Scientometrics (2014) 99:5–26
DOI 10.1007/s11192-013-1070-0

A multi-level network analysis of web-citations among the world's universities

George A. Barnett · Han Woo Park · Ke Jiang · Chuan Tang · Isidro F. Aguillo

Received: 2 March 2013 / Published online: 5 July 2013

Abstract This paper describes the results of a multi-level network analysis of web-citations among the 1,000 universities with the greatest presence on the world wide web. Using data from January 2011, it describes the web-citation network of the world's universities and ascertains the antecedent factors that determine its structure. At the university level, the network is composed of ten groups, and the most central universities are mainly from the United States. The factors that predict the structure of the network are, whether or not the universities are in the same country, the language of instruction, the size and excellence of the institution (university ranking and the number of Nobel Prizes received), if they offer doctoral degrees, and the infrastructure of its country. Physical distance was not a determinant of the network's structure. At the nation-state level, international connections among a nation's universities are composed of a single cluster with the United States, United Kingdom and Germany at the center. The structure of the international network may be predicted by the countries' overall hyperlink connections, international co-authorships, student flows and the number of Nobel Prizes won by its citizen.

G. A. Barnett · K. Jiang
Department of Communication, University of California, Davis, CA, USA
e-mail: gbarnett@ucdavis.edu

K. Jiang
e-mail: jenny.jiangke@gmail.com

H. W. Park (✉)
Department of Media & Information, Yeungnam University, Gyeongsan-si, South Korea
e-mail: hanpark@ynu.ac.kr

C. Tang
Chengdu Branch of the National Science Library, Chinese Academy of Sciences, Chengdu, People's Republic of China
e-mail: tangc@clas.ac.cn

I. F. Aguillo
The Cybermetrics Lab, National Research Council (CSIC), Madrid, Spain
e-mail: isidro.aguillo@csic.es

Keywords Webometrics · Hyperlink network analysis · Informetrics · World university ranking · International networks · Social network analysis

With the rapid development of information technology, the worldwide web has emerged as a significant channel to improve scholarly communication, accelerate international academic collaboration and the formation of an international academic network. During the last two decades, the international internet-based academic network has been studied at the university level (Almind and Ingwersen 1997; Ingwersen 1998). Scholars have examined the structure of hyperlink network among world universities by analyzing university web visibility (Lee and Park 2012), which has been shown to be an important indicator of a university's academic, scientific and educational competences (Ortega and Aguillo 2009).

Several studies on university web links have explored the academic connections in certain disciplines within particular countries and regions, such as Spain (Ortega and Aguillo 2007), the United Kingdom (Thelwall 2002a, b; Payne and Thelwall 2004; Li et al. 2003), the United States (Tang and Thelwall 2004), Canada (Vaughan and Thelwall 2005), and the European Union (Heimeriks and van den Besselaar 2006; Ortega and Aguillo 2008; Thelwall and Zuccala 2008; Seeber et al. 2012). Research on university web visibility has even expanded to a global scale (Ortega and Aguillo 2009; Lee and Park 2012) focusing on two points, the increased centralization of university web visibility, and the decentralization or diversification of university web visibility (Lee and Park 2012).

The globalization of research has substantially promoted international academic collaborations (Glänzel 2001, p. 101), and dramatically increased the number of co-authored scientific papers (Leydesdorff et al. 2013). However, while the global collaboration network is growing larger and denser, the central group in this network is becoming smaller (Leydesdorff and Wagner 2008). This trend of centralization resonates well with Wallerstein's world system theory (Wallerstein 1974; Chase-Dunn and Grimes 1995), which illustrates the "broader spectrum of global economics (Lee and Park 2012, p. 202)" According to the world system theory, a core group of a few nations with strong economies could use knowledge from the global collaboration network more efficiently than peripheral countries (Leydesdorff and Wagner 2008, p. 317).

International information flows have been extensively examined from the world system perspective (Barnett et al. 1996; Barnett and Salisbury 1996; Moon et al. 2010; Nam and Barnett 2011). World system theory may also be applied to the world's university hyperlink network since research has found that universities may be categorized into the core, the periphery and the semi-periphery according to the economic status of these universities' country/regional location (Lee and Park 2012). Within the structure of the hyperlink network of world universities, those in the core are in proximity to each other, whereas universities from the semi-periphery countries are closer to the core universities than to each other (Park and Thelwall 2006). Scholars generally come to a same conclusion. The U.S. universities are located in the center of hyperlink network, while Asian universities are in the periphery of the network (Lee and Park 2012; Ortega and Aguillo 2009). A similar centralized structure of university hyperlink network appears at the linguistic level. English is the dominant language for university web pages (Thelwall et al. 2003), and universities from English speaking countries (including those where English is not the native language but it is strongly used in the academia, like Scandinavian ones) have greater web visibility than universities from non-English speaking countries (Chung and Park 2012; Lee and Park 2012).

However, the idea of centralization in the study of web visibility is confronted by factors that lead to decentralization such as geographical proximity (Barnett and Choi 1995), linguistic variation (Barnett and Choi 1995), cultural similarity (Barnett and Sung 2006), financial capacity (Lim and Park 2013), and religious homogeneity (Barnett et al. 1999). In light of the above research, scholars found that the structure of the university hyperlink network may not be fully determined by the global economy, and they have examined regional clusters in the network indicative of decentralization and diversification (Park et al. 2011; Chung et al. 2013).

Earlier regional research on university web links has provided several exemplars for the study of decentralization. For example, Thelwall (2002b) explored regional (geographic distances) trends and linguistic patterns (Thelwall et al. 2003) of university web site interlinking. Researchers have expanded the studies of geographic and linguistic influences to the global scale. Lee and Park (2012) claimed that geographical proximity and linguistic division are two important factors for the formation of regional clusters in the hyperlink network of world universities. Ortega and Aguillo (2009) also described regional clusters in terms of national sub-networks, emphasizing the geographic proximity and a common language. In addition, scholars have taken regional macro-policy into account and considered EU's joint R&D policies as a factor for the formation of international clusters of European universities (Lee and Park 2012).

The above analysis sheds lights on two contradictory trends in the study of the structure of the hyperlink network among world universities (Barnett 2011): centralization and decentralization. This contradiction coincides with Tehranian (1990) dual effects hypothesis claiming that new technologies have paradoxical effects on the structure of power. However, the studies of these complex network attributes do not simply rely on the analysis at the single-level of elementary features of network, but a multilevel analysis is needed so as to illustrate the complex structural characteristics (Monge and Contractor 2003). In the study of the hyperlink network among the world's universities, the multilevel analysis includes not only the analysis of descriptive features of this network, such as density, degree, betweenness, eigenvector, page rank, inter-link and co-link, but also the analysis of the properties of nodes relating to university and multiplex networks relating to nation-state. Scholars have conducted multilevel analyses of the hyperlink network among the world's universities at both the university level and the country level, and they attempted to explore the mechanism of this network on the basis of homophily and physical proximity (Monge and Contractor 2003). Many scholars have predicted that the likelihood of communication is substantially greater when similarity and physical proximity of certain variables can be found at both the university and country level.

At the university level, scholars have mainly focused on these variables: size and research output of the university, prestige of the university, the tuition language of the university, the physical distance between them, whether or not they are in the same country, and if they offer doctoral degrees. They have found that larger universities receive more web links (Thelwall 2002a; Seeber et al. 2012); the reputation of the universities influences numbers of web links, and top-ranked universities have exclusive relationships with one another (Opsahl et al. 2008; Seeber et al. 2012); research productivity is positively correlated to the number of web links among the most prestigious (Seeber et al. 2012); geographic proximity and language similarity ease communication, and thus may be in the origin of a higher number of web links among these universities (Thelwall 2002a, b; Thelwall et al. 2003; Wagner 2008; Ortega and Aguillo 2009; Rivera et al. 2010; Seeber et al. 2012; Lee and Park 2012); universities in the same country and awarding PhD at the same time tend to be more interconnected, and the large generalist PhD awarding universities tend to have more web links across countries (Doz et al. 2000; Seeber et al. 2012).

Compared to the university level, fewer studies have been conducted at the national level. Scholars have described the network structure among world universities mainly in terms of international co-authorships (Leydesdorff and Wagner 2008) and country size (Heimeriks and Van den Besselaar 2006).

Although scholars, such as Seeber et al. (2012), have explored the main factors determining the structure of the hyperlink network among universities by comparing the strength of the above factors at both of the university and national level, their studies are limited to a few countries and there is a need of larger and more diverse group of countries for further study to increase generalizability. Research is needed to examine more variables, especially at the national level related to the structure of the hyperlink network among universities to describe these mechanisms in a more comprehensive way.

Research aims

This paper describes a multi-level network analysis to study the URL-citations among 1,000 universities located in 58 countries with the greatest presence on the world wide web. By analyzing the latest data from January 2011, it describes the URL-citations network of world's universities at both institutional and national level. Because past research has focused on the hyperlink connection among universities this suggests the following research questions: what is the structure of the international university URL-citation network? Does it differ in some systematic way from the international university hyperlink network, from the perspective of world system's theory?

Also, the described research ascertains the antecedent factors that determine the structure of the network among world universities. It examines how such factors suggested by past research impact the network's structure. They include the physical distances between universities, whether or not they are in the same country, the language of instruction, the size, prestige and research excellence of the institution (university ranking and the number of noble prizes received), and if they grant doctoral degrees. This paper enriches the study of the academic network at both the university and national level with the analysis of the countries' overall hyperlink connections, international bandwidth capacity, GDP, GDP/capita, student flows and the number of Nobel Prizes won by its citizens.

Methods

Although there was previous research analyzing the world university academic network, they only included a limited number of institutions, mostly from Western European and North American countries. In order to provide a global perspective, we chose the top 2000 higher education institutions as ranked in the ranking web of universities (www.webometrics.info), plus an additional group of 100 universities from BRIC (Brazil, Russia, India, China) and CIVETS (Colombia, Indonesia, Vietnam, Egypt, Turkey, South Africa) countries to increase the presence of the emerging economies.

Data on the international university hyperlink network were collected using the Google search engine during January 2011. Due to several problems with the Yahoo Site Explorer (now no longer available), it was decided to use an alternative collection methodology described by Peter Hirst in 2008 (http://web.archive.org/web/20080820131725/http://www.universitymetrics.com/) as the G-factor, which consists in measuring the number of times two names of universities are mentioned together. But as the institutions' name can have several variants and this method can be very noisy, we used URL-citations as a way for estimating "hyperlinks", as proposed by Thelwall and Sud 2011; Thelwall et al. 2012. A 2,100 by 2,100 matrix of universities using Google web domains inter-URL-citations was generated. The search query used was "university A webdomain" site: university B webdomain. For example, "harvard.edu" site:stanford.edu was used in order to estimate the number of incoming links from Stanford University to Harvard University. Not all the URL-citations are really links, as for example email addresses in co-authored papers are included in the counting, but in most cases these citations are showing a bond between different institutions and the meaning and the aim is not very different from those intended in a true link analysis (De Maeyer 2012; Sams and Park 2013; Park 2010; Choi and Park 2014; Chung et al. 2013; Khan et al. 2013, 2014; Nam et al. 2013; Park et al. 2011). An intensive check-up (to get rid of noise) was performed because there were many zeros in the cells. Several universities had no ties to other institutions. When there are two domains for some institutions, the largest one was chosen. Finally, we ranked the most interlinked universities and built a 1,000 by 1,000 matrix of universities. This matrix of webdomains inter-citations was also aggregated to the national level.

Data on the antecedent variables came from a number of sources. The physical locations of the individual universities were from Google Maps. The location to distance conversion was performed using an R package called "fields" available at: http://cran.r-project.org/web/packages/fields/index.html. The process was completed using a function that automatically takes a vector of longitude/latitude coordinates and calculates the great-circle distances between all points in the coordinate list. The calculations are done using the 'spherical law of cosines' to convert the distances to arc measures, and assumes a spherical earth with radius of 6378.388 km.

The country in which the universities were located was determined by the country top-level domain (cTLD) of its website or by hand when education (mostly but not ever referring to US universities) is involved. Language of instruction was based on the country's national language in all instances with the exception where English is the language of instruction for higher education (i.e., India, Singapore).

The size of the individual universities for European institutions and whether they offered a doctoral program was taken from the European University Data Collection (EUMIDA 2010). The report was published in December 2010 (http://thedatahub.org/dataset/eumida). For the U.S. schools data were extracted from *College Handbook 2012*, which contain data on over 2,200 4-year colleges and universities (College Board 2012). For the universities from Canada, Asia, Africa, Latin America, and Oceania, the data was gathered from the universities' websites.

The prestige of the universities was obtained from their score on the *U.S. News, World's Best Universities 2012* (http://www.usnews.com/education/worlds-best-universities-rankings/top-400-universities-in-the-world), while the research excellence accounts for the numbers of Nobel Prizes its faculty have received since its inception in 1900. Shared prizes were as each counted as awards to individuals. U.S. News' rating was selected because it contained the most schools in the sample (361). It is based on the QS World University Rankings (www.topuniversities.com/university-rankings/world-university-rankings). However, the latter contained fewer scores for the sample schools (291). The Shanghai Academic Ranking of World Universities (http://www.shanghairanking.com) rated even fewer schools (102). The three indicators of the quality of the universities are very similar. The U.S. News and QS ratings were correlated 1.0, and both measures were correlated to the Shanghai ratings .68. Additionally, data were gathered on the university ratings on six different academic disciplines (Biology, Economics, Computer Science, Mathematics, Physics, International Relations and Political Science, and Medicine) to determine if there were more URL-citations between institutions that were highly rated in specific fields. The results indicated that all discipline specific ratings were strongly related to the schools' overall ratings. The mean correlation between the individual fields and the overall ratings were .83. Universities that were highly rated in one field tended to be highly rated in all fields. Therefore, it was decided to focus only on the overall institutional ratings.

Data on the number of Nobel Prizes won by specific universities and the winners' country of residence (at the time of the award) were gathered from the official website of the prize (http://www.nobelprize.org). Only those subjects where the recipients were academics (Physics, Chemistry, Physiology or Medicine, and Economics) were counted. The Peace and Literature Prizes were not included because the recipients usually are not affiliated with universities.

At the national level, the data on the overall hyperlinks among countries were extracted from the data reported by Barnett and Park (2012). They were collected in November 2010 using Yahoo. International bandwidth capacity, GDP and population were obtained from TeleGeography (http://www.telegeography.com/). Data on international student at the tertiary level were obtained from UNESCO (http://stats.uis.unesco.org/unesco), for the latest year available (usually 2010) for each of the 58 countries in the hyperlink network. Finally, Leydesdorff provided data on international co-authorships and international citations. The co-authorship data comes from Leydesdorff and Wagner (2008). Collected in 2005, it is based on data from the *Science Citation Index* and thus excludes data from the social science and arts and humanities. Table 1 lists the antecedent variables and their sources.

Table 1 Antecedent variables and their sources

Variable	Source
University level	
Physical location	Google maps
Country	cTLD of website
Language of instruction	Country of university
Size of university	Europe—(EUMIDA) (http://thedatahub.org/dataset/eumida)
	U.S.—*College Handbook 2012*
	Asia, Africa, Latin American & Canada—Universities' Websites
Prestige	*U.S. News, World's Best Universities 2012* (http://www.usnews.com/education/)
Nobel Prizes	(http://www.nobelprize.org)
National level	
Total hyperlinks	Barnett and Park (2012)
GDP & population	TeleGeography (http://www.telegeography.com/)
Student exchange	UNESCO (http://stats.uis.unesco.org/unesco)
International co-authorships	Leydesdorff and Wagner (2008)
International citations	*Science Citation Index*

The measures of centrality and the clusters within the network were calculated using the R package ‘network’ (Handcock et al. 2003). Correlations among the network’s measures of centrality and the antecedent variables were performed and linear regression analysis was conducted to predict the structure of the university hyperlink network. At the national level, UCINET was employed to determine measures of density, centrality and the clusters (hierarchical analysis) within the network (Borgatti et al. 2002). To calculate the correlations between the web-citations network and the others used to predict its structure the quadratic assignment procedure (QAP) from UCINET was used (Krackhardt 1987).

Results

A description of the university URL-citation network

The world’s universities’ web-citation network had over 9,600,000 links among the universities. The network density was .606, 60.6 % of the possible links among the universities were present, far higher than the figure obtained by Seeber et al. (2012) for a group that included besides top research-oriented universities a large number of small and medium size European higher education institutions. The average number of links among the universities was 24.0 (standard deviation = 2,208.6). The universities with the greatest number of links (322,000) were between the Universität Trier and Rheinisch Westfalische Technische Hochschule Aachen, two German institutions that host very huge and popular bibliographic systems (DBLP and SunSite).

The ten most central universities based on five different measures (in-degree, out-degree, total degree, betweenness and eigenvector centrality) are listed in Table 2. The University of Wisconsin, Madison had the greatest number of links from other institutions (in-degree), 978. The average for all 1,000 universities was 606 (SD = 183.6). Of the ten most central institutions in terms of in-degree, eight were from the United States, and one each from the United Kingdom and Argentina. For outgoing links, the University of California, Berkeley was the most central with 959 links. The mean for all 1,000 universities was 606 (SD = 224.0). Again, eight of the ten most central schools were from the U.S. with one from the U.K. and one from Malaysia. When degree was totaled, the University of Wisconsin, Madison was the most central with 1,970 total links with other schools. The average was 1,212 (SD = 376.9). Nine of the ten were from the U.S. and one from the U.K.

Betweenness is the extent to which a node lies on the shortest path connecting others in the network or the proportion of all paths linking *j* and *k* that pass through *i*, summed for all nodes. Betweenness may be taken as a measure of the concept of structural hole (Burt 1992), such that a node with high betweenness centrality is considered to be the occupant of a structural hole. Further, nodes that serve as liaisons or bridges between groups have high betweenness relative to others in the network (Monge and Contractor 2003). It is considered as an identifier of gatekeeping potential, since a high betweenness value indicates that information flows are restricted to paths that run through those nodes. In this case, it may be considered an indicator of social or intellectual capital, and thus of a university’s academic influence (Leydesdorff 2007). Table 2 presents the ten universities with the highest betweenness. The University of Wisconsin, Madison was the most central institution by this measure. Nine of the ten universities on this list were from the U.S. Oxford was the only exception.

Table 2 University hyperlink network—ten most central institutions

In-degree	Links	Out-degree	Links
University of Wisconsin Madison	978	University of California Berkeley	959
Stanford University	966	Stanford University	966
University of Minnesota	965	Harvard University	940
Michigan State University	962	Massachusetts Institute of Technology	949
Universidad Nacional de la Plata	960	University of Oxford	917
University of California Berkeley	959	Universiti Kebangsaan Malaysia	927
University of Cambridge	958	University of Minnesota	965
Indiana University	957	University of Washington	956
University of Texas Health Science	956	University of Wisconsin Madison	978
University of Washington	956	Carnegie Mellon University	928
Mean number of links	606		606
Standard deviation	183.6		224

Total degree	Links	Betweenness
University of Wisconsin Madison	1970	University of Wisconsin Madison
Stanford University	1961	Massachusetts Institute of Technology
University of Minnesota	1958	University of California Berkeley
University of California Berkeley	1955	University of Washington
University of Washington	1949	Stanford University
University of Cambridge	1946	University of Minnesota
University of Texas Health Science	1946	University of Oxford
Massachusetts Institute of Technology	1943	Harvard University
Indiana University	1942	Carnegie Mellon University
Michigan State University	1941	Cornell University
Mean number of links	1212	
Standard deviation	376.9	

Eigenvector
University of California Berkeley
University of Cambridge
Harvard University
Massachusetts Institute of Technology
University of Oxford
Stanford University
University of Minnesota
University of Texas Health Science Center
University of Washington
Carnegie Mellon University

Eigenvector centrality is an indicator of a node's overall centrality in a network (Bonacich 1972). The measure takes into account the positions of a node's contacts such that it becomes more central if it is tied to more central nodes. Again, the ten most central nodes on this measure are presented in Table 2. They are all equivalent in eigenvector centrality. Eight of the ten are U.S. institutions (University of California Berkeley, Harvard University, Massachusetts Institute of Technology, Stanford University, University of Minnesota, University of Texas Health Science Center, University of Washington, Carnegie Mellon University), and two are from the U.K. (University of Cambridge and University of Oxford).

A hierarchical cluster analysis of the web-citation network using the average similarities between the members of the clusters resulted in ten groupings of universities. The average similarity method was chosen because the network consisted of such a large number of nodes. The single link method would not account for sufficient information about the relations among the nodes and the complete link method would consider too much. They are listed in Table 3. The clusters may be differentiated by a combination of many factors including the language of instruction spoken at the universities, how central they are in the network, the prestige of the universities and the home countries international bandwidth connections. Any single factor alone does not characterize cluster membership. Of special note is Cluster 2, which is composed of high prestige schools that are central in the network. The 82 highest ranked institutions in eigenvector centrality are all members of this grouping.

A description of the university hyperlink network: the national level

The connections among the universities were aggregated to the national level. The U.S. is at the center of the network with countries such as Peru, Indonesia and Iceland at the periphery. The network density was .924. Almost all countries had links to all others through their universities' URL-citations. Table 4 provides the centralities of the individual countries in terms of degree (in and out), share, eigenvector centrality and betweenness, as well as the number of its universities that are among the 1,000 in the network. The United States is by far the most central country in the university web-citation network. Over 30 % of the total links involve the U.S. Its universities had more than four million outward links to institutions in other countries and nearly 1.9 million links from other countries. From a web point of view English-speaking Canada can be regarded as part of the U.S. network, which would, in part, explain the higher values for U.S. indicators. This is indicative of a pull effect. Its eigenvector centrality is more than ten times that of the next most central country, Germany. The U.S. was followed by Germany, U.K.

Table 3 Groupings of universities from cluster analysis

Cluster	Defining attributes
1.	German, Swiss & Italian, not English, central, low prestige, less bandwidth connections
2.	English (U.S., Canada, U.K. Australia), central, high prestige, strong bandwidth connections
3.	Low prestige, peripheral, less bandwidth connections
4.	English, not French, peripheral, no Ph.D's, strong bandwidth connections
5.	Continental Europe, not English
6.	Chinese, less bandwidth connections
7.	French, not English, peripheral, lower prestige
8.	English, primarily (Jesuit Institutions), peripheral, low prestige
9.	English, peripheral
10.	Japanese & other Asian, peripheral, little bandwidth connections

Table 4 International university hyperlink network—national level centralities

	In-degree	Out-degree	Share	Eigenvector	Betweenness	Number of schools
ar	21102	9259	.002	.146	.151	3
at	168778	159697	.017	1.529	.218	9
au	362674	307335	.033	2.966	.196	28
be	100394	73180	.009	.704	.216	8
bg	3818	1941	.000	.024	.030	1
br	244604	49809	.018	1.617	.229	17
ca	566259	750292	.065	7.041	.229	38
ch	207058	256765	.022	1.693	.192	10
cl	15180	22998	.002	.172	.120	3
cn	154745	66426	.012	1.612	.155	11
co	30598	8774	.003	.186	.116	3
cr	7681	2932	.001	.040	.101	1
cz	76223	36266	.006	.626	.142	9
de	1133169	661893	.090	11.913	.229	67
dk	92609	128097	.010	1.446	.116	8
ec	24861	905	.002	.254	.008	1
ee	22481	5904	.002	.095	.054	2
es	403074	183454	.033	3.334	.229	43
fi	105029	75628	.009	.760	.185	9
fr	276260	153720	.022	2.176	.200	54
gr	68226	31383	.005	.458	.142	6
hk	123745	67766	.011	1.501	.229	7
hu	43471	19480	.003	.429	.077	4
id	32237	3794	.002	.231	.084	4
i.e.	49442	63028	.005	.420	.167	6
il	102568	116886	.009	1.223	.177	7
in	27872	9758	.002	.270	.100	3
ir	1921	1166	.000	.018	.007	1
is	31502	4172	.002	.331	.070	1
it	285905	150149	.022	2.119	.229	37
jm	3813	1249	.000	.020	.033	1
jp	364687	119799	.028	4.257	.196	34
kr	101865	22508	.008	1.270	.107	8
lt	10712	3415	.001	.042	.029	3
mx	78062	52887	.006	.764	.229	5
my	121984	4822	.009	.720	.078	5
nl	269643	170389	.023	2.207	.196	13
no	90290	56079	.007	.705	.192	5
nz	66636	52068	.007	.367	.110	7
pe	62380	2832	.005	.524	.071	2
ph	1114	791	.000	.008	.001	1
pl	75914	26665	.006	.699	.148	10

Table 4 continued

	In-degree	Out-degree	Share	Eigenvector	Betweenness	Number of schools
pr	100543	7262	.008	1.496	.026	1
pt	88109	78798	.008	.759	.200	8
ro	5871	2517	.000	.035	.020	2
ru	15974	4678	.001	.107	.064	3
sa	34702	40596	.004	.691	.069	4
se	163871	178382	.018	1.504	.167	12
sg	35935	85288	.007	.611	.177	2
si	30604	21384	.002	.310	.064	2
sk	12945	7634	.001	.062	.022	3
th	59837	13751	.005	.697	.145	11
tr	49828	13820	.004	.538	.147	6
tw	360038	48157	.027	4.446	.229	17
uk	654953	1153384	.085	8.139	.229	67
us	1897001	4023702	.306	140.140	.229	359
ve	27165	3318	.002	.236	.098	2
za	75035	23965	.006	1.156	.167	5
Mean	166258.563	166258.563	.017	3.756	.135	17.210
SD	298337.656	546907.938	.042	18.186	.072	65.440

and Canada on all measures. These four countries accounted for the majority (53.1 %) of the universities in the network. Overall, the network has a Gini-coefficient of .672 indicating a moderate core-periphery structure. Further, the number of a country's universities was distributed according to the power law ($R^2 = .935$, $F = 802.6$, $p < .000$, $a = 423.30$, $b = -1.36$), indicating that few countries accounted for the majority of links in the network. The hierarchical cluster analysis indicated that all the countries were part of a single group, centered about the U.S. and the U.K., with Canada and Western Europe near the center.

Predicting the structure of the university URL-citation network

The previous section described the structure of the university web-citation network at both the institutional and national level. This section describes the results of attempts to identify the determinants of the structure. The first two variables examined were the physical distances between universities and whether or not they were located in the same country. It was predicted that the closer the schools, the greater the number of university-to-university links. The correlation between web-citations and physical distance was essentially zero (.005). This coefficient was calculated without a number of French universities that had the same webdomain for multiple campuses located in different cities. Thus, physical distance did not determine the network's structure, suggesting that the Internet has made distance irrelevant for research collaboration.

Regarding the impact of political boundaries on the structure of the network, 78.4 % of bilateral relations were to universities outside the country, and 21.6 % of the relations were domestic. For those relations without a link, 93.9 % were international, and only 6.1 % were within the same country. Table 5 summarized these results in a contingency table.

Table 5 Domestic and international university hyperlinks

	Domestic	International	Total
Links	130,815	475,188	606,003
	13.1 %	47.5 %	60.6 %
No links	23,800	369,197	392,997
	2.4 %	37.0 %	39.4 %
Total	154,615	844,386	999,000
	15.5 %	84.5 %	100.0 %

The correlation between whether or not two universities are located in the same country and the number of web-citations was .065. While not large, this coefficient was calculated from 999,000 relationships, indicating that universities tend to have connections with others in the same country. Also, these relations tend to be stronger. The mean link strength (number of hyperlinks) to domestic universities was 1,415.0 versus only 42.5 to international institutions.

To predict a university's centrality in the network, correlation and linear regression analysis were employed. The preferred methods of regression analysis would be to use a type of count data model, such as Poisson or negative binomial regression due to skew in the number of web citations between two institutions (Seeber et al. 2012). However, because the analyzed network only considered the number of universities cited by an individual university, the number of ties is not skewed (See Table 6). Also the sample was further restricted to the 1,000 most interlinked universities, eliminating those extreme cases and limiting the variance and skew in the data. Thus, linear regression may be justified in this case. It should be noted that two independent variables, the number of noble prizes and the size of the universities were highly skewed. They were transformed by the natural log

Table 6 Means, standard deviations and skew of centrality measures and predictors

Variable	Mean	Std. deviation	Skew
In-degree	606	183.62	−.34
Out-degree	606	223.98	−.13
Total degree	1212.01	376.86	−.04
Betweenness	391	353.42	1.39
Eigenvector	.79	.16	−.78
Size	23676.75	31147.53	1.45[a]
Ph.D	.93	.26	−3.29
U.S. news	52.76	18.01	.77
Noble Prizes	.41	1.97	1.29[a]
Cluster 2	.28	.45	.97
English	.52	.5	−.06
German	.09	.28	2.96
Spanish	.06	.24	3.57
French	.08	.27	3.05
Chinese	.04	.19	4.91
Japanese	.03	.18	5.15
Italian	.04	.19	4.91
Bandwidth	9138409.59	6960797.67	−.10

[a] Value of skew is after transformation by natural log

for the analysis. Table 6 presents the variables' means, standard deviations and skews. Table 7 the correlations among the predictors and the various measures of centrality. In general, centrality in the network was predicted by the size of the university, whether or not it offered a doctoral degree, its international rating, the number of Nobel Prizes its faculty members have won, if "English and not French" was its language of instruction, and its nation's international bandwidth capacity. In terms of receiving citations from other universities, "German and not English" was significant. Bandwidth capacity was not a good predictor. However, for out-degree centrality, connecting to other universities, "all languages except English and German" were negatively related to centrality.

Because many of the antecedent predictors are interrelated, multiple regression analysis was performed to determine their independent and the combined effects on the universities' centrality in the network. The best predictive models are presented in Table 8. R^2 ranged from .31 to .58 depending on the measure of centrality. Across all the indicators, the log of the size of the university, its nation's international bandwidth capacity and its U.S. News' rating were significant predictors of a university centrality in the URL-citation network. English was not a significant predictor of in-degree. It was, however, significant for out-degree, betweenness and eigenvector centrality. Whether or not a university offered a doctoral degree was not significant because all prestigious universities offered doctorates.

Predicting the structure of the university hyperlink network: national level

A number of different networks were used as predictors of the structure of the university URL-citation network at the national level. These include international co-authorships (Leydesdorff and Wagner 2008), international citations according to the web of Science, international student exchanges, and the total hyperlinks among the individual countries. The QAP correlations indicate that all of these networks are significantly related to the university URL-citations network. Co-authorships from 2006 correlate with university citations .772 ($p < .000$), citations, $r = .967$ ($p < .000$), student flows $r = .270$ ($p < .002$), and total hyperlinks $r = .545$ ($p < .000$). Because each network was composed of a slightly different set of nations, the number of countries used to compare these networks was 52, with the exception of student flows, which were 48. Missing were China, Columbia, Ecuador, Jamaica, Israel, Singapore, South Africa, and Taiwan. These missing nodes may account for the low correlation between student flows and university mentions, as for example Asians account for 52 % of all international students (OECD 2012).

To predict a country's centrality in the national university URL-citations network both correlation and linear regression analysis were employed. While the citation data at the university level were not skew, the indicators of centrality and a number of their predictors were highly skewed. In-degree, out-degree, total degree, the number of Noble prizes, bandwidth, GDP, and population were transformed by the natural log to normalize the distributions for these variables. Eigenvector centrality was transformed by the natural log and its range truncated by assigning the United States the next greatest value, that of Germany.

Table 9 presents the means, standard deviations and skews, and Table 10 the correlations among the predictors and the various measures of centrality. For all measures, centrality in the network was predicted by the number of Nobel Prizes won by its citizens, the country's international Internet bandwidth capacity and its GDP. GDP/capita was a significant predictor of all centrality measures except eigenvector centrality. English significantly predicted out-degree and eigenvector centrality. The population of a country was unrelated to its centrality in the network.

Table 7 Correlations among centrality measures and predictors

	In-degree	Out-degree	Total degree	Between	Eigenvector	ln size	Ph.D	US news	ln Nobles	Cluster 2
In-degree	1									
Out-degree	.707*	1								
Total degree	.907*	.939*	1							
Betweenness	.841*	.839*	.909*	1						
Eigenvector	.894*	.892*	.966*	.795*	1					
ln size	.197*	.078	.143*	.183*	.122*	1				
Ph.D	.176*	.106*	.149*	.171*	.144*	.133*	1			
US news	.477*	.447*	.501*	.580*	.422*	−.051	.075	1		
ln Noble Prizes	.263*	.277*	.293*	.426*	.223*	−.014	.054	.519*	1	
Cluster 2	.418*	.497*	.499*	.514*	.449*	.074	.091*	.370*	.195*	1
English	.015	.424*	.259*	.224*	.242*	−.078	−.210*	.243*	.098*	.275*
German	.113*	.069	.096*	.051	.087*	−.064	.045	−.104	.009	−.114*
Spanish	.016	−.169*	−.092*	−.086*	−.073	.177*	.058	−.134	−.054	−.046
French	−.170*	−.125*	−.157*	−.107*	−.200*	−.038	.070	−.027	−.019	−.115*
Chinese	.062	−.143*	−.054	−.035	−.018	−.004	.055	.106	−.038	−.029
Japanese	−.064	−.157*	−.124*	−.072	−.128*	−.024	.053	.078	−.014	−.02
Italian	.017	−.067	−.032	−.055	−.025	.105*	.035	−.131	−.038	−.076
Bandwidth	−.006	.322*	.188*	.163*	.142*	−.131*	−.252*	.146*	.115*	.166*

	English	German	Spanish	French	Chinese	Japanese	Italian	Bandwidth
English	1							
German	−.317*	1						
Spanish	−.270*	−.080	1					
French	−.309*	.038	−.078	1				
Chinese	−.107*	−.060	−.051	−.059	1			

Table 7 continued

	English	German	Spanish	French	Chinese	Japanese	Italian	Bandwidth
Japanese	−.194*	−.058	−.049	−.056	−.037	1		
Italian	−.202*	−.060	−.051	−.059	−.038	−.037	1	
Bandwidth	.616*	.135*	−.252*	−.088*	−.214*	−.176*	−.188*	1

* Correlation is significant

Table 8 Multiple regression—predicting network centrality

	In-degree			Out-degree			Betweenness			Eigenvector		
R^2	.350			.489			.579			.310		
F	47.94			85.16			122.25			39.94		
p	.000			.000			.000			.000		
	β	t	p	β	t	p	β	t	p	β	t	p
Size (log)	.279	6.49	.000	.123	3.22	.001	.282	8.13	.000	.150	3.36	.001
English	−.025	−.516	.606	.356	8.50	.000	.185	4.86	.000	.214	4.40	.000
Bandwidth	.268	5.70	.000	.302	7.31	.000	.336	8.94	.000	.208	4.33	.000
Rating	.465	10.53	.000	.323	8.25	.000	.502	14.12	.000	.348	7.65	.000

Table 9 Means and standard deviations of centrality measures and predictors—national level

	Mean	Std. deviation	Skew	N
Ln In-degree	11.05	1.51	−.37	58
Ln Out-degree	10.25	1.9	.129	58
Ln Total degree	11.20	1.57	−.23	58
Ln Eigenvector	.11	.02	3.21	58
Betweenness	4.31	2.33	−.27	58
Ln Nobles	.84	1.28	1.76	58
Ln bandwidth	13.76	2.30	3.03	53
Ln GDP	5.81	1.46	.02	56
Ln Pop	3.18	1.45	.59	56
GDP/cap	23.42	18.83	.84	56
English	.19	.4	1.63	58

Because many of the antecedent predictors are interrelated (See Table 10), multiple regression analysis was performed to determine their independent and combined effects on the nations' centrality in the network. The best predictive models are presented in Table 11. In-degree centrality is positively predicted by a country's GDP per capita and its population ($R^2 = .52$). This suggests that economically well off countries with large populations, such as the United States and Germany have a large number of incoming links to their universities. Out-degree centrality is positively predicted by, a country's GDP per capita, its population and whether English is its language of instruction ($R^2 = .67$). Prestigious universities from wealthy English-speaking countries with large populations such as the United States and United Kingdom tend to have a large number of outward citations. Betweenness is predicted by GDP per capita and a country's population ($R^2 = .39$). Connected countries with economically well off populations such as the United States, United Kingdom, Germany, Spain and Italy were high in betweenness. Finally, the number of Nobel Prizes, English and GDP predict eigenvector centrality ($R^2 = .64$). Countries' overall centralities in the international university URL-citations network tend to have prestigious universities as indicated by the number of Noble prize winners, a large GDP and tend to use English. Examples of this would include the Germany, Japan, the U.K., Canada and the United States.

Table 10 Correlations among centrality measures and predictors—national level

	In-degree	Out-degree	Total degree	Eigenvector	Betweenness	Nobles
In-degree	1					
Out-degree	.88**	1				
Total degree	.99**	.91**	1			
Eigenvector	.71**	.70**	.73**	1		
Betweenness	.82**	.86**	.83**	.53**	1	
Nobles	.62**	.72**	.65**	.75**	.51**	1
Bandwidth	.51**	.52**	.52**	.42**	.40**	.48**
GDP	.69**	.68**	.69**	.63**	.66**	.63**
Population	.24	.14	.22	.33*	.21	.24
GDP/cap	.37**	.30*	.32*	.21	.57**	.26
English	.18	.31*	.23	.32*	.23	.19

	Bandwidth	GDP	Population	GDP/cap	English
In-degree					
Out-degree					
Share					
Eigenvector					
Betweenness					
Nobles					
Bandwidth	1				
GDP	.61**	1			
Population	.32*	.77**	1		
GDP/cap	.36**	.22	−.24	1	
English	.08	.12	. 03	.16	1

In-degree, out-degree, total degree, noble prizes, bandwidth, GDP, and population have been transformed by the natural log. Eigenvector centrality has been transformed by the natural log and its range truncated

* Correlation is significant, $p < .05$ level (2-tailed)

** Correlation is significant, $p < .01$ level (2-tailed). $N = 56$

Discussion

The results of this research indicate that the global university URL-citations network is clustered into ten groupings that may be differentiated by the language of instruction spoken at the universities, how central they are in the network, their prestige and the home countries international bandwidth connections. At the center of this network are prestigious institutions primarily from the United States and the United Kingdom. The overall structure of the network could not be predicted by the physical distance between the schools, suggesting that the Internet is making distance irrelevant for research collaboration. Universities tended to link to others from the same country. In general, centrality in the network was predicted by the size of the university, whether it offered a doctoral degree, its U.S. News rating, the number of noble prizes its faculty have won, its language of instruction, and its nation's international bandwidth capacity. At the national level, the countries formed a single group centered about the U.S. and the U.K. with almost all countries linked together. The United States is by far the most central country followed by

Table 11 Multiple regression—predicting network centrality—national level

	In-degree[a]			Out-degree[a]			Betweenness			Eigenvector[b]		
R^2	.524			.670			.505			.642		
F	33.78			35.12			22.99			31.05		
p	.000			.000			.000			.000		
	β	t	p	β	t	p	β	t	p	β	t	p
Nobles[a]										.553	5.07	.000
English				.184	2.27	.028				.183	2.15	.000
Population[a]	.482	.4.80	.000	.398	4.70	.000	.443	4.33	.000			
GDP/capital	.722	.7.19	.000	.797	9.28	.000	.720	7.03	.000			
GDP[a]										.258	2.41	.020

[a] In-degree, out-degree, noble prizes and population have been transformed by the natural log

[b] Eigenvector centrality has been transformed by the natural log and the range truncated

Germany, U.K. and Canada. Together, they accounted for the majority of the universities in the network. Overall, the international network had a core-periphery structure with a few countries accounting for the majority of links in the network. International co-authorships, citations, international student exchanges and the number of links among the individual countries are strongly predictive of the structure of the network. The regression analysis, indicated that centrality in the international network was predicted, by a country's population and its GDP. Depending on the measure it may also be predicted by whether or not English was its language of instruction for higher education and the number of noble prizes won by its citizens.

Although these results are based on URL-citations rather than the hyperlinks among universities, they are consistent with the Seeber et al. (2012) who found that the European university hyperlink network displayed a center-periphery structure with centrality a function of the universities' reputation. This study extends their conclusions to the global academic community. The reported findings are also consistent with Ortega and Aguillo (2009, p. 278) who conclude that:

> 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.

The results reveal a center-periphery structure of the network consistent with World System theory. At the national level, there is a single cluster centered about the United States, United Kingdom and Germany. There are strong correlations between the various measures centrality and population, GDP and GPD/capita. Thus, the international university URL-citation network is similar to other global information networks with core-periphery structures that have been characterized as supportive of World System Theory. These include the telephone (Barnett 2001), the Internet (Barnett and Park 2005), student flows (Barnett and Wu 1995; Chen and Barnett 2000), and patents, trademarks and copyrights (Nam and Barnett 2011).

While the results reported here suggest that the URL-citation network is similar in structure to other academic and international networks and provides a useful indicator of the relations among these institutions, the overall results are somewhat erratic accounting for anywhere from 30 to 56 % of the variance at the university level and 51–67 % of the variance at the national level depending on the network indicator. Thus, future research is required to more precisely determine what is signified by URL-citations. Are they simply textual references or do they represent formal research collaborations, or something else entirely?

This study analyzed the global university URL-citation network as a multiple level system, at the level of single universities and at the nation-state level. Social networks exist simultaneously at several levels, from the individual component and its links that form groups, which are also tied together to create the network as a whole (Barnett and Park 2012). The solitary entities are embedded in superordinate systems composed of components of the same taxonomic category. This is known as hierarchicalization. For example, academic departments are the subordinate components to schools or colleges that compose universities. These institutions together are organized into a national system of higher education, which in turn creates the global academic community. Thus, the networks that tie together the components of this system can be examined at multiple levels as was done in this case.

The Internet has often been described as an ever evolving and self-organizing system (Barnett et al. 2001; Barnett 2005; Chakrabarti et al. 1999; Shumate and Lipp 2008; Weber and Monge 2011). This implies that the international university URL-citation network may also be described as a self-organizing system. The academic network may be considered an autopoietic or self-replicated system that evolved from traditional scientific activities including co-authorship, the practice of citing the research of others and numerous other behaviors that required the sharing of information among scholars (Manturana and Varela 1980). An autopoietic system can be defined as “a network of processes that produces all the components necessary to embody the very process that produces it” (Krippendorf 1991, p. 138). The network recursively produces all its components through the interaction in this historical reproductive network of postings on university websites and the links from one institution to another.

It should be noted that there are a number of environmental constraints that limit the possible states into which this system (network) may evolve. These limits are due issues of information property, the policies of individual universities and national governments, as well as the policies of scientific funding agencies like the U.S. or European National Science Foundations. Further, academic networks have co-evolved along with other global institutions. Universally, higher education is developing a common curriculum especially in the science (Lechner and Boli 2005). This may be reflected in pattern of universities’ hyperlinks and web-citations.

Future research is planned to examine how the network is changing over time (Park and Leydesdorff 2010; Kwon et al. 2012). Are the international university networks becoming denser as a function of the globalization of higher education and the increased scientific productivity of countries from outside the core (Leydesdorff and Wagner 2008)? What are the effects of the improvements in information technology and the strengthening of global communication ties on this network (Park et al. 2011a, b)?

Acknowledgments The authors would like to thank Loet Leydesdorff, Jessie Hammond, Haley Pleasant, and Jeanette Ruiz without whom this paper would not have been possible. The corresponding author acknowledges a partial support from the SSK (Social Science Korea) Program funded by National Research Foundation of South Korea; NRF-2010-330-B00232.

References

Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: Methodological approaches to 'Webometrics'. *Journal of Documentation, 53*(4), 404–426.

Barnett, G. A. (2001). A longitudinal analysis of the international telecommunications network: 1978–1996. *American Behavioral Scientist, 44*(10), 1638–1655.

Barnett, G. A. (2005). An Introduction to self-organizing systems. In G. A. Barnett & R. Houston (Eds.), *Advances in self-organizing systems* (pp. 3–32). Cresskill: Hampton Press.

Barnett, G. A. (2011). Communication and the evolution of SNS: Cultural convergence perspective. *Journal of Contemporary Eastern Asia, 10*(1), 43–54.

Barnett, G. A., & Choi, Y. (1995). Physical distance and language as determinants of the international telecommunication network. *International Political Science Review, 16*, 249–265.

Barnett, G. A., Chon, B. S., Park, H. W., & Rosen, D. (2001). An examination of international Internet flows: An autopoietic model. Paper presented to the International Communication Association, Washington.

Barnett, G. A., Jacobson, T. L., Choi, Y., & Sun-Miller, S. L. (1996). An Examination of the international telecommunication network. *The Journal of International Communication, 3*, 19–43.

Barnett, G. A., & Park, H. W. (2005). The structure of international internet hyperlinks and bilateral bandwidth. *Annals of Telecommunication, 60*, 1115–1132.

Barnett, G. A., & Park, H. W. (2012). Examining the international internet using multiple measures: New methods for measuring the communication base of globalized cyberspace. *Quality and Quantity*. doi: 10.1007/s11135-012-9787-z.

Barnett, G. A., & Salisbury, J. G. T. (1996). Communication and globalization: A longitudinal analysis of the international telecommunication network. *Journal of World-Systems Research, 2*(16), 1–17.

Barnett, G. A., Salisbury, J. G. T., Kim, C., & Langhorne, A. (1999). Globalization and international communication networks: An examination of monetary, telecommunications, and trade networks. *The Journal of International Communication, 6*, 7–49.

Barnett, G. A., & Sung, E. (2006). Culture and the structure of the international hyperlink network. *Journal of Computer-Mediated Communication, 11*(2), 17–38.

Barnett, G. A., & Wu, Y. (1995). The international student exchange network: 1970 and 1989. *Higher Education, 30*, 353–368.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology, 2*(1), 113–120.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). UNICET 6 for Windows: Software for social network analysis [Computer Program]. Needham, MA: Analytic Technologies.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge: Harvard University Press.

Chakrabarti, S., Dom, B. E., Kumar, S. R., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Gibson, D., & Kleinberg, J. (1999). Mining the Web's link structure. *Computer*, *32*(8), 60–67.

Chase-Dunn, C., & Grimes, P. (1995). World-systems analysis. *Annual Review of Sociology, 21*, 387–417.

Chen, T. M., & Barnett, G. A. (2000). Research on international student flows from a macro perspective: A network analysis of 1985, 1989 and 1995. *Higher Education, 39*, 435–453.

Choi, S., & Park, H. W. (2014, forthcoming). An exploratory approach to a Twitter-based community centered on a political goal in South Korea: Who organized it, what they shared, and how they acted. *New Media & Society*.

Chung, C. J., & Park, H. W. (2012). Web visibility of scholars in media and communication journals. *Scientometrics, 93*(1), 207–215.

Chung, C. J., Cho, S. E., & Park, H. W. (2013). SNS use by the Korean Government: A case of Me2Day. *Asian Journal of Communication*, *23*(6).

Chung, C. J., Barnett, G. A., & Park, H. W. (2013). Inferring international dotcom web communities by link and content analysis. *Quality & Quantity*. doi:10.1007/s11135-013-9847-z.

College Board. (2012). *College handbook 2012*. New York: College Board.

De Maeyer, J. (2012). Towards a hyperlinked society: A critical review of link studies. *New Media & Society*. doi:10.1177/1461444812462851.

Doz, Y. L., Oik, P. M., & Ring, P. S. (2000). Formation processes of R&D consortia: Which path to take? Where does it lead? *Strategic Management Journal, 21*, 239–266.
EUMIDA (2010) European University data collection. Retrieved July 12, 2012 from http://thedatahub.org/dataset/eumida.
Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics, 51*, 69–115.
Heimeriks, G., & Van Den Besselaar, P. (2006). Analyzing hyperlinks networks: The meaning of hyperlink based indicators of knowledge production. *Cybermetrics*, 10(1). http://cybermetrics.cindoc.csic.es/articles/v10i1p1.html. Accessed 21 June 2013.
Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation, 54*, 236–243.
Khan, G. F., Yoon, H. Y., & Park, H. W. (2013). Social media communication strategies of government agencies: Twitter use in Korea and the U.S. *Asian Journal of Communication, 23*(6)
Khan, G. F., Yoon, H. Y., Kim, J. Y., & Park, H. W. (2014). From E-Government to social government: The case of twitter use by Korea's Central Government. *Online Information Review, 38*(1).
Krackhardt, D. (1987). QAP partialling as a test for spuriousness. *Social Networks, 9*, 171–186.
Krippendorf, K. (1991). Society as self-referential. *Journal of Communication, 41*, 136–140.
Kwon, K.-S., Park, H. W., So, M. H., & Leydesdorff, L. (2012). Has globalization strengthened South Korea's national research system? National and international dynamics of the triple helix of scientific co-authorship relationships in South Korea. *Scientometrics, 90*(1), 163–176.
Lechner, F. J., & Boli, J. (2005). *World culture: Origins and consequences*. Oxford: Blackwell.
Lee, M., & Park, H. W. (2012). Exploring the web visibility of world-class universities. *Scientometrics, 90*, 201–218.
Leydesdorff, L. (2007). ''Betweenness centrality'' as an indicator of the ''interdisciplinarity'' of scientific journals. *Journal of the American Society for Information Science and Technology, 58*(9), 1303–1309.
Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics, 2*, 317–325.
Leydesdorff, L., Wagner, C., Park, H. W., & Adams, J. (2013). International collaboration in science: The global map and the network. *El professional de la Información, 22*(1), 87–94.
Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003). The relationship between the links/web impact factors of computer science departments in UK and their RAE (Research Assessment Exercise) ranking in 2001. *Scientometrics, 57*(2), 239–255.
Lim, Y. S., & Park, H. W. (2013). The structural relationship between politicians' web visibility and political finance networks: A case study of South Korea's national assembly members. *New Media & Society, 15*(1), 93–108.
Manturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel.
Handcock, M. S., Hunter, D. R, Butts, C. T, Goodreau, S. M., & Morris M. (2003). statnet: Software tools for the statistical modeling of network data. http://statnetproject.org. Accessed 21 June 2013.
Monge, P. R., & Contractor, N. S. (2003). *Theories of communication networks*. Oxford: Oxford University Press.
Moon, S., Barnett, G. A., & Lim, Y. S. (2010). The structure of International music flows using network analysis. *New Media and Society, 12*(3), 379–399.
Nam, Y., & Barnett, G. A. (2011). Globalization of technology: Network analysis of global patents and trademarks. *Technological Forecasting and Social Change, 78*(8), 1471–1485.
Nam, Y., Lee, Y.-O., & Park, H. W. (2013). Can web ecology provide a clearer understanding of people's information behavior during election campaigns? *Social Science Information, 52*(1), 91–109.
OECD (2012). Education at a glance 2012: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en. Accessed 21 June 2013.
Opsahl, T., Colizza, V., Panzarasa, P., & Ramasco, J. J. (2008). Prominence and control: The weighted rich-club effect. *Physical Review Letters, 101*, 168702.
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. (2008). Linking patterns in the European Union's Countries: Geographical maps of the European academic web space. *Journal of Information Science, 34*(5), 705–714.
Ortega, J. L., & Aguillo, I. F. (2009). Mapping world-class universities on the web. *Information Processing and Management, 45*, 272–279.
Park, H. W. (2010). Mapping the e-science landscape in South Korea using the webometrics method. *Journal of Computer-Mediated Communication, 15*(2), 211–229.
Park, H. W., Barnett, G. A., & Chung, C. J. (2011a). Structural changes in the 2003–2009 global hyperlink network. *Global Networks, 11*(4), 522–542.

Park, H. W., & Leydesdorff, L. (2010). Longitudinal trends in networks of university-industry-government relations in South Korea: The role of programmatic incentives. *Research Policy, 39*(5), 640–649.
Park, S. J., Lim, Y. S., Sams, S., Sang, M. N., & Park, H. W. (2011b). Networked politics on cyworld: The text and sentiment of Korean political profiles. *Social Science Computer Review, 29*(3), 288–299.
Park, H. W., & Thelwall, M. (2006). Web science communication in the age of globalization. *New Media & Society, 8*(4), 629–650.
Payne, N., & Thelwall, M. (2004). A statistical analysis of UK academic web links. *Cybermetrics, 8*(1), 2.
Rivera, M. T., Soderstrom, S. B., & Uzzi, B. (2010). Dynamics of dyads in social networks: Assortative relational, and proximity mechanisms. *Annual Review of Sociology, 36*, 11–91.
Sams, S., & Park, H. W. (2013). The presence of hyperlinks on social network sites: A case study of Cyworld in Korea. *Journal of Computer-Mediated Communication, 19*(1).
Seeber, M., Lepori, B., Lomi, A., Aguillo, I., & Barberio, V. (2012). Factors affecting web links between European higher education institutions. *Journal of Informetrics, 6*, 435–447.
Shumate, M., & Lipp, J. (2008). Connective collective action online: An examination of the hyperlink network structure of an NGO issue network. *Journal of Computer-Mediated Communication, 14*, 178–201.
Tang, R., & Thelwall, M. (2004). Patterns of national and international web inlinks to US academic departments: An analysis of disciplinary variations. *Scientometrics, 60*(3), 475–485.
Tehranian, M. (1990). *Technologies of power: Information machines and democratic prospects*. Norwood: Ablex.
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., & Sud, P. (2011). A comparison of methods for collecting web citation data for academic organisations. *Journal of the American Society for Information Science and Technology, 62*(8), 1488–1497.
Thelwall, M., Sud, P., & Wilkinson, D. (2012). Link and co-inlink network diagrams with URL citations or title mentions. *Journal of the American Society for Information Science and Technology, 63*(4), 805–816.
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-Centred European union link analysis. *Scientometrics, 75*(3), 407–420.
Vaughan, L., & Thelwall, M. (2005). A modeling approach to uncover hyperlink patterns: The case of Canadian universities. *Information Processing and Management, 41*(2), 347–359.
Wagner, C. S. (2008). *The new invisible college*. Washington, DC: Brookings Press.
Wallerstein, I. (1974). *The modern world system*. New York: Academic Press.
Weber, M., & Monge, P. R. (2011). Network evolution. In G. A. Barnett (Ed.), *Encyclopedia of social networks* (pp. 600–606). Los Angeles: Sage.