Qual Quant (2014) 48:563–575
DOI 10.1007/s11135-012-9787-z

Examining the international internet using multiple measures: new methods for measuring the communication base of globalized cyberspace

George A. Barnett · Han Woo Park

Published online: 29 September 2012

Abstract This article examines the network structure of the international internet using four different sources of data: (1) bilateral bandwidth between countries; (2) hyperlink connections among nations' domain names; (3) structural equivalence of nations from the perspective of websites, measured by the percentage of specific websites' traffic from individual countries; and (4) structural equivalence of nations from the national perspective, using the proportion of a country's 100 most-visited websites shared with other countries. Results indicate that the international internet network appears to consist of series of small worlds determined by language, geography, and historical circumstances. Therefore, one cannot depict the internet only through an examination of the hyperlink connections among nations. There is a need for multiple indicators to accurately describe the global internet.

Keywords Multiple measures · International internet structure · Bandwidth · Hyperlink network · National website traffic

1 Introduction

Social network analysis has recently grown because it enables an unobtrusive approach to deal with the relational connections of many interacting units at micro, meso, and micro levels (D'Esposito and Zaccarin 2011). The network perspective can be applied to discover relational processes among individuals, organizations, and nation-states. However, few studies have examined the international internet and its enabling Web-services structure. One reason is that the internet is a packet-switched network that, unlike a telephone network, does not devote a single circuit to each individual message. Packets may take multiple paths from

G. A. Barnett
Department of Communication, University of California, Davis, CA, USA

H. W. Park (✉)
Department of Media and Communication, Cyber Emotions Research Institute,
Yeungnam University, Gyeongsan, South Korea
e-mail: hanpark@ynu.ac.kr

origin to destination. Thus, the Web has no "engineered architecture," and, consequently, the origin and destination of individual messages cannot be determined (Barnett and Park 2005), and the flows of information between any two points can be tricky (Dodge and Zook 2009). Therefore, there is no "ideal" description of the network's structure. Still, the internet represents a self-organized system with a well-defined structure of linkage that implies an underlying social structure (Chakrabarti et al. 1999; Shumate and Lipp 2008). Thus, multiple indicators of the network and its structure are required to identify the hidden structure of internet-mediated communication among nations.

Past research examining international internet traffic has primarily focused on single measure, in other words, inter-domain hyperlinks (Barnett et al. 2001). A hyperlink is the technological capability that enables a website to link seamlessly with another, generally through a click of a mouse (Park et al. 2004). The World Wide Web may be defined as a distributed hypertext system consisting of a virtual network of content and hyperlinks, with billions of inter-linked pages (Kleinberg and Lawrence 2001). The research describing the structure of hyperlink connections among the nations of the world has been reviewed in detail (Barnett et al. 2011; Rosen et al. 2011; Park et al. 2011a). However, as Weber and Monge point out, the measure used in prior studies can be problematic since linking has a very low cost barrier (Weber and Monge 2011). As a result, hyperlink analysis may contain erroneous links, irrelevant information, and inconsequential relationships.

As an alternative, this paper suggests examining the structure of online relationships through the context or content of relationships among social systems. It examines the structure of the international internet using four different sources of data: (1) bandwidth capacity and the physical infrastructure between countries; (2) hyperlink connections among nations' domain names; (3) structural equivalence of nations from the perspective of websites, measured by the percentage of specific website's traffic from individual countries; and, (4) structural equivalence from the national perspective, using the ranking of website use by individual countries. As such, it addresses the inadequacies of hyperlink analysis. The use of multiple measures provides a more robust description of the network structure of the international internet, which makes possible the study of its evolution and better predictions of the role of the internet in global communication. Data from 2010 and 2011 are used to describe the relations among the four perspectives. Thus, the paper offers both an important methodological improvement as well as an empirical contribution that may prevent scholars and policy makers from drawing possibly misguided conclusions. Therefore, one can easily follow our recommendation and expand on the research that examines the international internet through a network analysis of hyperlinks among nations.

2 Data

2.1 Hyperlinks

Hyperlink connectivity data were collected through the Yahoo.com using LexiURL Searcher, a Web analysis tool developed by Thelwall, available at http://lexiurl.wlv.ac.uk (Thelwall 2009). The data were collected in November 2010. Google indexed between 42 and 52 billion websites in November 2011 (http://www.worldwidewebsize.com/, 2012, visited February 1). Yahoo indexed a maximum of 15 billion websites from 273 TLDs. This study examined over 14.3 billion hyperlinks.

For the United States, four TLDs reserved for the US institutions (.edu, .gov, .mil, and .us) were combined. Since .com, .org, and .net are not exclusive to the United States, they were not included as US websites. However, the most-widely used sites from these TLDs were decomposed, and additional links were added to the individual countries totals using the procedures described by Barnett et al. (2011).

Yahoo was employed because it was the only search engine that traced incoming and outgoing links between websites (Thelwall 2009). Also, there was no option in Google for retrieving bidirectional links between a pair of websites, although Google could run incoming link queries. However, both search engines do not properly index Web 2.0 (e.g., blogs, social networking sites, P2P sites) (Hsu and Park 2011; Park et al. 2011b). Thus, "dark" or "hidden" internet traffic was not included (http://www.ncess.ac.uk/events/item/?item=207, 2012, visited February 1). Moreover, there may be potential bias in the coverage of certain countries because of the unstructured nature of the Web across national borders. Although not providing comprehensive and unbiased coverage, search engines do return useful information. Search engines should be viewed as engineering products, not as mathematical tools (Thelwall 2009; Park 2012). While search engines collect data from the Web, the researcher must draw appropriate conclusions from the gathered information.

2.2 Bandwidth capacity

Data on the bandwidth capacity among nations was gathered from TeleGeography—Global Internet Geography (TeleGeography 2011). It reports bilateral internet bandwidth in megabits per second (MBPS). Data for 75 countries for July 2011 were analyzed.

2.3 Shared website use: website perspective

Data on shared website use were mined from *Alexa.com's* website during December 2011 (http://www.alexa.com/topsites, 2012, visited January 17). The 1,000 most-frequently visited websites over the prior 3-month period by global internet users were collected along with the information regarding the percentage of these users from individual countries. The most-frequently visited website was *google.com* with 48.93 % of internet users worldwide visiting the site daily. The thousandth site was *usmagazine.com* with 0.11 %. The ranks of websites by site visits were distributed by a power function ($R^2 = .958$, $p < .000$). The number of countries reporting using a particular website varied from forty-eight for *netlog.com* to one for *tataindicom.com*, which was only visited from Indian users. The average was 18.51 countries (SD = 13.74). The data collection resulted in a two-mode sociomatrix (country [249] by website [1,000]), where each cell s_{ij} was the percentage of use of website i attributable to nation j. This matrix was premultiplied by its transpose to create a symmetrical country (249) by country sociomatrix in which each cell indicates the total percentages of website use shared by a pair of countries. A network analysis of these data describes the structural equivalence of nations from the perspective of shared websites.

2.4 Shared website use: national perspective

Data on shared website use from a national perspective were also mined from *Alexa.com* during December 2011. The ranks of the one hundred most-frequently visited websites for each of 160 countries were collected producing a two-mode sociomatrix (country [160] by website [100]), where each cell s_{ij} was the rank of website i for nation j. The ranks were dichotomized, where one indicated that a website was among the hundred most-frequently visited by a country, and zero indicated that it was not. This matrix was also premultiplied by its transpose to create a symmetrical country (160) by country sociomatrix in which each cell indicates the number of a country's 100 most frequently visited websites shared by each other country. A network analysis of these data describes the structural equivalence of countries from the perspective of the nations through the creation of an affiliation matrix of countries.

The rank orders of the one hundred most-frequently visited websites were somewhat volatile. These data were collected twice, once in December 2011, and again throughout the first quarter of 2012. The two dataset correlated .872 ($p < .000$). The rank orders were least stable for Nicaragua, Oman, Saudi Arabia, Senegal, Thailand, Uganda, and Uzbekistan, countries toward the periphery of these networks.

These four data sets were analyzed and compared, using UCINET, the most popular network analysis software, and graphically displayed, using NetDraw (Borgatti et al. 2002; Borgatti 2005).

3 Results

3.1 International hyperlink network

The 2010 international hyperlink network has a density of .657, that is, 65.7 % of possible links among the ccTLDs exists. Japan (.jp) is the most central country with over 227 million inward and 79 million outward links: 9.0 % of the total international hyperlinks. The .uk had over 264 million inward and 100 million outward links: 9.0 % of the total international links. They are followed by Germany (.de), France (.fr), and .es (Spain). The Gini-coefficient for the distribution of international links is 0.996. Hyperlinks in this network are very concentrated in a few countries.

Table 1 presents 75 countries' centralities after the 120 most-widely used websites for the gTLDs (.com, .org, .net) have been decomposed and allocated to the ccTLDs based on the procedures described by (Barnett et al. 2011). Each of these websites was used by at least 0.5 % of the internet users worldwide on a daily basis for the last three months (www.Alexa.com). This smaller network is totally interconnected. Its density is 1.0. The distribution of hyperlinks is more equitable for these nodes. The Gini-coefficient was .762. As expected, the United States is the most central node in this network, with 13.2 % of the hyperlinks, 500 million inward and 751 million outward links. Japan is the next most central country with 10.9 %, 636 million inward and 382 million outward hyperlinks; followed by China, the United Kingdom, and Germany. These results are graphically displayed in Fig. 1.

Figure 1 shows that the network has regional groupings as determined through hierarchical cluster analysis. They include groups for Latin America, the former Soviet republics, Scandinavia, and the Middle East. These regional clusters are connected to the core through more central countries from those regions. For example, Mexico, Argentina, and Brazil connect the other Latin American countries to the most central nations (.jp, .uk, .de, .cn, .usa); Russia and certain eastern nations link the former Soviet republics (.ee, .kz, .by .az) to the core; and China connects Taiwan and Hong Kong to the more central countries.

3.2 Bandwidth capacity

The density of the internet bandwidth capacity is .146. The United States in the most central country in this network, with a capacity of over 17.2 million MBPS, or 14.7 % of the world's share. It is followed by Germany, 15.0 million (12.8 %); the United Kingdom, 11.5 million (9.8 %); and France, 10 million (8.5 %). The distribution of bandwidth links is also concentrated in a few countries. The Gini-coefficient is .775. Perhaps the most interesting

Table 1 Eigenvector centralities

Bandwidth hyperlink web use Website country				
Argentina	3.69	5.52	6.14	19.57
Australia	9.38	16.40	15.71	19.34
Austria	6.69	14.20	3.86	17.34
Bangladesh	0.02	0.01	3.44	17.59
Belarus	0.33	0.69	0.87	11.29
Belgium	6.88	9.24	5.00	18.70
Bolivia	0.14	0.15	0.02	17.15
Brazil	9.86	12.00	17.94	17.23
Bulgaria	1.54	0.71	0.31	15.94
Canada	16.69	30.03	33.65	18.69
Chile	2.78	1.68	1.53	17.61
China	9.86	25.34	26.18	4.61
Colombia	3.22	1.65	3.09	18.27
Czech Rep.	4.55	3.97	1.34	12.96
Denmark	9.12	5.42	1.52	17.47
Ecuador	0.30	0.37	0.20	18.47
Egypt	1.01	0.16	4.20	15.33
Estonia	0.21	0.27	0.02	13.01
Finland	1.30	2.74	1.76	17.47
France	58.80	43.47	20.54	15.44
Germany	63.35	57.31	27.38	15.43
Greece	1.85	2.44	3.79	16.34
Guatemala	0.07	0.10	0.14	18.79
Hong Kong	2.69	5.84	2.40	14.25
Hungary	3.83	2.64	0.93	14.02
India	2.86	0.12	59.84	19.03
Indonesia	0.23	0.52	13.60	18.42
Ireland	3.66	9.63	2.80	7.89
Israel	1.00	0.76	1.37	18.95
Italy	13.99	23.86	20.91	17.04
Japan	10.47	54.72	25.96	12.00
Kazakhstan	0.30	0.61	1.31	10.20
Korea	2.35	7.27	15.33	13.35
Latvia	0.14	0.36	0.13	13.46
Lithuania	0.34	0.43	0.12	12.28
Luxembourg	1.06	1.06	0.17	16.84
Macedonia	0.01	0.11	0.01	14.47
Malaysia	1.01	2.10	6.25	18.66
Mexico	6.38	5.33	21.16	19.61

Table 1 continued

Bandwidth hyperlink web use Website country				
Moldova	0.36	0.08	0.01	13.78
Morocco	0.48	0.15	0.84	15.81
Netherlands	53.09	12.73	9.05	17.98
New Zealand	0.41	3.86	0.63	17.78
Norway	3.81	4.14	2.57	17.89
Oman	0.07	0.02	0.21	14.44
Pakistan	0.48	0.16	12.78	20.95
Panama	1.24	0.30	0.02	18.42
Peru	1.80	0.71	2.34	19.08
Philippines	1.13	1.55	6.82	19.09
Poland	8.05	11.24	6.41	10.66
Portugal	2.48	3.73	3.97	18.55
Qatar	0.15	0.01	0.31	17.37
Romania	3.16	2.00	4.79	17.52
Russia	6.82	6.41	17.75	10.91
Saudi Arabia	0.83	0.09	3.35	13.86
Senegal	0.04	0.05	0.02	13.26
Serbia	0.58	0.41	0.24	16.55
Singapore	3.33	3.19	4.03	18.01
Slovakia	3.26	1.86	0.15	14.51
Slovenia	0.44	0.70	0.09	16.81
South Africa	0.21	1.71	6.36	18.04
Spain	21.48	27.47	18.33	18.26
Sri Lanka	0.07	0.02	0.65	18.87
Sweden	10.26	8.33	5.29	17.91
Switzerland	6.72	11.58	2.96	18.51
Taiwan	2.22	10.24	2.78	10.44
Thailand	0.54	2.09	7.95	19.02
Turkey	9.46	4.06	6.44	11.98
Ukraine	2.32	1.71	3.50	10.85
UAE	1.08	0.58	0.98	18.24
UK	66.85	66.85	41.36	18.88
USA	56.20	52.33	87.66	18.38
Uruguay	0.07	0.56	0.01	17.93
Venezuela	1.31	0.72	4.24	17.37
Vietnam	0.43	2.24	2.22	7.75
Mean	7.11	7.91	7.65	16.01
SD	14.70	14.29	14.43	3.27
1st D %	21.00	29.80	47.50	40.20

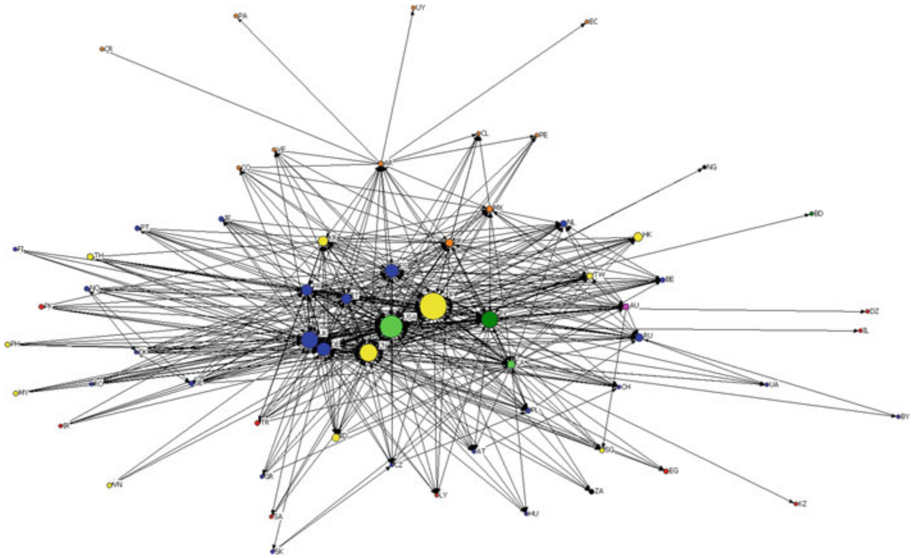

Fig. 1 International hyperlink network 2010, the size of the nodes is equal to the number of received links. *Blue nodes* are European; *yellow* East Asia; *light green* North America; *dark green* South Asia; *orange* Latin America; *red* Middle East and North Africa; *black* Africa; and *magenta* Oceania. One million links is required for a line. *Thicker lines* indicate a greater number of links. (Color figure online)

characteristic of the internet's physical infrastructure is indicated by the countries' "betweenness" centrality. The United States has a normalized betweenness centrality of 34.2, indicating that over a third of all internet traffic must flow through the country. This is nearly twice that of Germany (18.0 %) and almost three times the United Kingdom (12.8 %). Clearly, the United States acts as a hub in the international internet network. The centralities for the bandwidth network are also presented in Table 1.

Figure 2 graphically displays the bandwidth network, showing the United States as a hub. Also prominent in the figure are regional clusters in Latin America (bottom left), East Asia (top left), Russia, the Nordic countries (bottom right), and Southeastern Europe (top right). These groupings were confirmed through cluster analysis.

3.3 Shared website use: website perspective

The density of the internet based on shared websites is .162. The United States is the most central country using 128 of the 1,000 most-frequently visited websites. It had the greatest degree (51.4), betweenness (6.72), and eigenvector centralities (16.93). The United Kingdom is the second most central, with 125 websites and degree (50.2), betweenness (5.97), and eigenvector (16.58). Germany, India, Canada, and France were third, depending on the measure. The Gini-coefficient is .930, indicating that the Web-based network is very centralized among a dozen core countries (USA, UK, China, Germany, Brazil, France, India, Italy, Japan, Spain, and Russia). The centralities for the shared website network are also presented in Table 1.

Figure 3 graphically represents the relations among countries based on their shared use of websites. At the center are the core countries, with the exception of China and Russia. These two countries are the central nodes in regional clusters. China is grouped with Hong Kong,

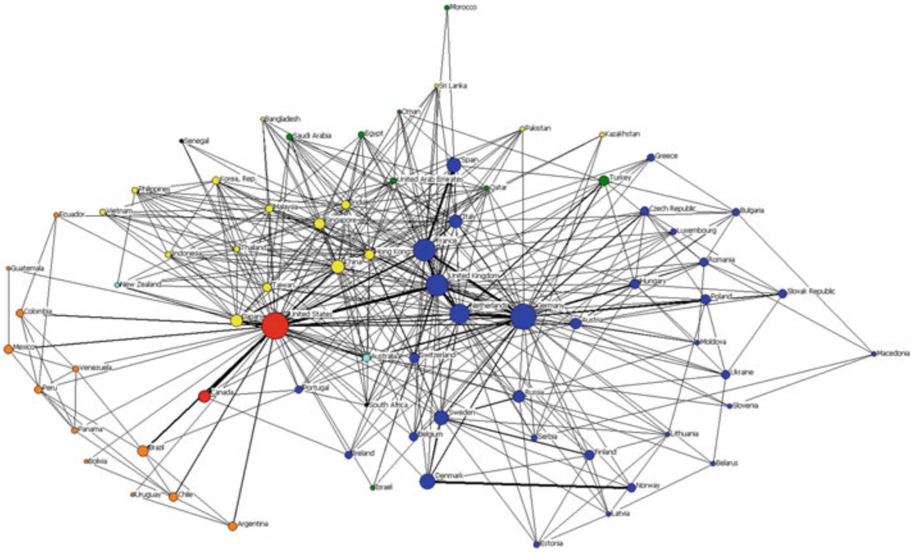

Fig. 2 International internet network based on bandwidth capacity, the nodes' sizes are equal to the square root of the total of a country's internet bandwidth capacity. *Thicker lines* indicate a greater bilateral bandwidth capacity

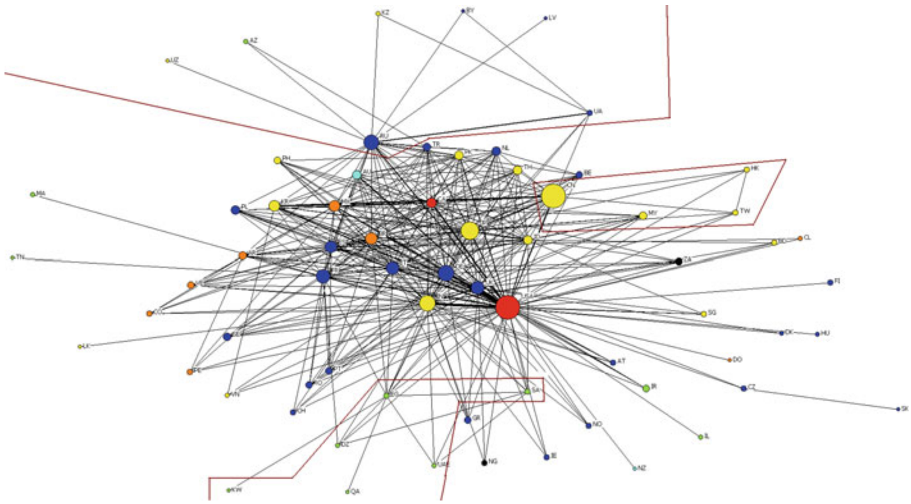

Fig. 3 International internet network based on shared website use, the size of the nodes is equal to the log of the total of the percentages of website use by a country. The mean percentage plus one standard deviation is required for a line. *Thicker lines* indicate a greater number of links

Taiwan, and Malaysia (to the upper right). Russia forms a clique with Azerbaijan, Belarus, Kazakhstan, Latvia, Ukraine, and Uzbekistan (along the top). Also (at the bottom), is another cluster composed of Arab nations (Egypt, Saudi Arabia, Algeria, Kuwait, Qatar, and UAE). These groups were verified using cluster analysis.

3.4 Shared website use: national perspective

The density of the internet based on the proportion of shared websites is .970. Almost all 160 countries jointly used at least one website in common with another nation. On the average, they shared 24.85 websites among the 100 most-frequently visited sites. There was little variance in this number (SD = 2.86). Pakistan was the most central country, followed by Costa Rica, Philippines, and Kenya. The Gini-coefficient is only .121, indicating that the network is not centralized but rather, that nations are grouped into clusters based upon language, geography, and historical circumstances. The eigenvector centralities for the proportion of shared websites are also presented in Table 1.

Figure 4 graphically represents the relations among countries based on the rankings of shared website use. Most prominent are three regional clusters, a grouping of Chinese countries (at the top right) a group of former Soviet republics, and (to the left) a large cluster of Arab countries. French speaking African nations cluster (to the lower left). There is a large grouping of Spanish-speaking Latin American countries (in the center to the left). There is a tight grouping of developed English-speaking countries (to the right of center), and (slightly to the right) there are a group of Asian countries with historical ties to the United Kingdom and the United States, where a great deal of English is spoken. These were confirmed by a cluster analysis. While the continental European countries appear just below the English-speaking group, they were not grouped by the cluster analysis, perhaps, because they speak a variety of different languages. Also, missing from the figure are Iran, Japan, Poland, Turkey, and Vietnam: these countries failed to have the minimum 24 shared links required to be connected to the other countries. Each of these countries has a sufficiently large population to sustain the production of websites and speaks a language that is not

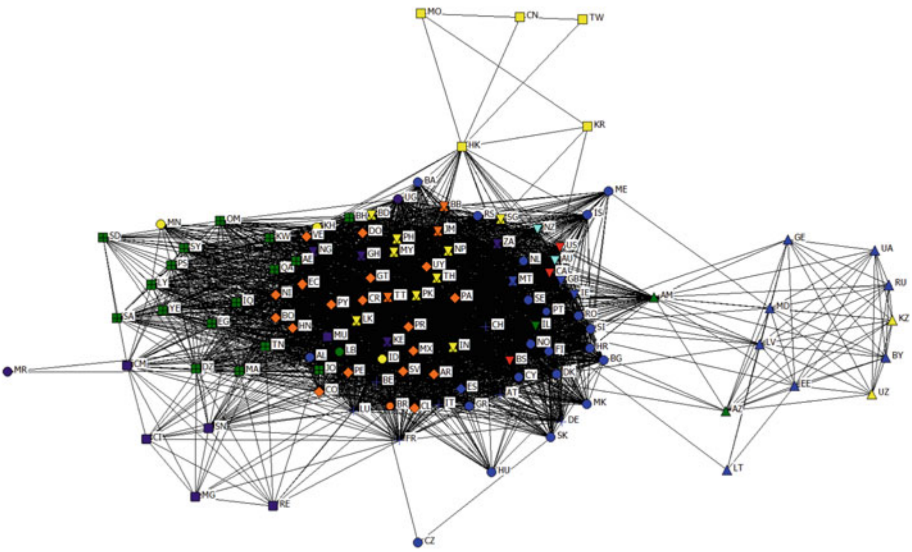

Fig. 4 International internet network based on countries' ranks of website use, shape indicates to which cluster a country belongs. A *circle* indicates that the country did not cluster. *Blue nodes* are European; *red* North America; *yellow* Asia; *orange* Latin America; *green* Middle East and North Africa; *magenta* Africa; and *light blue* Oceania. The *shape* indicates the country to which the cluster belongs. A *circle* indicates that the country did not cluster. (Color figure online)

Table 2 Correlations among networks: overall

	1	2	3	4
1. Bandwidth	1.0			
2. Hyperlinks	.442*	1.0		
3. Websites (%)	.478*	.350*	1.0	
4. Websites (country)	.064	.076*	.121*	1.0

$N = 75$, $p < .05$

Table 3 Correlations among networks: eigenvector centralities

	1	2	3	4
1. Bandwidth	1.0			
2. Hyperlinks	0.840*	1.0		
3. Websites (%)	.626*	.696*	1.0	
4. Websites (country)	−.092	.030	−.118	1.0

$N = 75$, $p < .000$

widely spoken outside their borders, which thus creates a unique self-sufficient Web environment. Therefore, the description of the internet based on the proportion of a country's shared websites as determined from its 100 most-frequently visited websites produced a unique picture.

3.5 The relations among the networks

Table 2 presents the Quadratic Assignment Procedure (QAP) correlations among the four networks for 75 countries for which there is data for all the network indicators (Dekker et al. 2007). The results indicate that three of the measures of the international internet (bandwidth, hyperlinks, and shared website use) are moderately related with coefficients ranging from .328 to .571, while website use from the country's perspective is at best weakly related to the other three measures.

Table 3 presents the correlations among the eigenvector correlations for the four measures for the same countries. The correlations for the first three measures are much higher, ranging from .626 to .849, indicating that these measures produced similar core-periphery structures. The measure based on the proportion of shared websites of a country's 100 most-frequently visited sites was not related to the other indicators. This is because the network described by this measure did not have a core-periphery structure. Rather, it might be described as a small world with a series of clusters based upon language, geography, and history.

4 Discussion

The results of the network analysis of the four different operationalizations of the internet provided four different, though somewhat related, descriptions of the structure of the network. The bandwidth network describes the physical network, and hyperlinks compose the navigational network (WWW), while the structural equivalence of websites describes the international distribution of information, either from the perspective of countries or websites. These differences may be due to a number of factors that may differentially impact the structures of these four networks including individual countries use of unique websites. For example, *Google*, the most-widely used search engine, uses a number of different domains (*google.com, google.co.in, google.de, google.com.hk google.co.uk, google.co.jp, google.fr, google.com.br, google.es*) for different countries. Typically, these sites use the language of the host country, resulting in differential preferences based on language and culture. For example, the former French colonies Algeria and Cameroon use *google.fr,* although they do not have a direct bandwidth connection to France. This is one source of the differences between the bandwidth and the website-use measures. Also, nations that share a common culture use the same websites and share preferential attachment to certain websites. This is seen clearly in the clusters of the Latin American, Arab, Russian, and Chinese nations across the different measures. Finally, as Barabási pointed out, as the internet grows, new websites tend to link with older established sites in order to increase their credibility and visibility (Barabási 2002).These websites are generally in the United States and other core nations. One result of preferential attachment may be that the countries that adopted the internet earlier are more central in the hyperlink network than countries that have experienced rapid recent growth, such as Brazil, India, and China, whose usage behavior reflects a greater reliance on domestic websites. Preferential attachment does not seem to occur when the network is examined from the perspective of the individual countries. In this case, the network appears to consist of series of small worlds determined by language, geography, and historical circumstances.

Clearly, one cannot depict the internet only through an examination of the hyperlink connections among nations. There is a need for multiple indicators to accurately describe the global internet (Doreian 2008). Traditionally, the different indicators of a network's structure were simply summed (multiplexed), perhaps after normalizing (Hartman and Johnson 1989). What is being proposed here is similar to the notion of the multitheoretical-multilevel framework proposed by Monge and Contractor (Monge and Contractor 2003), and Carley's (Carley 1999), notion of an ecology of networks, which suggests that a variety of network properties at different levels should be analyzed when examining social systems. In this case, the correlations among the eigenvector centralities, and the overall networks, suggest that it may not be sensible to combine all four data sets, but rather only the hyperlinks, bandwidth, and structural equivalence from the website should be treated as multiple indicators of the same network. Given the low correlations between the proportion of share websites and the other network measures, the former should be treated as a separate network.

While this article suggested examining the structure of the international internet using four different methods, a fifth may be suggested, one that uses keyword searches of online documents such as *Google Maps* to identify pairs of countries that are referred to in a single record (Zook et al. 2011). Given a sufficiently large sample of materials, a description of the structural equivalence of nations on the internet may be developed. Currently, research is underway that describes the relations among the countries of the world based on their co-occurrence in news stories from various global media that contain the root word "terror," obtained through the LexisNexis database. One goal of this research is to predict the resulting structure of web-based news networks from the four networks that describe the structure of the international internet.

Future research will examine the residuals of the correlations among the various measures of the internet's international structure. This will allow a better determination of the reasons for the systematic differences among the indicators. Additionally, each of the four measures will be examined longitudinally to determine how the internet has evolved over time, and to identify the existence of causal relations among the different indicators.

5 Summary and conclusions

This paper examined the network structure of the international internet using four different sources of data: (1) bilateral bandwidth between countries; (2) hyperlink connections among nations' domain names; (3) structural equivalence of nations from the perspective of websites, measured by the percentage of specific website's traffic from individual countries; and (4) structural equivalence of nations from the national perspective, using the proportion of a country's 100 most-visited websites shared with other countries. It described the methods to obtain the data used to depict the network in multiple ways, and the differences in the structures resulting from the four different perspectives. The use of multiple measures provided a more nuanced description of the structure of the international internet. There were moderate correlations among the measures. The reasons for the discrepancies included individual countries use of unique websites, nations with a common culture using the same websites, and preferential attachment.

Acknowledgments I would like to thank Jeanette B. Ruiz and Zhige Xin for their assistance in the data collection process, and acknowledge support from the Army Research Labs (W911NF-11-1-0405).

References

Barabási, A.: Linked: The New Science of Networks. Perseus, Cambridge (2002)

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

Barnett, G.A., Chon, B.S., Rosen, D.: The structure of international internet flows in cyberspace. Netw. Commun. Stud. (NETCOM) **15**(1–2), 61–80 (2001)

Barnett, G.A., Chung, C.J., Park, H.W.: Uncovering transnational hyperlink patterns and web-mediated contents: a new approach based on cracking .com domain. Soc. Sci. Comput. Res. Eval. (SSCORE) **29**(3), 369–384 (2011)

Borgatti, S.P., Everett, M.G., Freeman, L.C.: UCINET for windows: software for social network analysis. Analytic Technologies, Harvard (2002)

Borgatti, S.P.: Netdraw Network Visualization. Analytic Technologies, Harvard (2005)

Carley, K. M.: On the evolution of social and organizational networks. In Knoke, D., Andrews, S.B. (eds.) Networks in and Around Organizations. JAI Press, Greenwich (1999)

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

D'Esposito, M., Zaccarin, S.: Editorial: applied and methodological issues in the analysis of network data. Qual. Quant. **45**, 985–987 (2011)

Dekker, D., Krackhardt, D., Snijders, T.A.B.: Sensitivity of MRQAP test to collinearity and autocorrelation conditions. Psychometrika **72**(4), 563–581 (2007)

Dodge, M., Zook, M.: Internet-based measurement. In: Kitchin, R., Thrift, N. (eds.) International Encyclopedia of Human Geography, Elsevier, Oxford (2009)

Doreian, P.A.: Multiple indicator approach to blockmodeling signed networks. Soc. Netw. **30**, 247–258 (2008)

Hartman, R.L., Johnson, J.D.: Social contagion and multiplexity: communication networks as predictors of commitment and role ambiguity. Human Commun. Res. **15**, 523–548 (1989)

Hsu, C.-L., Park, H.W.: Sociology of hyperlink networks of web 1.0, web 2.0, and twitter: a case study of South Korea. Soc. Sci. Comput. Rev. **29**(3), 354–368 (2011)

Kleinberg, J., Lawrence, S.: The structure of the web. Science **294**, 1849–1850 (2001)

Monge, P.R., Contractor, N.S.: Theories of Communication Networks. Oxford University Press, Oxford (2003)

Park, H.W.: How do social scientists use link data from search engines to understand internet-based political and electoral communication. Qual. Quant. **46**(2), 679–693 (2012)

Park, H.W., Barnett, G.A., Chung, C.J.: Structural changes in the global hyperlink network 2003–2009. Glob. Netw. **11**(4), 522–544 (2011a)

Park, S.J., Lim, Y.S., Sams, S., Sang, M.N., Park, H.W.: Networked politics on cyworld: the text and sentiment of korean political profiles. Soc. Sci. Comput. Rev. **29**(3), 288–299 (2011b)

Park, H.W., Kim, C.S., Barnett, G.A.: Socio-communicational structure among political actors on the web. New Media Soc. **6**(3), 403–423 (2004)

Rosen, D., Barnett, G.A., Kim, J.H.: Social networks and online environments: when science and practice co-evolve. Soc. Netw. Min. **1**, 27–42 (2011)

Shumate, M., Lipp, J.: Connective collective action online: an examination of the hyperlink network structure of an NGO issue network. J. Comput. Mediat. Commun. **14**, 178–201 (2008)

TeleGeography.: Global Internet Geography. TeleGeography, Inc., Washington, DC http://www.telegeography.com/research-services/global-internet-geography/(2011)
Thelwall, M.: Introduction to Webometrics. Morgan & Claypool, New York (2009)
Weber, M.S., Monge, P.R.: The flow of digital news in a network of sources, authorities, and hubs. J. Commun. **61**, 1062–1081 (2011)
Zook, M., Devriendt, L., Dodge, M.: Cyberspatial proximity metrics: conceptualizing distance in the global urban system. J. Urban Technol. **18**(1), 11–93 (2011)