CHAPTER—1

Introduction

[This chapter introduces background of the research problem and motivation for research. It mentioned the objectives, specific research questions and hypotheses. The methodology has been clearly specified. The broad area of research, scope and limitation are clearly mentioned to keep the focus of research in right direction.]
## Contents: Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Research Motivation</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Statement of the Research Problem</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Objectives of the Study</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Hypotheses</td>
<td>6</td>
</tr>
<tr>
<td>1.6 Methodology</td>
<td>7</td>
</tr>
<tr>
<td>1.6.1 Selection of Webometric Tools and Techniques</td>
<td>7</td>
</tr>
<tr>
<td>1.6.1.1 Search Engines</td>
<td>7</td>
</tr>
<tr>
<td>1.6.1.2 Web Crawlers</td>
<td>8</td>
</tr>
<tr>
<td>1.6.1.3 Visualization Tools and Techniques</td>
<td>8</td>
</tr>
<tr>
<td>1.6.1.4 Professional Search Engine Optimization (SEO) Tools</td>
<td>9</td>
</tr>
<tr>
<td>1.6.1.5 Web Citation Databases</td>
<td>9</td>
</tr>
<tr>
<td>1.6.2 Collection of URLs</td>
<td>9</td>
</tr>
<tr>
<td>1.6.3 Data Collection</td>
<td>10</td>
</tr>
<tr>
<td>1.6.4 Calculation of Web Impact Factors (WIFs)</td>
<td>11</td>
</tr>
<tr>
<td>1.6.5 Choice of Indicators for Ranking</td>
<td>11</td>
</tr>
<tr>
<td>1.6.6 Data Analysis</td>
<td>12</td>
</tr>
<tr>
<td>1.6.7 Data Validation</td>
<td>12</td>
</tr>
<tr>
<td>1.6.8 Data Interpretation</td>
<td>12</td>
</tr>
<tr>
<td>1.6.8.1 Arithmetic Mean (AM)</td>
<td>12</td>
</tr>
<tr>
<td>1.6.8.2 Standard Deviation (SD)</td>
<td>13</td>
</tr>
<tr>
<td>1.6.8.3 Co-efficient of Variation (CV)</td>
<td>13</td>
</tr>
<tr>
<td>1.6.8.4 Chi-square Test</td>
<td>13</td>
</tr>
<tr>
<td>1.6.8.5 Pearson’s Correlation Coefficient</td>
<td>13</td>
</tr>
<tr>
<td>1.6.8.6 Spearman’s Rank Correlation</td>
<td>14</td>
</tr>
<tr>
<td>1.7 Definitional Analysis</td>
<td>14</td>
</tr>
<tr>
<td>1.7.1 Indian Universities</td>
<td>14</td>
</tr>
<tr>
<td>1.7.2 The World Wide Web</td>
<td>15</td>
</tr>
<tr>
<td>1.7.3 Hyperlinks</td>
<td>15</td>
</tr>
<tr>
<td>1.7.4 Webometrics</td>
<td>15</td>
</tr>
<tr>
<td>1.7.5 Link Topology</td>
<td>15</td>
</tr>
<tr>
<td>1.7.6 Web Impact Factors</td>
<td>15</td>
</tr>
<tr>
<td>1.7.7 Webometric Tools and Techniques</td>
<td>16</td>
</tr>
<tr>
<td>1.8 Profile of the Study Area</td>
<td>16</td>
</tr>
<tr>
<td>1.9 Scope and Limitation</td>
<td>19</td>
</tr>
<tr>
<td>1.10 Conspectus</td>
<td>19</td>
</tr>
<tr>
<td>1.11 Summary</td>
<td>21</td>
</tr>
<tr>
<td>References</td>
<td>21</td>
</tr>
</tbody>
</table>
1.1 Introduction

In ancient India, Monastic order of education under the supervision of guru was the form of education in the area of philosophy, religion and ancillary branches. Secular Buddhist institutions, urban centers for learning were established e.g. Nalanda, Takshila, Ujjain, Vikramshila and Vallabhi\(^1\). Takshila was famous for medicine while Ujjain was renowned for astronomy. Nalanda covers all branches of knowledge. During Gupta period, India became a renowned center for higher education with students and scholars not only from several parts of India but also from abroad.

The foundation for modern higher education was laid by the Britishers setting up network of schools to impart western education in English medium. First college was set up at Serampore (near Calcutta) in 1818 and then at Agra, Bombay, Madras, Nagpur, Patna and Calcutta. McCauley’s policy in 1835 to promote European education through English language and Sir Charles Wood’s Dispatch of 1854 recognized the need for mass education, administration of education in each province and the establishment of a few universities in Bombay, Calcutta and Madras in 1857. First Indian Education Commission in 1882 recommended for the initiative of private agencies in expansion of higher education. After independence, Maulana Azad, the first education minister envisaged strong central government intervention to spread education throughout the country. Several other issues were addressed like increased investment on higher education, growing awareness on

\(^1\) Vallabhi, famous for Jain center is located in Saurashtra in Gujarat. Vallabhi was famous for its catholicity and the students from all over the country, including the Brahmana boys, visited it to have higher education in secular and religious subjects.
education, formation of several Five-Year plans, judicious framing of policies on higher education, establishing the apex body to monitor, regulate and ensure the qualities of education.

Since the inception of civilization, communication has taken place in many ways and many forms like from sign language to body language, from spoken words to written words, clay tablets and stone inscriptions to print on paper by virtue of Gutenberg’s invention of printing press, from print documents to digital documents. In the family of digital documents, web documents are considered as the most popular and prominent form. It has a tremendous impact on scholarly communication through Internet (Joshi, 2000) including formal publication of research results, accessibility of books and journals in anywhere and at any time. The existing forms of delivery of education in the 18th and 19th century have come across tremendous change in 21st century due to change in information and communication technologies (ICTs). ICT has increased the flexibility of delivery of education so that access to knowledge may take place anytime, anywhere and at anyplace. The influence of globalization and privatization on higher education has led to increased competitive environment and therefore, marketing is inevitable process of higher education. Traditional classroom teaching was transformed into learning studios featuring wireless technology, video conferencing, enquiry-based teaching, and a place where faculty share their experiences with other faculties in a collaborative and supportive way, small group discussion, e-learning, virtual learning and web-based learning.

The Web is a reflection of human culture, massive socio-cultural network of Web resources authored by millions of people and organizations around the world. The significance of the Web in social and economic life of human being is inseparable since mid 1990s and needs no introduction. The influence of the web on academic especially higher education and research activities is no exception. As a consequence, there is an urgent need to understand the potential of the Web for generating new type of knowledge (Davenport & Cronin, 2000) through the power of hyperlinking nature and behavior of the web. The potential of hyperlink analysis was recognized due to the facilities extended to conduct various types of searches by search engine AltaVista. Borgman and Furner (2002) pointed out the possibility for analyzing web data with established citation techniques (Ingwersen, 1998; Larson, 1996; Rodriguez i Gairín, 1997; Rousseau, 1997) to interpret inter-relationship between web documents. This led to the birth of new field Webometrics. The term was coined by Almind and Ingwersen in 1997 for denoting the quantitative analysis of web related data. According to Björneborn and Ingwersen (2004) Webometrics is “the study of quantitative aspects of the construction and use of information resources structure and technologies on the web drawing on Bibliometrics and Informetrics approach”.

Web is an infinite source of resources. The growth and impact of Web is very difficult to calculate. Lawrence and Gill (1999) reported in the journal ‘Nature’ that publicly indexable webpages containing more than 6 terabyte of text data exists in about three million servers. National Association of Software and Service Companies, India (NASSCOM, n.d.) reported that there are about 60 million Internet users in India in June 2010. At present, almost all the educational institutes including Indian universities are having their own websites for providing
information support to the students, research scholars and faculty members. The University Grants Commission (UGC) and Ministry of Human Resource Development (MHRD), Government of India have already taken initiative for the computerization of universities through Information and Library Network (INFLIBNET), an Autonomous Inter-University Centre (IUC) of UGC, making availability of e-resources through Indian National Digital Library in Engineering Science & Technology (INDEST) consortium and UGC-Infonet programme, website development and technical supports. The rapidly developing electronic and digital media have a tremendous impact on the universities and research institute for the subscription to e-journals, databases, e-books, electronic publishing and digital library. Many research results, scholarly articles and research papers, and books are available only in the digital media i.e. through the Web. In fact, there is a paradigm shift in higher education, shift in rapid technological advancement, shift from information society to knowledge society, from cooperation to fierce competition, shift from quality assurance at national level to at par global level.

In order to stay in the competitive market, it is understood that there is a need to maintain a quality education, not only at the state level, regional and national level but also at the global level. Bibliometric laws and citation techniques generally were used to measure the quality at the print environment, not only at the individual level but also at the institutional level. But, from the middle of 1990s onwards, the form of scholarly communication has changed from print to digital media. All stakeholders in educational sector increasingly depend on the Web. Hence, traditional bibliometric and informetric laws have been applied for measuring the quality of higher educational institution through the web by assessing the performance of the websites in terms of link analysis, web content analysis and web citation analysis. One important measure may be the application of Web Impact Factor (WIF) to evaluate the quality of websites of institutes or universities having valid domain names. The calculation of WIFs for Indian universities may be performed using search engines that support special search techniques required for webometric research. Webometrics tries to measure the World Wide Web (WWW) in order to get knowledge about number and type of hyperlinks, reference structure of the WWW and its usage pattern.

The changing attitude of the people, the impact of globalization and the advent of Internet have opened up the opportunities for all the stakeholders, who may be interested to know the best universities not only at national level but also at global level. The main reason for ranking the universities is the vast growth in higher education across the world in recent decades. As a result, ranking the academic institutes is of immense attraction not only from public initiative but also from private sector. The first ranking system for the universities in the world is dated back to 1983. A few existing popular international initiatives on ranking the universities are: (a) Academic Ranking of World Universities (ARWU), Shanghai; (b) Center for College Affordability & Productivity, USA; (c) Global Universities Ranking, Russia; (d) Macleans, Canada; (e) International College & Universities (4ICU); (f) Ranking Web of World Universities (RWWU) by Cybermetrics Lab.

---

1 US News and World Report published the America's Best Colleges in 1983.
2 A research body in Spain headed by Prof. Isidro Aguillo, who is responsible for publishing World Universities' Ranking on the Web twice a year covering more than 20,000 Higher Education Institutions worldwide.
(webometrics.info); (g) Times Higher Education Ranking (THE) etc. Recently, India is having National Assessment and Accreditation Council (NAAC) for ranking and accreditation of higher educational institutes.

The measuring the Web presence of universities may be done through various aspects of websites. Web presence implies the existence of domain names, space occupied in the Web for uploading the files covering the objectives, activities, admission, course materials, details of staff, academics, notices, etc. Several other indicators, which represent the web presence by the number of webpages, visibility as measured through the number of hyperlinks, rich files, Google scholar, Age of website, etc. The performance of the website may require to measure through analysis of hyperlinks. The idea of measuring average link frequencies i.e. WIF was developed by Peter Ingwersen in 1998 (Ingwersen, 1998). It is a fact that Rodriguez i Gairin (1997) had introduced the concept of information impact on the Internet in a Spanish documentation journal. Web Impact Factors (WIFs) are web versions of the Impact Factors (IF). Generally, it is defined as the number of pages linking to a site or area in Internet divided by the number of pages in that site. A high value of WIFs indicates a site with great impact because there are relatively many pages linking to the site. Till today, researchers’ need is to depend on the commercial search engines to collect the necessary data for calculating WIF. The popular search engines helpful in this endeavor are AltaVista, Web Crawler, Yahoo!, Hotbot, etc. It is known to everybody that the search engine’s coverage and capability are limited with respect to the volume of webpages. The data for link analysis have been collected for universities in India using relevant commercial search engine with the provision of special keywords for Webometrics studies. Such data sets can be utilized to measure Web presence of Indian universities (Jalal, Biswas & Mukhopadhyay, 2009a) and their impact through WIF on the teaching-learning-research process.

The literature survey has revealed that large number of studies on academic interlinking were conducted on UK universities (Thelwall, 2001b, 2002a, 2002b), other countries and geographical areas e.g. Asia-Specific region (Thelwall & Smith, 2002), EU countries (Thelwall et al., 2002), China (Tang & Thelwall, 2002) and Canada (Vaughan & Thelwall, 2005). Several studies have found that the number of hyperlinks to websites varies with the quality of the university using various definitions of quality.

1.2 Research Motivation

As electronic resources gradually increase and easily available through the Web, many research scholars depend on the Web along with print documents. Society journals, conference proceedings, scholarly forth-coming articles, born-digital journals, wikis, blogs, discussion forum and other forms of digital documents appear in every discipline. Web is an important component of scholarly communication. Borgman (2000, p.144) described the scholarly communication "the study of how scholars in any field use and disseminate information through formal and informal channels". Many researchers (Thomas & Willett, 2000; Thelwall, 2001a; Smith & Thelwall, 2002; Wilkinson et al. 2003; Thelwall, Vaughan & Björneborn, 2005; Noruzi, 2006; Elgohary, 2008) have conducted the webometric studies on different countries but there was no such comprehensive
study that has been conducted in India so far to identify the details of their web performance, link analysis, and structure of websites using webometric techniques. Therefore, it is the high time to evaluate websites of the universities in India with respect to their web presence and popularity of websites in the web-space and to discover link pattern and web integration. The study may help university authorities to know the status of the universities’ website in comparison with other universities to take some effective measures for improving the impact of websites. The webometric study aims to reveal a particular university’s position in the list of universities ranking in India through hyperlink analysis using webometric tools and techniques.

As the Web is continuously developing, users’ need are constantly changing, organizations have increasingly focused on developing sophisticated information and communication technologies to introduce e-services to their customers. Therefore, more and more emphasis have been given to develop good websites through selective and rich content, updated information, proper maintenance, published R&D results, to launch attractive package for students and staff. A pilot study (Jalal, Biswas & Mukhopadhyay, 2008) was conducted to know the web presence of universities in India and found that more than 450 recognized universities in India are having their own domain names. Through literature survey, it was realized that was no comprehensive webometric study of Indian universities. So, there exists a gap in knowledge and in order to fulfill the gap, there is a need to conduct a comprehensive study concentrating on the following aspects:

- Web presence of Indian universities;
- Hyperlink analysis of Indian universities;
- Designing and developing a ranking system for Indian universities;
- Formation of link topology in Indian academic web space.

With this understanding, the research problem has been identified in the area of webometrics and broadly expressed as follows:

"Indian Universities on the Web: Analysis of Hyperlinks through the Application of Webometric Tools and Techniques"

1.3 Statement of the Research Problem

The title of the present research work ‘Indian Universities on the Web: Analysis of Hyperlinks through the Application of Webometric Tools and Techniques’ represents a generic statement of problem. However, a close analysis of research problem formulated above reveals that it is readily amenable to a set of five interlinked questions. These are as follows:

1. What is the status of the web presence of Indian universities? How can it be measured?
2. Is it possible to rank Indian universities at the national, regional and state levels by applying appropriate webometric indicators?
3. Is there any correlation amongst web ranking systems for Indian universities? If so, what is the degree of correlation?
4. Is it possible to identify link topology for selected Indian academic web space? What factors are associated with the formation of such link topologies?

5. What are the possible reasons for hyperlink among the Indian academic web space?

1.4 Objectives of the Study

The primary objective of this research study is to understand the present situation of web presence, nature and features of link relationship, ranking based on website performance of Indian universities based on hyperlink analysis through webometric tools and techniques. To achieve the objective, the current research project precisely makes an endeavor to reach the following specific objectives:

1. To find out various types of hyperlinks, explore the web presence and calculate various web impact factors for selected Indian universities;

2. To develop a new method of ranking using appropriate webometric indicators to rank the Indian universities at the state, regional and national level;

3. To compare and correlate the existing ranking approaches (e.g. NAAC, RWWU, THE) with the newly developed ranking approach;

4. To identify the patterns of hyperlinks and to find out the existence and nature of link topology in the web space of the selected Indian universities; and

5. To identify the motivational factors for the creation of hyperlinks in the websites of the universities.

1.5 Hypotheses

Generally, hypothesis is more specific than a research problem with an aim to formalize the research questions. Hypothesis may be of two types: Null hypothesis usually denoted by $H_0$ and Alternative hypothesis denoted by $H_1$. According to R.A. Fisher, null hypothesis is the hypothesis, which is tested for possible rejection under the assumption that it is true. The literature study helps to formulate the following hypotheses, which will ultimately give direction to keep the study in right way. Following hypotheses are set for the present study as mentioned below:

1. There is a high degree of web presence of Indian universities;

2. It is feasible to rank Indian universities by the application of appropriate webometric indicators;

3. Webometric techniques are helpful to compare and correlate newly developed Indian University Ranking System with other quality ranking systems (RWWU, NAAC, 4ICU etc.);
4. The websites of Indian universities have well-structured pattern; the topology of Indian academic web space as constituted by the selected India universities can be structured through the application of webometric tools and techniques; and

5. Classification and categorization of hyperlinks help to understand the motivational factors for the creation of hyperlinks.

1.6 Methodology

"Theory suggests what should be measured. Data limits what can be measured"

(WISER, 2001: Part-B, p.18)

The methodology of the study covers selection of relevant search engines and data collection procedures and finally analysis of data for the purpose of fulfilling the objectives of the study. The methodology basically answers two crucial questions: (i) How was the data collected? (ii) How was the data analyzed?

This section sketches in detail the various data collection methods during period of study. This is very crucial because the result may differ significantly due to change in data collection procedure. The validity and reliability of results are fully dependable on data collection methodology. The methodology also addressed the problems, which may encounter during data collection and suggests some way out to minimize the problems. The data are collected in a systematic manner using webometric tools and techniques. The specific methodologies wherever applicable are given in respective chapters.

1.6.1 Selection of Webometric Tools and Techniques

Webometric research may be small-scale or large-scale endeavor. In the case of small-scale, it may be possible to browse each webpage manually for counting the exact number of webpages, total links, self-links of the website, incoming links and outgoing links. But this technique is not practically feasible for the large-scale data. As a result, there is a need to choose appropriate webometric tools for data collection. The tools may be categorized as commercial search engines and personal web crawlers. Besides, there are some visualization tools e.g. Pajek, Ucinet 6.0 etc. which are helpful to visually display the pattern of hyperlinks or structure of the website.

1.6.1.1 Search Engines

Selection of search engine is the pivotal task and fundamental pillar on which the significance of whole research work depends. Wrong selection of search engines may ultimately leads to misinterpretation and misleading results. Although there are some limitations for commercial search engines as pointed out by Rousseau (2001) and Bar-Ilan (2001) to accept them as data collection tool for Webometric research, still commercial search engines may be appropriate tool for data collection opined by many webometricians. AltaVista, Google and Yahoo! have been selected for our study to collect web pages and link data during the specified
period as mentioned in the respective tables. As Google has bigger index size, it may be quite reasonable to use Google for retrieving the value of webpage (i.e. size) but Google is not suitable for collecting link data because it does not support the linkdomain command. As AltaVista is being purchased by Yahoo!, it would be logical to use Yahoo! as data collection tool for link analysis because of its wide spectrum of data coverage. In the case of Scholar value and Rich File, the Google Scholar and Google in general are used to retrieve the data for the selected set of web sites under study.

1.6.1.2 Web Crawlers

For data collection, Thelwall (2001a), University of Wolverhampton, UK has created specialist information science personal Web Crawlers such as LexiURL and SocScibot specifically for link analysis to study academic websites. These are designed for non-commercial use and especially designed to crawl data for academic web domain. There are many other web crawlers. One such crawler is Websphinx, which is an interactive customizable web crawler base on Java environment (Websphinx, n.d.). It has basically two parts: Crawler Workbench and the Websphinx class. It browses and processes web pages automatically and retains all the pages and links that it has crawled until one clears the crawler. It visualizes a collection of web pages as a graph.

1.6.1.3 Visualization Tools and Techniques

Visualization tools are helpful for graphically representing the link pattern and relationship among various domains. There are many visualization tools such as Ucinet, Pajek, Netdraw, Websphinx etc. Pajek is one of the popular softwares used for visualizing the large network links (Pajek, n.d.). It is associated with SocScibot tool.

Figure 1.1— The graphical structure of Burdwan University Website (26/01/10)
Chapter 1: Introduction

**blue** : for links (the A tag)
**red** : for tables (TABLE, TR and TD tags)
**green** : for the DIV tag
**violet** : for images (the IMG tag)
**olive** : for forms (FORM, INPUT, TEXTAREA, SELECT and OPTION tags)
**orange** : for linebreaks and blockquotes (BR, P, and BLOCKQUOTE tags)
**black** : the HTML tag, the root node
**gray** : all other tags

For example, the structure of the website for the University of Burdwan (i.e. http://www.buruniv.ac.in) can be known through its webpages as Graph\(^4\). The graph is generated instantly having online connectivity and all the specifications as mentioned above. Graph is generated from http://www.aharef.info/static/htmlgraph/

Ucinet 6.0 is a social network analysis programme developed by Steve Borgatti and Martin Everett and distributed by analytic technologies. Ucinet works with freeware program, which is called Netdraw 2.097 for visualizing networks and reads Ucinet system files, Ucinet DL files, Pajek files, etc.

1.6.1.4 Professional Search Engine Optimization (SEO) Tools

Alexa is a powerful tool to discover information about websites and various statistics. It can be used to discover how popular a site is, to find new sites, to learn who owns a site, the age of the site and lot more. Linkvendor’s professional SEO tools have so many facilities. One important service is ‘Searchmetrics report’, which conducts an extensive site analysis based on the most important search engine key figures. Other tools are used like linkvendor.com for page rank and webconfs.com for domain age calculation.

1.6.1.5 Web Citation Databases

During the study citation databases have been consulted especially for ranking the universities (Chapter 5) and for finding the correlation between WIFs and research rating as measured through publication of the institute. The popular citation databases are Google Scholar (SC), Web of Science (WoS) and Scopus. In order to get the citation data, Scopus has been consulted during June 2011. Publication, citations and h-index values for each university (e.g. University of Delhi) for the period of last 5 years i.e. 2006 to 2010 has been collected under the institutional affiliation using query syntax:

\[
\text{AFFL ("University of Delhi") AND (PUBYEAR AFT 2005 AND PUBYEAR BEF 2011)}
\]

1.6.2 Collection of URLs

There are more than 570 universities in India. The URL of each university has been collected by conducting Internet search, literature search and through various

\(^4\)The program has been developed by Marcel Salathe under http://www.aharef.info/static/htmlgraph/
Chapter 1: Introduction

websites. The UGC website [www.ugc.ac.in](http://www.ugc.ac.in) helps to find out most of the URLs of universities under different states. Each URL has been double checked. It has been observed that some universities are having more than one URL. One such example is Rabindrabharati University having two web URLs i.e. [http://www.rabindrabharatiuniversity.net](http://www.rabindrabharatiuniversity.net) and [http://www.rbhu.ac.in](http://www.rbhu.ac.in). A university may have two URLs but it opens the same site. One such example is IIT Kharagpur, West Bengal ([http://www.iitkgp.ac.in](http://www.iitkgp.ac.in) and [http://www.iitkgp.ernet.in](http://www.iitkgp.ernet.in)). It is also experienced during the research that couple of universities' URL has been changed and six universities have no URLs.

1.6.3 Data Collection

The data are collected from concerned universities websites with the help of selected search engines using a few selective query statements. The important query syntaxes, which have been used for the data collection purpose, are *link*, *host*, *linkdomain* and combination of these commands using Boolean operator (AND, OR, NOT) or alternatively, one space used instead of 'AND' and '−' is used to denote instead of NOT. Using commercial search engine, the following key data are generally collected:

i. Total Number of pages
ii. Number of links to websites
iii. Number of self-links
iv. Number of external-links
v. Number of Rich Files
vi. Number of publications and citations, scholarly articles

Yahoo! Explorer, Google and Google Scholar have been used heavily for data collection during the study.

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Syntaxes</th>
<th>Results</th>
<th>Supported by</th>
</tr>
</thead>
<tbody>
<tr>
<td>site</td>
<td>site:buruniv.ac.in</td>
<td>Reports total number of webpages under the domain <em>buruniv.ac.in</em></td>
<td>Yahoo!, AltaVista. Google, MSN, Teoma, Exalead, Gigablast</td>
</tr>
<tr>
<td>domain</td>
<td>site:buruniv.ac.in</td>
<td>Reports total number of webpages under the domain <em>buruniv.ac.in</em></td>
<td>Yahoo! AltaVista. AllTheWeb Google</td>
</tr>
<tr>
<td>link</td>
<td>link:buruniv.ac.in</td>
<td>Reports the total links</td>
<td>Exalead</td>
</tr>
<tr>
<td>linkdomain</td>
<td>linkdomain:buruniv.ac.in</td>
<td>Reports the total links</td>
<td>Yahoo! AltaVista. AllTheWeb</td>
</tr>
<tr>
<td>linkdomain &amp; domain</td>
<td>linkdomain:buruniv.ac.in AND domain:buruniv.ac.in</td>
<td>Reports total links within the domain i.e. self-links</td>
<td>Yahoo! AltaVista. AllTheWeb</td>
</tr>
</tbody>
</table>
### Chapter 1: Introduction

<table>
<thead>
<tr>
<th>linkdomain &amp; domain</th>
<th>Inlinks/external links</th>
<th>Yahoo!, AltaVista, AllTheWeb</th>
</tr>
</thead>
<tbody>
<tr>
<td>linkdomain: buruniv.ac.in</td>
<td>Retrieved total number of webpages in the respective language</td>
<td>Yahoo!, AltaVista, AllTheWeb</td>
</tr>
<tr>
<td>domain: buruniv.ac.in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>language: english</td>
<td>Retrieved the filetype of the particular domain</td>
<td>Google, Gigablast, Exalead</td>
</tr>
<tr>
<td>filetype: allintitle: site: buruniv.ac.in filetype: php</td>
<td>Finds pages that contain the specified word or phrase in the title</td>
<td>AltaVista, Google, AllTheWeb</td>
</tr>
<tr>
<td>title: text</td>
<td>Reports co-link pages between University of Burdwan &amp; Delhi Universit</td>
<td>Yahoo!</td>
</tr>
<tr>
<td>co-link: linkdomain: buruniv.ac.in AND domain: du.ac.in</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 1.6.4 Calculation of Web Impact Factors (WIFs)

The WIF analysis provides a way to know the international visibility and relative importance of a particular university’s website. It is the quantitative tools for ranking, evaluating, comparing websites at the host level, top-level and sub-level domains. There are three types of WIFs: \( WIF_{\text{simple}} \), \( WIF_{\text{revised}} \) and \( WIF_{\text{overall}} \). The following formula may be used for the calculation of WIFs:

\[
\begin{align*}
A &= \text{Number of external backlinks to a given site} \\
B &= \text{Number of self-links to a given site} \\
C &= \text{Total number of links to a site} \\
D &= \text{Total number of webpages}
\end{align*}
\]

Therefore, \( WIF_{\text{revised}} = (A/D) \), \( WIF_{\text{selflink}} = (B/D) \) and \( WIF_{\text{overall}} = (C/D) \)

#### 1.6.5 Choice of Indicators for Ranking

Webometric indicators are the guidelines for evaluating websites of universities or institutions for the purpose of ranking them. There is no unique and supreme method, which is undisputed and uncontroversial for ranking the universities. But each type of method is having certain qualities and limitations. There are various approaches and methodologies applied for ranking Indian universities (Jalal, Biswas & Mukhopadhyay, 2009c). A few important and well-known methodologies are: (a) Ranking based on NAAC, (b) Ranking based on Times Higher Education (THE), (c) Ranking Based on International College & University (ICU) Ranking, (d) Ranking Web of World Universities (RWiWU), (e) Webometric Ranking of Indian University (WRUI), which has been proposed in the thesis. Each of these methodologies has their own indicators, some are common to each other, and some are unique. Besides, each indicator is assigned various weights by various ranking methods.
1.6.6 Data Analysis

Data analysis is the crux of any analytical and empirical research. Data itself is meaningless unless it is been analyzed and given the contextual meaning. Some basic statistical calculations were made wherever required. Correlation testing is used for data analysis in order to find out the relationship among variables. For example, to know whether there is any positive relationship between RWWU rank and NAAC rank for all the universities having a valid NAAC ranking. Correlation testing can be done using either Spearman’s Rank Correlation or Pearson’s correlation coefficient but rank correlation would seems to be more appropriate. Data analysis can be done through calculating various WIFs.

1.6.7 Data Validation

Data validation is a process to determine whether the data is accurate and complete. Validity can be judged by asking some questions and will look for answer, which may support by other researchers. The data may be raw or it may be of secondary in nature. But, the data used in the thesis is mostly raw one, which are collected directly from website using special webometric search keywords or data may be collected using personal web crawlers.

Several previous web based studies (Tang & Thelwall, 2003; Thelwall & Tang, 2003; Thelwall & Price, 2003; Thelwall & Harris, 2004, Thelwall, 2004) have taken the help of Kolmogorov-Smirnov test (Corder & Foreman, 2009), the principal goodness of fit test for normal and uniform data sets. Spearman’s coefficient is preferred to Pearson’s coefficient in this study because web data are non-parametric by nature and skewed distribution of data. This is significant methodological issue for the selection of testing tool for data validation. Also, Chi-Square test may be followed wherever applicable.

1.6.8 Data Interpretation

Data can be interpreted using some statistical tools or graphical representation. Graphical tools like line chart, bar chart, multiple bars, and pie chart etc may be highly useful to interpret the data. Following are some common statistical tools, which can be applied.

1.6.8.1 Arithmetic Mean (AM)

The arithmetic mean is calculated as the sum of numbers divided by the number of items in the list. The arithmetic mean is simply called mean. The mean may be population mean or it may be sample mean. If we denote a set of data, e.g. \( X = \{x_1, x_2, ..., x_n\} \), then the mean (\( \bar{x} \)) is typically denoted as:

\[
\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i = \frac{1}{N} (x_1 + x_2 + \ldots + x_n)
\]

In the case of weightage average, the above formula will be as follows:

\[
\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i f_i, \text{ Where } N = \sum_{i=1}^{N} f_i
\]

The value of arithmetic mean signifies that the particular value represents the whole set of data.
1.6.8.2 Standard Deviation (SD)

Standard deviation reflects the variability or dispersion of a data set or distribution function. The low standard deviation implies that data tend to be very close to the mean, while high standard deviation indicates that the data are “spread out” over a large range of values. This calculation is described by the following formula:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}$$

where, $\sigma$ is denoted as standard deviation and $\mu$ is the arithmetic mean of the values $x_i$, and $N= Total number of observations.$

1.6.8.3 Coefficient of Variation (CV)

The coefficient of variation (CV) is a ratio of standard deviation $\sigma$ to mean ($\mu$). Mathematically, CV can be represented as

$$CV = \frac{\sigma}{\mu}$$

It is a measure of dispersion of probability distribution. In other words, it is a relative measure, whereas standard deviation is an absolute measure. It is a useful measure for comparing the degree of variation from one data series to another, even if the means are drastically different from each other.

1.6.8.4 Chi-square Test

Chi-square is used to compare observed data with data we would expect to obtain according to a specific hypothesis. A chi-square test (i.e. $\chi^2$ test) is any statistical hypothesis test in which the test statistic has a chi-square distribution when the null hypothesis is true. Chi-square is the sum of the squared difference between observed value ($o$) and the expected value ($e$) divided by the expected value ($e$) in all possible categories. The formula for calculating the Chi-square test is given below:

$$\chi^2 = \frac{(o-e)^2}{e}$$, where $o = observed value & e = expected value$

1.6.8.5 Pearson’s Correlation Coefficient (r)

Correlation coefficient is usually denoted by $r$, which measures the strength of linear relationship between two variables. In other words, it is a technique for investigating the relationship between two variables and measures the association between them. The formula can be given as follows:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}$$

where, $n$ is the number of pairs and $r$ denotes correlation coefficient.

Correlation coefficient may be positive or negative. Positive correlation indicates that upward relationship between two variables i.e. both the variables
increase or decreases together, whereas negative correlation indicates that as one variable increases with the decrease in values of other variable and vice versa.

1.6.8.6 Spearman’s Rank Correlation

It is a non-parametric method to measure the correlation of two variables (say X and Y). If precise value of the variable (e.g. X) is not available, the data may be ranked in order of size or importance.

In case of Unique Rank

The formula for the calculation of rank correlation in case of having unique rank of the variable (X & Y) is as follows:

\[ Rs = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \]

Where, \( D \) is the difference between rank for the variables X & Y and \( N \) is the number of institutions to be ranked.

In Case of Repeated Rank

With reference to ranking the Indian universities, if it is found the same score for more than one university, common ranks are given to the repeated items. The common rank is the average of the ranks, which these items would have assumed if they were slightly different from each others and the nest item will get the rank next to the rank already assumed. This is called the case of repeated rank problem. As a result, following adjustment is required to the existing above formula of Spearman’s rank correlation formula. This part of the adjustment i.e. \( \frac{m(m^2 - 1)}{12} \) has to be added for every repeated item in the series to \( \sum D^2 \). Therefore, the new formula will be as follows:

\[ Rs = 1 - \frac{6 \sum D^2 + \frac{m(m^2 - 1)}{12}}{N(N^2 - 1)} \]

Where, \( m \) = number of times an item is repeated and \( N \) = total number of items in a series.

1.7 Definitional Analysis

1.7.1 Indian Universities

According to Wikipedia, “A University is an institution of higher education and research, which grants academic degree in a variety of subjects”. The original Latin word “universitas” was used at the time of emergence of urban town life and medieval guilds, to describe specialized “associations of students and teachers with collective legal rights usually guaranteed by charters issued by princes, prelates, or the towns in which they were located”. Indian higher education system is one of the largest educational systems in the world. Indian university comprises of deemed university, state university, institute of national importance, Open University and private university. The Indian universities are affiliated and
Chapter 1: Introduction

1.7.2 The World Wide Web

The World Wide Web, popularly known as Web, is a system of interlinked hypertext documents enabling users to browse and search for information through a web browser using standards like HTTP and TCP/IP via the Internet. “The W3 world view is of documents referring to each other by links. For its likeness to a spider's construction, this world is called the Web”.


The Web is one of the services that run on the Internet. It is a collection of interconnected contents or resources connected through hyperlinks and URLs.

1.7.3 Hyperlinks

A hyperlink is nothing but a reference to a document that the users can directly follow, or that is followed automatically. A hyperlink is a connection between any two documents. A hyperlink has an anchor within a document from which the hyperlink can be followed; that document is known as its source document. The target of a hyperlink is the document that the hyperlink leads to.

1.7.4 Webometrics

The science of webometrics made an attempt to measure the World Wide Web to know the types of hyperlinks, structure of the Web and its usage patterns. According to Björneborn and Ingwersen (2004), the definition of webometrics is "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches.”

Alternatively, webometrics is "the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study" (Thelwall, 2009), which emphasizes a small subset of relatively applied methods for use in the wider social sciences.

1.7.5 Link Topology

According to Wiktionary, link topology is the study of the link structure of the World Wide Web (simply Web). A link topology describes the link structure among nodes. These nodes may be the websites or webpages.

1.7.6 Web Impact Factors

It is the link-based website impact measure known as Web Impact Factor (WIF). It is the quantitative tool for evaluating and ranking websites, top-level domain and sub-level domain. Ingwersen (1998) developed the concept of measuring the average link frequencies as a quantitative tool to evaluate the website. Noruzi (2007) defined Web Impact Factors (WIFs) as $WIF = \frac{A}{D}$.
Where, \( A \) implies total links retrieved from particular Top-Level Domain (TLD) or Sub-Level Domain (SLD) or domain as indexed by search engine and \( D \) implies total number of webpages which are indexed by the search engines. The value of WIF is calculated for a particular website for a given point of time to determine the relative standing of website in a particular field.

### 1.7.7 Webometric Tools and Techniques

Webometric study is mainly based on the website statistics. Generally, commercial search engines, Web Crawlers and the Internet Archive for historical information concerning webpages are common phenomenon. The website statistics can be collected generally using search engine. The popular search engines for webometric study are AltaVista, Yahoo!, Hotbot, Alltheweb, and Google (partially) etc. These search engines, besides their regular activities have special keywords through which webometric data collection is performed.

Besides, there are a few personal Web crawlers such as SocSciBot 3.0 and LexiURL which are designed to collect the website statistics and also there are associated tools for data analysis. Popular visualization tools e.g. SocSciBot tools, NetDraw (Network visualization programme), Ucinet 6.0, Pajek are employed to generate the visualization of hyperlinks pattern. Some of the common bibliometrics and statistical techniques are like citation analysis, calculation of impact factor, bibliographic coupling, co-link analysis, measurement of degree of collaboration, performance analysis, correlation analysis, rank correlation, growth of hyperlinks using growth curves etc.

### 1.8 Profile of the Study Area

The study made an attempt to cover all the Indian universities, which are recognized and affiliated by the University Grants Commission (UGC). UGC is a central government body responsible for co-ordination and maintenance of standards, for releasing of grants for higher educational institutes in India. There are six types of universities in India: Central Universities (CU), State Universities (SU), Institutes of National Importance (INP), Deemed Universities (DU), Open Universities and Private Universities (PU).

Table 1.2—Category of Indian universities till April 2011

<table>
<thead>
<tr>
<th>Category</th>
<th>Indian Universities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Having websites</td>
</tr>
<tr>
<td>Deemed Universities</td>
<td>129</td>
</tr>
<tr>
<td>Central Universities</td>
<td>41</td>
</tr>
<tr>
<td>State Universities</td>
<td>262</td>
</tr>
<tr>
<td>Institutes of National Importance</td>
<td>39</td>
</tr>
<tr>
<td>Open Universities</td>
<td>13</td>
</tr>
<tr>
<td>Private Universities</td>
<td>81</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>565</strong></td>
</tr>
</tbody>
</table>

Source: UGC Website [http://www.ugc.ac.in](http://www.ugc.ac.in) (Accessed dated May 5, 2011)
It is found that 98.95% of Indian universities are having websites. The present study covers state universities, deemed universities, central universities, institute of national importance and open universities.

All the states of India may be grouped into four main regions like: Eastern, Western, Southern and Northern for making the study more significant and fruitful. It is worthwhile to mention that the private universities have not been taken into consideration for the present study. Following map indicated clearly the number of universities having websites under each state in India. These universities exclude private universities and open universities.

Figure 1.2—Map of India locating states with number of universities, April 2011

The map indicates that Tamil Nadu is having the highest number (56) of universities followed by Maharashtra (43) and Andhra Pradesh (41).

Following table clearly mentioned the distribution of various categories of universities under different 28 states and 3 union territories in India.
**Table 1.3—State-wise distribution of universities in India in 2011**

<table>
<thead>
<tr>
<th>S.N</th>
<th>Code</th>
<th>States</th>
<th>SU</th>
<th>DU</th>
<th>CU</th>
<th>INP</th>
<th>OU</th>
<th>PU</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AP</td>
<td>Andhra Pradesh</td>
<td>30</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>AR</td>
<td>Arunachal Pradesh</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>AS</td>
<td>Assam</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>BR</td>
<td>Bihar</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>CH</td>
<td>NCT of Chandigarh</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CT</td>
<td>Chhattisgarh</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>DL</td>
<td>NCT of Delhi</td>
<td>5</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>GA</td>
<td>Goa</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>GJ</td>
<td>Gujarat</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>10</td>
<td>HR</td>
<td>Haryana</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>HP</td>
<td>Himachal Pradesh</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>JK</td>
<td>Jammu &amp; Kashmir</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>JH</td>
<td>Jharkhand</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>KA</td>
<td>Karnataka</td>
<td>18</td>
<td>15</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>15</td>
<td>KL</td>
<td>Kerala</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>MH</td>
<td>Maharashtra</td>
<td>19</td>
<td>21</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>17</td>
<td>MN</td>
<td>Manipur</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>ML</td>
<td>Meghalaya</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>MZ</td>
<td>Mizoram</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>MP</td>
<td>Madhya Pradesh</td>
<td>15</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>21</td>
<td>NL</td>
<td>Nagaland</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>OR</td>
<td>Orissa</td>
<td>12</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>PY</td>
<td>NCT of Pondicherry</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>PB</td>
<td>Punjab</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>RJ</td>
<td>Rajasthan</td>
<td>14</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>16</td>
<td>40</td>
</tr>
<tr>
<td>26</td>
<td>SK</td>
<td>Sikkim</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>27</td>
<td>TN</td>
<td>Tamil Nadu</td>
<td>23</td>
<td>29</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>28</td>
<td>TR</td>
<td>Tripura</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>UP</td>
<td>Uttar Pradesh</td>
<td>21</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>13</td>
<td>52</td>
</tr>
<tr>
<td>30</td>
<td>UK</td>
<td>Uttarakhand</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>31</td>
<td>WB</td>
<td>West Bengal</td>
<td>20</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>266</td>
<td>130</td>
<td>42</td>
<td>39</td>
<td>13</td>
<td>81</td>
<td>571</td>
</tr>
</tbody>
</table>

_Source (i) For SU, DU & PU, please see UGC website: http://www.ugc.ac.in.
(ii) For OU, please see http://www.dec.ac.in: (iii) University Handbook, 2010
(iii) NCT => National Capital Territory_

It is clear from the above table that there are 571 universities in India; 130 deemed universities; 42 central universities including 16 new central universities; 266 state universities; 39 institute of national importance and 13 open universities and 81 private universities in India.

1 ISO 3166-2:IN codes are defined states and union territories. Each code consists of two parts, separated by hyphen. The first part is IN and second part is two letters.
1.9 Scope and Limitation

The study has been taken up to focus on Indian universities, which are recognized and affiliated by UGC, India to critically review the web performance using webometric analysis. The study covers all the universities under the following categories like Central Universities (CU), Deemed Universities (DU), Institute of National Importance (INP) and State Universities (SU).

Table 1.4—Various categories of university

<table>
<thead>
<tr>
<th>Category</th>
<th>Universities having websites</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU</td>
<td>41</td>
</tr>
<tr>
<td>DU</td>
<td>129</td>
</tr>
<tr>
<td>INP</td>
<td>39</td>
</tr>
<tr>
<td>SU</td>
<td>262</td>
</tr>
<tr>
<td>Total</td>
<td>471</td>
</tr>
</tbody>
</table>

An attempt has been taken up to study the websites of Indian universities in order to analyze the different types of links and calculation of Web Impact Factors. The present study also covers 13 Open Universities (OU) as recognized by the Distance Education Council (DEC), India. These 13 exclusively open universities are already covered under the above categories (IGNOU, New Delhi under CU and remaining 12 under SU). So, OU category has not been shown here separately.

But, the study does not include Private Universities (PU) as such but a preliminary study had been conducted to know the existence of their domain names.

1.10 Conspectus

The investigation made for this research work has been organized into eight chapters followed by glossary, bibliography and annexure.

Chapter 1 introduces the problems that lead to carry out this research work. The general introduction and background of the study help to formulate the objectives. It deals with research design and research approach undertaken throughout the study. It gives the guidelines to keep alive the focus of the study in the right direction. It also describes the hypotheses and methodology in detail to achieve the objectives. Scope of research and possible limitations of the study are mentioned.

Chapter 2 covers comprehensive review of literature including both Indian and global literature in order to put the current research into context. The review covers the sub-domain of Webometrics like webpage content analysis, web usage analysis, web link structure analysis, web usage analysis, web technology analysis, web citation analysis and web impact factor. The actual scenario of the trends in web link analysis and web impact factors for webometrics study can be viewed through this chapter. It helps to find out the critical gap in Webometric research in the context of academic domain especially Indian universities.
Chapter 3 deals with historical development of the concept webometrics and its various phases of hyperlink analysis using webometric tools and techniques. The part of content of this chapter is published in Information Studies (Jalal, Biswas & Mukhopadhyay, 2009b). Besides, the definitional aspects of various concepts are explained in this chapter.

Chapter 4 deals with the Indian higher educational system and specifically university system. Besides, the domain name system (DNS) is explained to shed a light upon to know comprehensively on various TLDs and SLDs as exists today. It deals with measuring the web presence and calculation of various web impact factors e.g. WIF-simple, WIF-self links, WIF-external links for Indian universities. Jalal, Biswas and Mukhopadhyay (2008) have shown the hyperlink analysis of universities in West Bengal in a pilot study and web presence of selected Asian countries (Jalal, Biswas & Mukhopadhyay, 2009a).

Chapter 5 deals with various aspects of webometric ranking of Indian universities. More specifically, UGC recognized and NAAC accredited universities have been covered for the comparison of ranking through various ranking methodologies like NAAC, Times Ranking, 4International Colleges & Universities (ICU) ranking. RWWU ranks etc. Jalal, Biswas and Mukhopadhyay (2009c) made an attempt to summarize the various ranking approaches- their indicators along with corresponding weights, comparison of ranks of Southern universities covered under NAAC accredited with Cumulative Grade Point Average (CGPA) method with other approaches e.g. RWWU, 4ICU Rank and WIF-inlinks.

Chapter 6 deals with link analysis and link topology. The chapter forms the basis of co-link technology, formation of micro-topology and its signification. The study of identification of characteristic of link patterns among selected Indian universities has been addressed. The co-link maps of Indian universities and the reasons for co-links are talked about. Jalal, Biswas, and Mukhopadhyay (2010) made an attempt to find out the link topology among the National Institute of Technologies (NITs) and central universities in India as case study.

Chapter 7 deals with the methodology applied to classifying the hyperlinks, grouping them to form a broad categorization. The motivation for hyperlinks of Indian academic web space especially for university domain system has been taken an attempt. The classification of links and reasons associated for the association of universities are discussed.

Chapter 8 deals with findings of the study. An attempt has made to put down couple of suggestions, which may be effective for the improvement of the performance of websites of the universities under study. Besides, initiatives have taken to identify some of the areas, where further research can be extended by the researchers. It outlines the significant contributions which have made the current study to enhance the existing knowledge base and highlights some strong areas of research. Further, future direction of research in the field of webometrics, especially hyperlink analysis in academic web space has been pointed out.

While writing the thesis, American Psychological Association (APA) style is being followed as reference style (APA style, n.d.). At the end of each chapter,
references have been listed in an alphabetical order and the bibliography, which includes chapter references and other consulted documents, is given at the end of the thesis using APA style as far as possible. For the case of online document citation, date of retrieval is included because "the source material may change over time" (apastyle.org)

1.11 Summary

The present chapter provides the research design for the concerned study. The layout of research project has been clearly mentioned with the background study, research motivation for undertaking the project, objectives, hypotheses, scope and limitations. In order to achieve the objectives, detailed methodologies have been laid out. Some statistical tools, which are mentioned in this chapter, are used later on for data analysis and data interpretation. The research study basically concentrates on the hyperlink analysis of the websites of Indian universities in the context of scholarly academic web space.

The next chapter deals with the ‘Review of Literature’, an instrumental source to establish the research study into proper perspective.

References


Chapter 1: Introduction


Rousseau, R. (2001). Evolution in time of the number of hits in keyword searches on the Internet during one year, with special attention to the use of the
word euro. Proceedings of the 8th International Conference on Scientometrics and Informetrics (pp.619-627). Sydney: ISSI.


Thelwall, M., & Harris, G. (2004). Do the web sites of higher rated scholars have significantly more online impact? Journal of the American Society for Information Science and Technology, 55(2), 149-159.


