network is constituted by national sub-networks (Fig. 2), betweenness centrality may be considered as a suitable indicator in order to show universities with more international scope. Thus, we can identify the outstanding universities in each country such as MIT in the US, Cambridge in the UK or ETHZ in Switzerland. Thus, these universities connect local web spaces internationally. However, there are no German or Spanish universities in the top positions, although both countries have a good position in the network. We suggest that as there is a linguistic factor in the relationships between countries, the German-speaking network is represented by ETHZ and the Spanish-speaking one by the Autonomous National University of Mexico (UNAM). Moreover, the betweenness index is rather close to the degree indicators, so we can state that these universities are the most important in their national or linguistic sub-network.

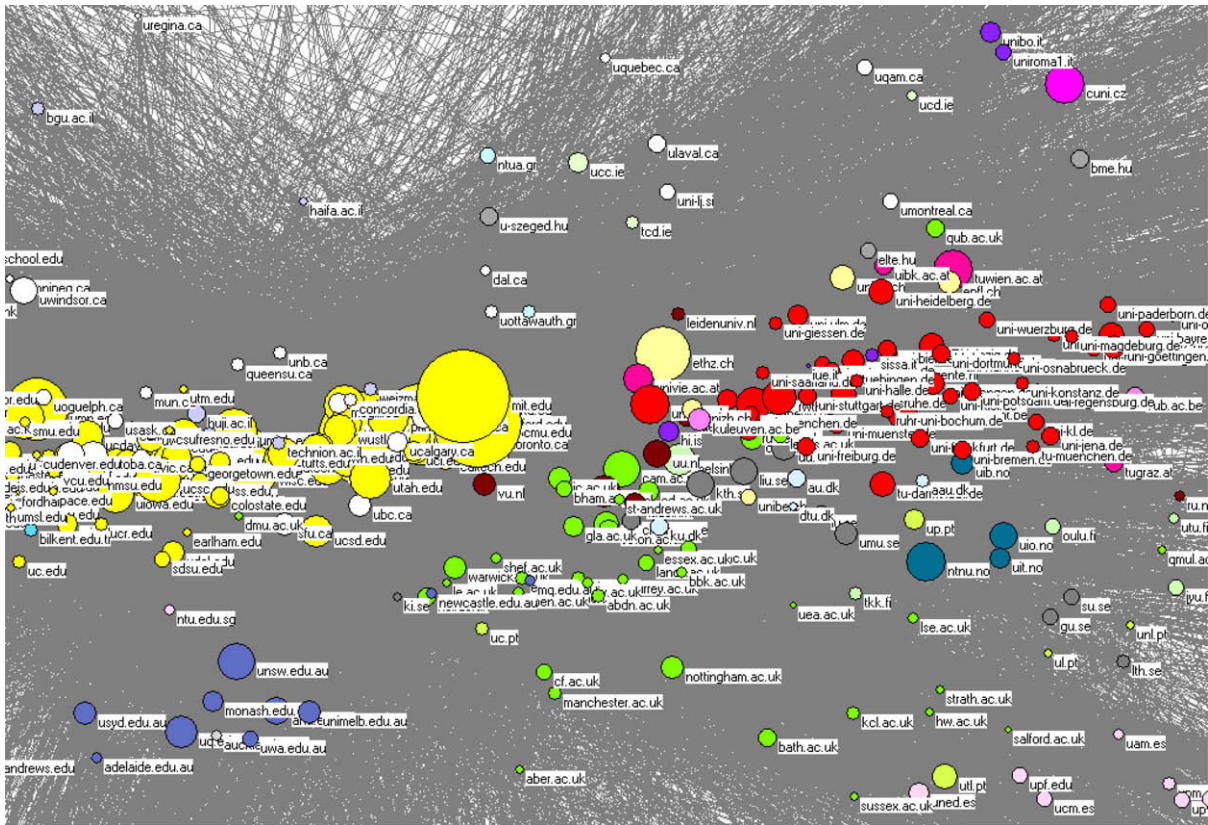


Fig. 3. Detailed view of the central core of the network.

**Table 2**  
First 10 universities by their InDegree.

University	Web domain	InDegree	NrmInDeg
Massachusetts Institute of Technology	mit.edu	781	78.1
University of California, Berkeley	berkeley.edu	735	73.5
Stanford University	stanford.edu	731	73.1
University of Illinois at Urbana-Champaign	uiuc.edu	666	66.6
Harvard University	harvard.edu	634	63.4
University of Michigan	umich.edu	634	63.4
University of Wisconsin-Madison	wisc.edu	629	62.9
University of Texas at Austin	utexas.edu	589	58.9
Cornell University	cornell.edu	557	55.7
University of Washington	washington.edu	555	55.5

**Table 3**  
First 10 universities by their OutDegree.

University	Web domain	OutDegree	NrmOutDeg
University of Wisconsin-Madison	wisc.edu	470	47.0
Stanford University	stanford.edu	418	41.8
University of Florida	ufl.edu	412	41.2
University of California, Berkeley	berkeley.edu	411	41.1
University of Washington	washington.edu	390	39.0
Massachusetts Institute of Technology	mit.edu	378	37.8
University of Illinois at Urbana-Champaign	uiuc.edu	369	36.9
Carnegie Mellon University	cmu.edu	365	36.5
University of Pennsylvania	upenn.edu	360	36.0
Harvard University	harvard.edu	356	35.6

**Table 4**

First 10 universities by their betweenness in each countries.

Country	University	web domain	Betweenness	nBetweenness
US	Massachusetts Institute of Technology	mit.edu	65422	6.54
UK	University of Cambridge	cam.ac.uk	20037	2.00
CH	Swiss Federal Institute of Technology Zurich	ethz.ch	18584	1.86
FR	Jussieu Campus	jussieu.fr	13280	1.33
JP	University of Tokyo	u-tokyo.ac.jp	12529	1.25
FI	University of Helsinki	helsinki.fi	9489	0.95
MX	Autonomous National University of Mexico	unam.mx	7019	0.70
CA	University of British Columbia	ubc.ca	6813	0.68
TW	National Taiwan University	ntu.edu.tw	6604	0.66
BO	University of Bologna	unibo.it	6397	0.64

## 5. Discussion

For some while now, the use of search engine data has been discussed because of the instability of their results over a short time period (Bar-Ilan, 1998; Rousseau, 1997), the weakness of their search operators (Ingwersen, 1998) and the unreliability of their databases (Sullivan, 2003). However, more recent studies have shown that current search engines have improved their consistency and reliability (Bar-Ilan, 2002, 2004, 2005a). Although their technical features have considerably improved, the coverage of their databases and the harvesting process are key issues to discuss. Bar-Ilan (2005b) detects that some search engines have serious problems indexing and retrieving non-Latin characters such as Japanese, Chinese or Russian. Vaughan and Thelwall (2004) showed that there is a local bias in favour of US and against East Asian web sites which are underrepresented in the search engines. Our work may be affected by these biases because large East Asian universities' web domains are remotely located in the graph (Fig. 2), although they have many web pages. The great presence of the US universities may be slightly affected by these coverage biases as well. It is not easy to solve these coverage problems because we will have to compare how different search engines harvest web pages. Using a crawler will avoid these problems; however, to crawl 1000 web domain is a hard task. Hence, interpreting these results must take into account these biases.

The link flows and web page distribution in the geographical map (Fig. 1) follow a similar pattern to the European Union (Ortega & Aguillo, 2008a,b). Countries with many web pages attract and make more links than others, confirming the strong relationship between web pages and links (Katz & Cothey, 2006; Thelwall & Harries, 2003). The network graph also shows similar results to previous work. The World-class universities are grouped in local or national sub-networks which are connected with other sub-networks for linguistic or geographical reasons (Heimeriks et al., 2006; Ortega et al., 2008). Several "gateway" universities act as hubs/authorities that connect the national communities or sub-networks between them (Barabasi & Albert, 1999; Kleinberg, 1999). This causes the reduction of the distances between nodes and explains the emergence of small-world phenomena on the Web (Bjorneborn, 2003).

## 6. Conclusions

The world-class university network graph is comprised of national sub-networks that merge in a central core where the principal universities of each country pull their networks toward international link relationships. This network rests on the United States, which dominates the world network in conjunction with the aggregation of the European ones, especially the British and the German sub-networks. This situation may be caused mainly by the technological development of these countries and the production of international content, that is, English web pages. This second reason might explain the apparent backward situation of some East Asian countries.

## References

- Aguillo, I. F., Granadino, B., Ortega, J. L., & Prieto, J. A. (2006). Scientific research activity and communication measured with cybermetrics indicators. *Journal of the American Society for Information Science and Technology*, 57(10), 1296–1302.
- Aguillo, I. F., Ortega, J. L., & Fernandez, M. (2008). Webometrics ranking of world universities: Introduction, methodology and future developments. *Higher Education in Europe*, 33(2–3).
- Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- Barabasi, A. L., Albert, R., & Jeong, H. (2000). Scale-free characteristics of random networks: The topology of the world-wide web. *Physica A*, 281(1–4), 69–77.
- Bar-Ilan, J. (1998). On the overlap, the precision and estimated recall of search engines, a case study of the query "Erdos". *Scientometrics*, 42(2), 207–228.
- Bar-Ilan, J. (2002). Methods for measuring search engine performance over time. *Journal of the American Society for Information Science and Technology*, 53(4), 308–319.
- Bar-Ilan, J. (2004). The use of web search engines in information science research. *Annual Review of Information Science and Technology*, 38, 231–288.
- Bar-Ilan, J. (2005a). Expectations versus reality – Search engine features needed for web research at mid 2005. *Cybermetrics*, 9(1), <<http://www.cindoc.csic.es/cybermetrics//articles/v9i1p2.html>>.
- Bar-Ilan, J. (2005b). Comparing rankings of search results on the web. *Information Processing & Management*, 41(6), 1511–1519.
- Bjorneborn, L. (2003). *Small-world link structures across an academic web space: A library and information science approach*. Copenhagen: Royal School of Library and Information Science. <<http://vip.db.dk/lb/phd/phd-thesis.pdf>>.
- Chen, C. (2003). *Mapping scientific frontiers: The quest for knowledge visualization*. London: Springer.

- Cothey, V. (2005). Some preliminary results from a link-crawl of the European Union research area web. In P. Ingwersen & B. Larsen (Eds.), *Proceeding of the 10th international conference of the international society for scientometrics and informetrics*. Stockholm: Karolinska University Press.
- de Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge, UK: Cambridge University Press.
- Faba-Pérez, C., Zapico-Alonso, F., Guerrero-Bote, V. P., & Moya-Anegón, F. de (2005). Comparative analysis of webometric measurements in thematic environments. *Journal of the American Society for Information Science and Technology*, 56(8), 779–785.
- Heimeriks, G., & Van Den Besselaar, P. (2006). Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production. *Cybermetrics*, 10(1,1). <<http://www.cindoc.csic.es/cybermetrics/articles/v10i1p1.html>>.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236–243.
- Jenks, G. F. (1963). Generalization in statistical mapping. *Annals of the Association of American Geographers*, 53, 15–26.
- Katz, J. S., & Cothey, V. (2006). Web indicators for complex innovation systems. *Research Evaluation*, 15(2), 85–95.
- Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), 604–632.
- Kretschmer, H., & Kretschmer, T. (2006). Application of a new centrality measure for social network analysis to bibliometric and webometric data. In *Proceeding of the IEEE international conference on digital information management (ICDIM)*. Bangalore, India: IEEE.
- Ortega, J. L., & Aguillo, I. F. (2007). La Web académica española en el contexto del Espacio Europeo de Educación Superior: Estudio exploratorio. *El profesional de la información*, 16(5), 417–425.
- Ortega, J. L., & Aguillo, I. F. (2008a). Visualization of the Nordic academic web: Link analysis using social network tools. *Information Processing & Management*, 44(4), 1624–1633.
- Ortega, J. L., & Aguillo, I. F. (2008b). Linking patterns in the European Union's countries: Geographical maps of the European academic web space. *Journal of Information Science*, 34(5), 705–714. <[http://internetlab.cindoc.csic.es/cv/11/Ortega\\_Aguillo\\_2008.pdf](http://internetlab.cindoc.csic.es/cv/11/Ortega_Aguillo_2008.pdf)>.
- Ortega, J. L., Aguillo, I. F., Cothey, V., & Scharnhorst, A. (2008). Maps of the academic web in the European higher education area – An exploration of visual web indicators. *Scientometrics*, 74(2), 295–308.
- Polanco, X., Boudourides, M., Besagni, D., & Roche, I. (2001). Clustering and mapping European university web sites sample for displaying associations and visualizing networks. In *Proceeding of the NTIS&ETK 2001 conference*. Hersonissos, Crete.
- Rousseau, R. (1997). Situations: An exploratory study. *Cybermetrics*, 1(1). <<http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html>>.
- Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks*, 5, 269–287.
- Smith, A. G. (2008). Benchmarking Google scholar with the New Zealand PBRF research assessment exercise. *Scientometrics*, 74(2), 309–316.
- Sullivan, D. (2003). Google Dance Syndrome Strikes Again. SearchEngineWatch.Com. <<http://searchenginewatch.com/showPage.html?page=3114531>>.
- Thelwall, M. (2002a). A research and institutional size based model for national university web site interlinking. *Journal of Documentation*, 58(6), 683–694.
- Thelwall, M. (2002b). Evidence for the existence of geographic trends in university web site interlinking. *Journal of Documentation*, 58(5), 563–574.
- Thelwall, M., & Aguillo, I. F. (2003). La salud de las Web universitarias españolas. *Revista Española De Documentación Científica*, 26(3).
- Thelwall, M., & Harries, G. (2003). The connection between the research of a university and counts of links to its web pages: An investigation based upon a classification of the relationships of pages to the research of the host university. *Journal of the American Society for Information Science and Technology*, 54(7), 594–602.
- Thelwall, M., & Harries, G. (2004). Do the web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149–159.
- Thelwall, M., Tang, R., & Price, L. (2003). Linguistic patterns of academic web use in Western Europe. *Scientometrics*, 56(3), 417–432.
- Thelwall, M., & Zuccala, A. (2008). A university-centered European union link analysis. *Scientometrics*, 75(3), 407–420.
- Vaughan, L. (2006). Visualizing linguistic and cultural differences using web co-link data. *Journal of the American Society for Information Science and Technology*, 57(9), 1178–1193.
- Vaughan, L., & Thelwall, M. (2004). Search engine coverage bias: Evidence and possible causes. *Information Processing & Management*, 40(4), 693–707.
- Vaughan, L., & Thelwall, M. (2005). A modeling approach to uncover hyperlink patterns: The case of Canadian universities. *Information Processing & Management*, 41(2), 347–359.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393, 440–442.