Webometric Analysis of Web Presence and Links of SAARC Countries

Vijayakumar M.[*], B.U. Kannappanavar[**], and Santosh Kumar K.T.[***]

*Central Library, Birla Institute of Technology, Deemed University, Mesra, Ranchi, Jharkhand-835 215
**Central Library, Kuvempu University, Shankaragatta, Shimoga, Karnataka
***Dept of Library and Information Science, Sahyadri College, Shivamogga, Karnataka

ABSTRACT

This study focuses on the identification of web presence and their links among SAARC countries. The research explores that India possesses maximum of 14, 10, 00, 000 webpages; 58, 20, 000 external links; 1, 18, 00,000 internal links; and 9, 83, 00,000 over all links. In case of web impact factor of external and internal links, Sri Lanka claims highest and for over all links once again India claims highest. One can note that, except India no SAARC country possesses all sub-domains, but these possess only a few sub-domains like .edu, .gov, .net, and .org. When analysis of the links was done, it was found that Pakistan has a maximum of 3,610 links to India as compared to other SAARC countries. India once again claims top position among the SAARC countries for Wiser ranking.

Keywords: Web presence, link analysis, webometric analysis, SAARC countries

1. INTRODUCTION

Despite the traditional information systems such as libraries, which categorise their information resources based on special criteria, the webpage is receptive for each kind of information from individuals and institutions. In other words, there is a no structural control on the data entering to webpage. Therefore, to find relevant and high quality information, in the easiest and fastest possible way, is a serious challenge for web user. The webpage is full of information and services including social/economical/science and technology information each of which has its own users.

Webometric covers research of all network-based communication using infometric or other quantitative measure. Webometrics, in future, may become one of the most interesting research area for the vast collection of electronic information available on the publicly indexable web. Paisley rightly identified this area as the future domain of bibliometric research. So an attempt has been made to analyse the web presence and their links of SAARC countries with different parameters.

2. SCOPE AND LIMITATION OF STUDY

Due to time limit, the study is focused to analyse the webpages and their links that are accessible from SAARC countries like Afghanistan, Bangladesh, Bhutan, Maldives, Nepal, Pakistan, India, and Sri Lanka. The advanced search facility of AltaVista and scholar-google was used to collect the primary data. Data collection was conducted to retrieve the links as well as the web presence of SAARC countries. The limitation of the study was that, instead of collecting the primary data with different search engines, we used only one search engine-Alta Vista was used.

3. LITERATURE REVIEW

Before conducting this study a literature review was done to bridge the gap between the research all ready and yet to be done.

Thomas & Willett[1] made a webometric analysis for the departments of library and information science (LIS) in UK universities. The findings indicate that it is not possible to identify any significant correlation between the citation data and peer evaluations of research excellence embodied in the research assessment exercise (RAE) rankings.

Thelwall[2], *et al.* made a webometric study of the universities of 16 European countries to know academic interlinking. It is found that universities' websites were tended to link mostly to countries geographically nearer.

Similar patterns may also appear within a single country such as the UK[3] and Canada[4,5].

Mukhopadhyay[6] made webometric investigation at different levels of domain name system based on the hyperlink analysis. Various web impact factor (WIFs) were calculated for cc to *p*-level domain (TLDs) of SAARC countries; sub-level domains (SLDs) related to academic and research institutes and hosts under Indian Institute of Technology and Indian Institute of Management educational system in India.

Aguillo[7], *et al.* made an analysis of web presence of the universities in Latin America using cybermetrics indicators. The authors have studies the Brazilian universities to know the web presence through web visibility and domain size. The result shows that there is a tremendous increase in the commitment of the Brazilian universities to the web and tried to visualise the web presence using the co-link maps of 167 Brazilian universities.

Jalal[8], *et al.* made a comparison among different types of ranking approaches for Southern Region universities having valid NAAC score. It is found that there is a high correlation between NAAC ranking and WISER ranking.

4. OBJECTIVES

The motive beyond this research was to identify the top SAARC countries with impact in the web environment, and to identify some of the variables that might correlate to the web impact factor of these SAARC countries. In particular, the study tries to know the following objectives:

(a) To find top and sub level web domains possessed by SAARC countries websites

(b) To find external links, internal links and over all links of SAARC countries on the web.

(c) To find generic top level domains that are linked to and from Indian domains

(d) To find SAARC countries that are linked to and from Indian domains

(e) Ranking of SAARC countries based on web impact factor

(f) Ranking of SAARC countries based on WISER law.

5. RESEARCH METHODOLOGY

The research followed the descriptive approach. The research method used in this study is survey method. In order to collect data, the list of SAARC countries provided by Wikipedia was used. Before using the list, we have checked the access of each SAARC countries website. The advanced search facility of AltaVista was used for data collection. Search expressions are as follows:

To identify the web pages, inlinks, self links and total links, Ingwersen[9] has given the following formula:

Total web pages	domain:domain name
Total external links	linkdomain:domain name NOT domain:domain name
Total self links	linkdomain:domain name AND domain:domain name
Total links	linkdomain:domain name

To identify the generic TLD that are linked to Indian domains, the search expression is:

linkdomain:in AND domain:generic TLD

To identify generic TLD that are linked from Indian domains, the search expression is:

linkdomain:generic TLD AND domain:in

Web Impact Factor

The WIF provides quantitative tools for ranking, evaluating, categorising, and comparing websites of top/ sub-level domains. To identify the WIF, Ingwersen[9] has given the following formula:

WIF for total external=Total external links divided by total pages

WIF for total self links=Total self links divided by total pages

WIF for total links =Total links divided by total pages

WISER Ranking

Aguillo[10], *et al.* has given the formula for WISER ranking as:

WISER ranking = Log (Visibility 50 %) + Log (Size 20 %) + Log (Rich files 15 %) + Log (Scholar 15 %)

6. ANALYSIS AND FINDINGS

6.1 Top Level Domains and Webpages of SAARC Countries

A TLD is one of the domains at the highest level in the hierarchical domain name system of the internet. Originally, the TLD space was organised into three main groups-countries, categories, and multiorganisations. Countries are designated in the domain name system by their two-letter ISO country code.

The TLD of SAARC countries are reflected in Table 1. Regarding webpages, India possesses maximum of 14,10,00,000 webpages followed by Pakistan, Sri Lanka, and Maldives.

Table 1. Web pages of under top levels domains possessed by SAARC countries[11]

S. No.	SAARC country	Top level domain	Number of web pages	Rank
1.	Afghanistan	.af	5,00,000	7
2.	Bangladesh	.bd	12,20,000	5
3.	Bhutan	.bt	2,94,000	8
4.	India	.in	14,10,00,000	1
5.	Maldives	.mv	5,54,000	6
6.	Nepal	.np	16,10,000	4
7.	Pakistan	.pk	1,57,00,000	2
8.	Sri Lanka	.lk	68,20,000	3

6.2 Sub-level Domains and Webpages of SAARC Countries

An SLD is a domain that is part of a larger domain. A 'sub-domain' expresses relative dependence, not absolute dependence. Organisations should register the domain name within the appropriate SLDs that corresponds to the type of the organisation/entity as outlined:

.co   For registered companies/trademarks/banks

.firm   For proprietary concerns/partnership firms/ shops/liaison offices

.ac   For academic community

.edu   For academic community

.res   For research institutes

.gov   For government organisations

.mil   For military establishments

.net   For internet service providers

.org   For non-profit organisations

.ind   For individuals

.gen   For general/miscellaneous purpose

The SLDs possessed by SAARC countries is shown in Table 2. It reflects that the SLD like .edu, .gov and .org is possessed by all SAARC countries.

The SAARC countries like Sri Lanka, Pakistan, Maldives, and Bhutan possess four SLDs Nepal and Afghanistan possess five SLDs and Bangladesh possesses six SLDs. Another interesting note is that, India possesses all SLDs shown in the Table 2.

6.3 Links or Hyperlinks

Hyperlink (or link) is a reference to a document that the reader can directly follow, or that is followed automatically. Hyperlinking means to create an internal or an external link. External link means, the links that are linked from outside domain, internal links means, the links that are linked with in the same domain.

Table 3 shows that India possesses highest of 1,18,00,000 internal links; 58, 20, 000 external links; and 9,83,00,000 over all links followed by Pakistan 49,10,000 over all links and Sri Lanka 16,60,000 over all links.

Table 4 illustrates the SAARC countries which are linked to Indian domain. India got maximum links from Pakistan (3,610) followed by Sri Lanka (2,070) and Nepal (728).

Table 5 shows Pakistan occupies first position to get a maximum of 18, 300 links from India, when compared to other SAARC countries. The next position goes to Sri Lanka 5,100 links followed by Maldives 2,520 links and Nepal 1740 links.

6.4 Generic Top-level Domains

A generic TLD is one of the categories of TLDs maintained by the internet assigned numbers authority (IANA) for use in the domain name system of the internet. The core group of generic TLDs consists of .co, .ac, .edu, .gov, .mil, .net, .org, .info, .com, and others.

Table 2. Webpages possessed under sub-level domain by SAARC countries[11]

S. No.	SAARC Country	.co	.firm	.ac	.edu	.res	.gov	.mil	.net	.org	.ind	.gen
1.	Afghanistan	2	-	-	2,800	-	25,4000	-	3	29,500	-	-
2.	Bangladesh	-	-	32,000	25,000	-	1,29,000	561	8,740	57,300	-	-
3.	Bhutan	-	-	-	3,790	-	35,000	-	303	4,110	-	-
4.	India	2,41,00,100	28,000	27,0,000	90,100	7,08,000	14,50,000	1	2,57,000	9,18,000	68,400	26,700
5.	Maldives	-	-	-	3,130	-	25,000	-	162	6,680	-	-
6.	Nepal	-	-	-	1,21,000	-	41,100	256	7,700	2,33,000	-	-
7.	Pakistan	-	-	-	3,54,000	-	2,91,000	-	67,400	1,57,000	-	-
8.	Sri Lanka	-	-	88,300	1,640	-	138,000	-	-	878	-	-

Table 3. Links possessed by SAARC countries[11]

S. No.	SAARC country	Internal links	External links	Over links	Rank
1.	Afghanistan	31,40	13,700	223000	6
2.	Bangladesh	1,24,000	33,700	397000	5
3.	Bhutan	25,400	12,900	103000	8
4.	India	1,18,00,000	58,20,000	98300000	1
5.	Maldives	55,000	20,800	174000	7
6.	Nepal	1,59,000	68,300	753000	4
7.	Pakistan	16,30,000	9,90,000	4910000	2
8.	Sri Lanka	7,69,000	4,43,000	1660000	3

Table 4. SAARC countries that are linked to Indian domains[11]

S. No.	SAARC countries	Search command	Links to Indian domain	Rank
1.	Afghanistan	linkdomain:in AND domain:af	415	4
2.	Bangladesh	linkdomain:in AND domain:bd	276	5
3.	Bhutan	linkdomain:in AND domain:bt	48	7
4.	Maldives	linkdomain:in AND domain:mv	65	6
5.	Nepal	linkdomain:in AND domain:np	728	3
6.	Pakistan	linkdomain:in AND domain:pk	3,610	1
7.	Sri Lanka	linkdomain:in AND domain:lk	2,070	2

Table 5. SAARC countries that are linked from Indian domains[11]

S. No.	SAARC countries	Search command	Links to Indian domain	Rank
1.	Afghanistan	linkdomain:af AND domain:in	177	6
2.	Bangladesh	linkdomain:bd AND domain:in	1,660	5
3.	Bhutan	linkdomain:bt AND domain:in	62	7
4.	Maldives	linkdomain:mv AND domain:in	2,520	3
5.	Nepal	linkdomain:np AND domain:in	1740	4
6.	Pakistan	linkdomain:pk AND domain:in	18,300	1
7.	Sri Lanka	linkdomain:lk AND domain:in	5,100	2

Table 6. Generic TLD that are linked to Indian domains[11]

S. No.	SAARC countries	Search command	Links to Indian domain	Rank
1.	.co	linkdomain:in AND domain:co	2,350	7
2.	.firm	linkdomain:in AND domain:firm	-	-
3.	.ac	linkdomain:in AND domain:ac	1870	8
4.	.edu	linkdomain:in AND domain:edu	93,300	5
5.	.res	linkdomain:in AND domain:res	-	-
6.	.gov	linkdomain:in AND domain:gov	12,000	6
7.	.mil	linkdomain:in AND domain:mil	106	9
8.	.net	linkdomain:in AND domain:net	17,00,000	2
9.	.org	linkdomain:in AND domain:org	14,70,000	3
10.	.ind	linkdomain:in AND domain:ind	-	-
11.	.gen	linkdomain:in AND domain:gen	-	-
12.	.info	linkdomain:in AND domain:info	1,75,000	4
13.	.com	linkdomain:in AND domain:com	3,62,00,000	1

Table 7. Generic TLD that are linked from Indian domains[11]

S. No.	SAARC countries	Search command	Links to Indian domain	Rank
1.	.co	linkdomain:co AND domain:in	7	9
2.	.firm	linkdomain:firm AND domain:in	-	-
3.	.ac	linkdomain:ac AND domain:in	14	8
4.	.edu	linkdomain:edu AND domain:in	154	5
5.	.res	linkdomain:res AND domain:in	-	-
6.	.gov	linkdomain:gov AND domain:in	110	6
7.	.mil	linkdomain:mil AND domain:in	21	7
8.	.net	linkdomain:net AND domain:in	3,660	2
9.	.org	linkdomain:org AND domain:in	3,300	3
10.	.ind	linkdomain:ind AND domain:in	-	-
11.	.gen	linkdomain:gen AND domain:in	-	-
12.	.info	linkdomain:info AND domain:in	872	4
13.	.com	linkdomain:com AND domain:in	7,790	1

Table 6 shows that .com possess first position (3,62,00,000 links) among the other generic TLD of the SAARC countries which are linked to Indian domain. The next position goes to .net (17, 00,000) followed by .org (14, 70,000) and .info (1, 75,000). But India is failing to get the link from the generic TLD like of SAARC countries.

From Table 6 and 7, some similarities can be seen. The generic TLDs like .com once again got the first position (7,790 links) among the other generic TLDs SAARC countries which majority of Indian links are linked to .com. Even .net (3,660 links) gets second position among others which majority of Indian links is linked to .net followed by .org (3,300 links), .info (872 links), and .edu (154 links).

6.5 Web Impact Factor

A WIF is a fairly innovative measure of the extent to which a webpage is linked to by other webpages. The WIF is a form of measurement used to determine the relative standing of webpages in particular fields or a country. Here higher the impact factor, the higher the perceived reputation of the webpage.

The WIF of external links possessed by SAARC countries is shown in the Table 8. Here, Sri Lanka occupies first position with 0.06495 impact factor followed by Pakistan (0.06305) and Bhutan (0.04387), It may be noted that, even though India possesses maximum

Table 8. Ranking of web impact factor for external links possessed by SAARC Countries[11]

S. No.	SAARC country	Total webpages links	Total external links	WIF of external	Rank
1.	Afghanistan	5,00,000	13,700	0.0274	8
2.	Bangladesh	12,20,000	33,700	0.027623	7
3.	Bhutan	2,94,000	12,900	0.043878	3
4.	India	14,10,00,000	58,20,000	0.041277	5
5.	Maldives	5,54,000	20,800	0.037545	6
6.	Nepal	16,10,000	68,300	0.042422	4
7.	Pakistan	1,57,00,000	9,90,000	0.063057	2
8.	Sri Lanka	68,20,000	4,43,000	0.064956	1

number of webpages and external links, India fail to get more WIF for its external links.

Table 9 gives the ranks distribution for WIF of internal links possessed by SAARC countries. Just like in Table 8, the Table 9 also reflects that Sri Lanka and Pakistan has succeeded to get first and second position, respectively, for WIF of internal links. Bangladesh got third position with the WIF 0.1016. India could not get top four positions.

Table 9. Ranking of web impact factor for internal links possessed SAARC Countries[11]

S. No.	SAARC country	Total webpages	Total internal links	WIF of internal links	Rank
1.	Afghanistan	5,00,000	31,40	0.00628	8
2.	Bangladesh	12,20,000	1,24,000	0.101639	3
3.	Bhutan	2,94,000	25,400	0.086395	6
4.	India	14,10,00,000	1,18,00,000	0.083688	7
5.	Maldives	5,54,000	55,000	0.099278	4
6.	Nepal	16,10,000	1,59,000	0.098758	5
7.	Pakistan	1,57,00,000	16,30,000	0.103822	2
8.	Sri Lanka	68,20,000	7,69,000	0.112757	1

The WIF for over all links is shown in the Table 10. India possesses first position by possessing highest WIF of overall links (0.6971) among SAARC countries. Nepal and Afghanistan got third and fourth position, respectively, by possessing 0.4677 and 0.446 WIF for over all links.

After verification of WIF of external link, internal link, and over all link, some very different results were found. Sri Lanka and Pakistan got first and second position, respectively, in WIF of internal and external links, but they failed to retain the same positions in WIF for over all links. Similarly, India and Nepal failed to get any top three position in the WIF of external and internal links, but they retained the first and second position in WIF of over all links.

Table 10. Ranking of overall web impact factor for over all links possessed SAARC countries[11]

S. No.	SAARC country	Total webpages	Over all links	Over all WIF	Rank
1.	Afghanistan	5,00,000	223000	0.446	3
2.	Bangladesh	12,20,000	397000	0.32541	5
3.	Bhutan	2,94,000	103000	0.35034	4
4.	India	14,10,00,000	98300000	0.697163	1
5.	Maldives	5,54,000	174000	0.314079	6
6.	Nepal	16,10,000	753000	0.467702	2
7.	Pakistan	1,57,00,000	4910000	0.312739	7
8.	Sri Lanka	68,20,000	1660000	0.243402	8

6.6 WISER Ranking

An attempt has been made to rank the webpages and the links of SAARC countries using appropriate webometric indicators. The details of the indicators is already explained in the choice of indicator, in addition to this WISER ranking which was used explained in the research methodology.

WISER rank of by SAARC countries is shown in the Table 11. Here, India occupied highest rank followed by Pakistan and Sri Lanka. Table 12 shows the rich files.

Table 11. WISER ranking of SAARC countries websites

S. No.	SAARC country	Webpages (Size)	External links (Visibility)	Rich files	Google scholar	WISER ranking	Rank
1.	Afghanistan	500000	13700	57292	415000	17.56402	6
2.	Bangladesh	1220000	33700	79890	529000	18.59212	5
3.	Bhutan	294000	12900	25701	48100	16.02321	8
4.	India	141000000	5820000	1E+07	2200000	25.61075	1
5.	Maldives	554000	20800	22524	32300	16.2756	7
6.	Nepal	1610000	68300	91294	379000	18.93251	4
7.	Pakistan	15700000	990000	595061	834000	22.23944	2
8.	Sri Lanka	6820000	443000	223140	381000	20.76187	3

Table 12. Rich files[11,12]

S. No.	SAARC country	Rich files						
		pdf	ps	doc	ppt	xls	rtf	Total
1.	Afghanistan	54100	-	2010	72	1110	-	57292
2.	Bangladesh	69600	4	8650	471	1150	15	79890
3.	Bhutan	24000	-	1300	47	354	-	25701
4.	India	8690000	14400	1030000	81000	217000	13800	1E+07
5.	Maldives	22100	-	224	36	164	-	22524
6.	Nepal	78800	-	10800	439	1240	15	91294
7.	Pakistan	479000	21	87400	6580	21700	360	595061
8.	Sri Lanka	202000	442	18500	642	1530	26	223140

7. FINDINGS

• The TLDs of SAARC countries are reflected in Table 1. Regarding webpages, India possess highest (14,10,00,000) webpages followed by Pakistan, Sri Lanka, and Maldives.

• The SAARC countries like Sri Lanka, Pakistan, Maldives, and Bhutan possess four SLDs, Nepal and Afghanistan possess five SLDs and Bangladesh possess six SLDs.

• India possessed maximum of 1,18,00,000 internal links, 58, 20, 000 external links, and 9,83,00,000 over all links followed by Pakistan 49,10,000 over all links and Sri Lanka 16,60,000 over all links.

• India got maximum links from Pakistan (3,610) followed by Sri Lanka (2,070) and Nepal (728).

• Pakistan occupies first position to get a maxmimum of 18, 300 links from India, when compared to other SAARC countires.

• .com possesses first position (3,62,00,000 links) among the other generic TLDs of the SAARC countries which are linked to Indian domain.

• Some similarities can be seen: the generic TLDs like .com got the first position (7,790 links) among the other generic TLDs SAARC countries which majority of Indian links are linked to .com. Even .net (3,660 links) gets second position among others which majority of Indian links is linked to .net followed by .org (3,300 links), .info (872 links) and .edu (154 links).

• Sri Lanka occupies first position with 0.06495 WIF for external links followed by Pakistan (0.06305) and Bhutan (0.04387). It is to be noted that, even though India possess maximum number of web pages and external links, India fail to get more WIF for its external links.

• Sri Lanka and Pakistan has succeeded to get first and second position, respectively, for the WIF of internal links.

• For WIF for over all links, India succeeded to get first position in the WIF of overall links (0.6971) among SAARC countries.

• Sri Lanka and Pakistan got first and second position, respectively, in WIF of internal and external links, but they fail to retain the same positions in WIF for over all links. Similarly, India and Nepal could not get any top three position in the WIF of external and internal links, but they retained the first and second position in WIF of over all links.

• India possessed first position as per the WISER rank for SAARC websites.

8. CONCLUSIONS

It would be practical to do study comparing more countries and evaluating web with traditional publication productivity and indicators of fiscal and technical development. Web presence and links analysis of countries is promising in general terms, but in practice the web presence and links are not entirely equivalent to citations in the scholarly publications.

By keeping this study as a webometric analysis of web presence and their links of sub-domains like .ac, .edu, and .res possessed by SAARC countries can be done.

REFERENCE

1. Thomas, O. & Willett, P. Webometric analysis of departments of librarianship and information science. *J. Inf. Sci.*, 2000, **26**(6), 421-28.

2. Thelwall, M.; Tang, R. & Price, E. Linguistic patterns of academic web use in Western Europe. *Scientometrics*, 2003, **56**(3), 417-32.

3. Thelwall, M. A comparison of sources of links for academic web impact factor calculations. *Journal of Documentation*, 2002, **58**(1), 60-72.

4. Vaughan, L. & Thelwall, M. Scholarly use of the web: What are the key inducers of links to journal web sites? *JASIST*, 2003, **54**(1), 29-38.

5. Vaughan, L. Why are websites co-linked? The case of Canadian universities. *Scientometrics*, 2007, **72**(1), 81-92.

6. Mukhopadhyay, P.S. Measuring web impact factors: A webometric study based on the analysis of hyperlinks. *In* National Seminar on Information Support for Rural Development, India. IASLIC, December 2004.

7. Aguillo, I.F.; Ortega, J.L. & Granadino, B. Brazil academic webuniverse revisited: A cybermetric analysis, 2006. http://digital.csic.es/bitstream/10261/4200/1/R-12.pdf (accessed on July 2010).

8. Jalal, S.K.; Biswas, S.C. & Mukhopadhyay, P.S. Webometric indicators for ranking of accredited Indian universities: An evaluatory study. *In* XXVII All India Conference of IASLIC on Library/Information Users in Digital Era, organised by KIIT University, 26-29 December 2009, Bhubaneswar.

9. Ingwersen, P. The calculation of web impact factors. *Journal of Documentation*, 1998, **54**(2), 236-43.

10. http://www.webometrics.info/methodology.html

11. http://in.altavista.com/

12. http://scholar.google.co.in/ (accessed on 23 July 2010).

About the Authors

Dr Vijayakumar M. is working as In-Charge-Library, Birla Institute of Technology, Mesra, Ranchi. He has more than 15 year professional experience as a teacher and administrator in India and abroad. He has presented many papers in national/international conferences and also published many research papers in professional journals. He has guided 25 MPhil projects and 8 PG projects.

Dr B.U. Kannappanavar is presently working as Selection Grade Librarian at Sahyadri College, Shimoga, Karnataka. He has published four books in LIS and 102 papers in leading journals and presented 43 papers in national/international seminars and conferences. He is guide to PhD students in LIS at Kuvempu University. Under his guidance four scholars have awarded PhD.

Mr Santosh Kumar K.T. is Assistant Professor in the Department of Library and Information Science, Sahyadri College, Shivamogga, Karnataka. He obtained MSc (LIS) in 1995 from Kuvempu University and PGDLAN from University of Hyderabad in 2005. He has 13 years of teaching experience. He attended several national/international seminars and conferences and published more than 15 articles. His areas of interest are: Library classification and cataloguing; information sources and services; and application of information technology.