MEASURING QUALITY OF WEB SITE NAVIGATION

[1]G. SREEDHAR, [2]Dr.A.A. CHARI, [3]Dr.V.V.VENKATA RAMANA

[1]Sr. Lecturer, Dept. of Computer Science, Rashtriya Sanskrit Vidyapeetha, (Deemed University), Tirupati
[2]Professor, Dept. of OR & SQC, Rayalaseema University, Kurnool
[3]Scientist-E, National Informatics Centre, Hyderabad

ABSTRACT

The primary goal of this paper is to identify the elements involved in quality assessment of website structure. Always quality of website structure is based on its navigability, average number of clicks and structural complexity. The web tool PowerMapper is used to establish the sitemap for the website, path length metric is used to evaluate average number of clicks to get desired web page and web site structural complexity is determined with cyclomatic complexity. The quality of web site structure is measured in 10-point scale and scale is formulated by considering various levels of website tree structure. The status of web site structure is determined in 10-point scale and the value suggests the improvement of the site structure.

Key words: *sitemap, path length, navigability, cyclomatic complexity, web site structure*

1. INTRODUCTION

A website is a collection of web pages containing text, images, audio and video etc. Today, Web is not only an information resource but also it is becoming an automated tool in various applications. Due to the increasing popularity of web, one can be very cautious in designing the Website. Poor and careless web design [1] leads to hard ship to public utility and does not serve the purpose. Navigation places crucial role in the design of website structure because it determines the path to be traveled to reach a required web page. Normally a website structure resembles tree like structure starting from home page as root. In designing a website tree structure one has to concentrate in breadth and width sizes of tree. The home page of website designed in such a way that it should not be too much crowded with links and also it should not be too much empty. It was stated that number of links on each page should not exceed 20 links [4] and average number of clicks should not be more than 4 clicks to get a required page. Keeping view of this the developer must be very much careful in designing the website structure.

2. RELATED RESEARCH WORK

The quality of website can be assessed mainly in functionality and usability. The quality of website design [5], [6] is assessed using various qualitative measures of functionality. The notion of usability is a key factor to interact a website. The efficiency of usability is depended on website structure. The structure of website [7] should be in such a way that the user can easily interact website without any formal training. An effective web design [8] is one that makes it easier on users to navigate through the different pages on the site. The website structure [9] is represented by directed graph where each node represents a web page and edge represents link to that page. It is already investigated that web link structure can also be used for page ranking [10] and web page classification [11]. These works stress on the navigational relationship among web pages. Our present work focus on discovering the various elements involved in website structure and thus we can concentrate on these elements to improve quality of a website structure.

3. METHODOLOGY

The procedure for the quality assessment of website structure involves three modules: establishment of sitemap, computing path length metric and evaluating structural complexity of website. All these modules are incorporated in a web program.

3.1. Establishment of sitemap

Every website must have sitemap to know the organization of web pages in the website structure. The sitemap shows all web pages in a hierarchical tree with home page as root of the tree. A web tool PowerMapper is used in the procedure to construct a sitemap for the website. It selects URL address of website and generates the tree structure for all web pages of website. In this process only markup files (html, asp, php, xml, etc.,) are considered and remaining components like graphic files script files, etc., are not included because these files do not have any significance in website structure. The sitemap of a website may be organized into various levels depending on its design. Some websites have one or two levels and some may have three or more levels. A snapshot of Aligarh Muslim University's website sitemap is shown in figure 3

3.2. Evaluating Path length metric

A path length is used to find average number of clicks per page. The path length of the tree is the sum of the depths of all nodes in the tree. It can be computed as a weighted sum, weighting each level with its number of nodes or each node by its level using equation (1). The average no. of clicks is computed using equation (2). The width of a tree is the size of its largest level and the height of a tree is the length of its longest root path.

Path length = $\Sigma\, l_i.m_i$ (1)

where l_i is level number i, m_i is number of nodes at level i.

Avg no. of clicks = path length/n (2)

where n is the number of nodes in the tree An example tree is shown in figure 1.

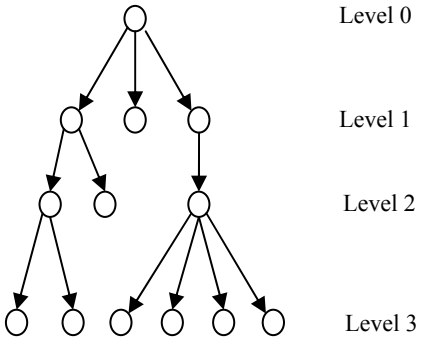

Figure 1. A tree with 3 levels

Path length = 0*1 + 1*3 + 2*3 + 3*6 = 27
Avg. no. of clicks = 27/13 = 2.07

3.3. Structural complexity

The structural complexity of website is determined with Mc. Cab's cyclomatic complexity metric [2]. This metric is used to know navigation path for a desired web page. The cyclomatic complexity metric is derived in graph theory as follows. A tree graph is constructed with home page as root. The tree consists of various sub trees and leaf nodes. An example tree is shown in figure 2.

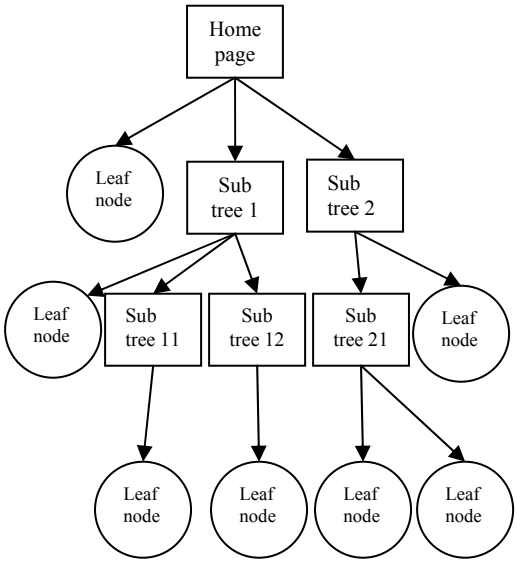

Figure 2. tree graph for a website

A tree graph is constructed for a website by considering various hyperlinks in the website. Each sub tree of the graph represents a web page which has further hyperlinks to the next web pages and leaf node represents a web page which does not have further links to any web pages. In tree graph, at each level all web pages that do not have further links are represented with one leaf node at that level and a sub tree at each level consists of links to the web pages to the next level. The cyclomatic complexity is calculated using equation (3) and it should not exceed 10 according to Mc. Cab.

web_site_complexity = (e-n+d+1)/n (3)

where e = number of web page links
n = number of nodes in the graph
d = number of leaf nodes in the graph

Fig.3 Sitemap of Aligarh Muslim University's website

4. EVALUATION

The websites of more than 50 Indian universities are considered in the evaluation process. The web program accepts URL address of each university's website and generates sitemap using PowerMapper web tool. The sitemap consists of all web pages of the website and displays in hierarchical tree structure in various levels. The average number of clicks needed to access a web page is computed with path length metric using equation (1) and (2). The cyclomatic complexity value is computed for the web site structure using equation (3). The quality of website structure for each university is evaluated in 10-point scale. The 10-point scale value for each university is based on organization of web pages in sitemap structure, cyclomatic complexity of website and average number of clicks. The computation procedure for 10-point scale is shown in table 1 and table 2. As an example, the sitemap of Assam university website is shown in the figure 4. The evaluation of navigability value of Assam university website is shown in table 3. The quality of some universities' website structure values are shown in table 4.

Sno	Quality parameter	10-point scale value evaluation

1	No. of links at web page of sitemap tree	If (no. of links in home page = total number of web pages) and (no. of links in home page <=20) then k1 = 10 else
		If (no. of links in a page between 10 and 20) then k1 = 10 else
		If (no. of links in a page = 9 or 21) then k1 = 9 else
		If (no. of links in a page = 8 or 22) then k1 = 8 else
		If (no. of links in a page = 7 or 23) then k1 = 7 else
		If (no. of links in a page = 6 or 24) then k1 = 6 else
		If (no. of links in a page = 5 or 25) then k1 = 5 else
		If (no. of links in a page = 4 or 26) then k1 = 4 else
		If (no. of links in a page = 3 or 27) then k1 = 3 else
		If (no. of links in a page = 2 or 28) then k1 = 2 else
		If (no. of links in a page = 1 or 29) then k1 = 1 else k1 = 0

2	Cyclomatic complexity	If (cyclomatic complexity <= 1) then k2 = 10 else
		If (cyclomatic complexity <= 2) then k2 = 9 else
		If (cyclomatic complexity <= 3) then k2 = 8 else
		If (cyclomatic complexity <= 4) then k2 = 7 else
		If (cyclomatic complexity <= 5) then k2 = 6 else
		If (cyclomatic complexity <= 6) then k2 = 5 else
		If (cyclomatic complexity <= 7) then k2 = 4 else
		If (cyclomatic complexity <= 8) then k2 = 3 else
		If (cyclomatic complexity <= 9) then k2 = 2 else
		If (Cyclomatic complexity <= 10) then k2 = 1 else k2 = 0

value= average(k1,k2)		

Table 1: Evaluating 10-point scale value for sitemap structure and cyclomatic complexity

Sno	Quality parameter index	10-point scale value
1	Average number of clicks per web page	If average number of clicks <= 2.5 then value = value + 0.75 else If average number of clicks < = 4 then value = value + 0.5 else If average number of clicks <=5 then value = value + 0.25

Table 2: Evaluating 10-point scale value for average number of clicks and broken link error index

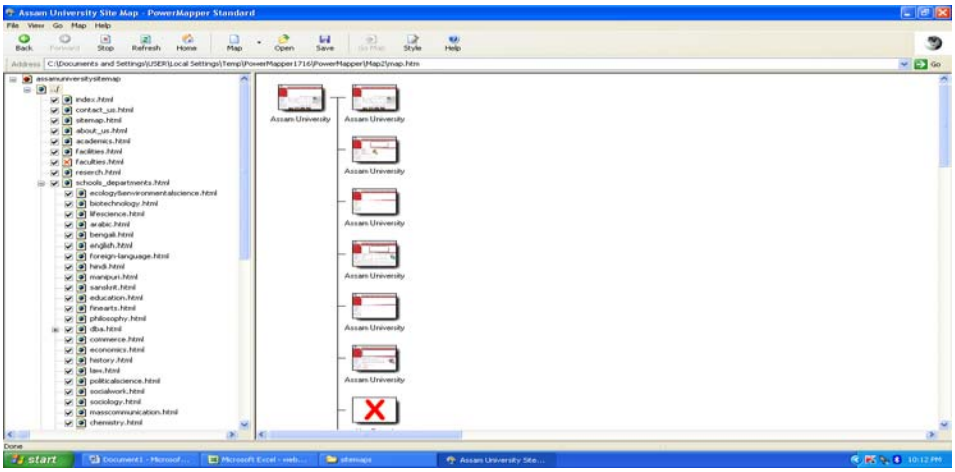

Figure 3: Assam University's website sitemap

University Name: Assam University Path Length = 149 Average number of clicks =1.886076				
Level No	Sub tree in web site structure		No. of web pages in sub tree	10-Point Scale Value
1	1		28	1
2	1		28	1
	2		3	3
3	1		1	1
	2		1	1
	3		4	4
	4		11	10
4	1		1	1
	2		1	1
				2.555556
Cyclomatic complexity	(e-n+d+1)/n	(256-52+26+1)/52 = 4.442308	10-point scale value = 7	
Path Length =149		Avg number of clicks = 1.886076		
10-point scale value for Assam University's website structure = 2.555556+7+0.25 = 4.7				

Table 3: Assam University's website structure quality evaluation

Sno	University Name	V_1	V_2	Avg(v1,v2)	P	10-point scale value	Remarks
1	Aligarh Muslim University	4.21875	8	6.109375	0	6.1	Needs improvement
2	Assam University	2.555556	7	4.777778	0.25	5	Poor design
3	Dravidian University	5.666667	3	4.333333	0.25	4.6	Very Poor design
4	Bharatiar University	2.55	6	4.275	0.25	4.5	Very Poor design
5	Bharatidasan University	4.380952	8	6.190476	0.5	6.7	Needs improvement
6	Bhavnagar University	2.833333	8	5.416667	0.75	6.2	Needs improvement
7	Sri Venkateswara University	3.058824	9	6.029412	0.75	6.8	Needs improvement
8	Rashtriya Sanskrit Vidyapeetha	2.365854	6	4.182927	0.75	4.9	Very poor design
9	Madurai Kamaraj University	4.346154	5	4.673077	0.75	5.4	Poor design
10	Guru Jambeshwar University	3.352941	6	4.676471	0.75	5.4	Poor design

Table 5: Quality of various universities' website structures

V1= sitemap evaluation value v2= cyclomatic complexity value P= path length metric

5. CONCLUSION

The main objective of the work is to measure quality of website navigation and suggesting the importance of navigability of website. In this paper an attempt is made to find key components like sitemap evaluation, average number of clicks per page, and cyclomatic complexity which greatly influences quality of website navigation. The web developer must concentrate these components in selection and design of website. We can further extend this work to identify other factors to improve the quality of website structure which is a part of the ideology of TQM which emphasizes the continuous improvement of design aspect and promote excellence of web design.

REFERENCES

[1] Folders, Vincent And Michel Will, "Web Pages that suck: Learn Good Design by looking at Bad Design", SanFrancisco, CA, SYBEX.

[2] Yanlong Zhang, Hong Zhu and Sue Greenwood, "Website Complexity Metrics for Measuring Navigability", Proceedings of the fourth conference on quality software (QSIC'04), 0-7695-2207-6/04, IEEE.

[3] www.websiteoptimiztion.com

[4] Benjamin Yen, Paul Jen-Hwa Hu, May Wang, "Toward and analytical approach for effective Website design: A frame work for modeling, evaluation and enhancement", Electronic Commerce Research and Applications 6 (2007), 159-170.

[5] G. Sreedhar and A.A. Chari, "An experimental Study to Identify Qualitative Measures for Website Design", Global Journal of Computer Science and Technology, University of Wisconsin, USA, September, 2009, pp.12.

[6] G. Sreedhar, A.A. Chari and V.V.Venkata Ramana, "Evaluating Qualitative Measures for Effective Website Design", International Journal on Computer Science and Engineering, vol.02, No.01S, 2010, pp.61-68.

[7] Benbunan-Fich, R, "Using Protocol Analysis to evaluate the usability of a Commercial Website, Information and Management, 39, 151-163, 2001.

[8] Ali Azad, "Elements of Effective Web Page Design", Global Competitiveness, January, 2001.

[9] Zheng Chen, Shengping Liu, Liu Wenyin, Geguang Pu, Wei-Ying Ma, "Building a Web Thesaurus form Web Link Structure", SIGIR 2003, July 28 – August 1, 2003, Toronto, Canada. 10.

[10] L. Page, S. Brin, R. Motwani and T. Winograd, "The PageRank Citation Ranking: Bring Order to the Web", Technical Report, Stanford University, 1998.

[11] E. Glover, K. Tsioutisiouliklis, S. Lawrence, D.Pennock, G. Flake, "Using Web Structure for Classifying and Describing Web Pages", in Proceedings of WWW2002, Hawaii, May 2002.

AUTHOR PROFILE:

Mr. G. Sreedhar is pursuing Ph.D in Computer Science from S.K.U, Anantapur. He received M.Phil (Computer Science) from Alagappa University, Karaikudi, T.N, India and M.C.A from S.K. University, Anantapur, A.P, India in 2005 and 1998 respectively. He is working as a Sr. Lecturer in Dept. of Computer Science at Rashtriya Sanskrit Vidyapeetha, Tirupati since 2001. He has over 10 years of teaching and research experience in the field of Computer Science.

Dr. A.A.Chari received Ph.D, M.Phil and M.Sc in Statistics from S.V University, Tirupati, A.P, India in 1986, 1977 and 1976 respectively. He is working as a Professor in Dept. of Operations Research and Statistical Quality Control, Rayalaseema University, Kurnool, A.P, India. He has 32 years of teaching and research experience in the field of statistics. He has large number of published papers in national, international and highly reputed IEEE journals. He guided number of research scholars in the field Reliability, Optimization & SQC, Computer Science and other related fields.

Dr. V.V. Venkata Ramana received Ph.D. from Sri Krishnaevaraya University, Anantapur, A.P, India in 1997, while M.Phil and M.Sc in OR & SQC from S.V. University, Tirupati, A.P, India, 1986 and 1983 respectively. He is working as Scientist – E in National Informatics centre, Hyderabad. He has a number of published papers in National, International Journals. He guided a number of research scholars in the field of OR & SQC, Computer Science and other related fields.