Quantitative Evaluation of Web Page Attributes for Improving the Quality of Website

Deepthi. M
Software Engineering
AMC Engineering College
Bangalore, India

Shalini. S
Computer Science
AMC Engineering College
Bangalore, India

Abstract—**World Wide Web is getting more common now a days and the data in internet is growing day by day. For all our day to day use we depend on internet. So the need for relevant data in the websites is becoming necessary. To ensure this, evaluating website quality is an essential. There are few ways to analyze and evaluate the quality of the website in quantitative form. Various guidelines have been proposed for the same. But still it is not clear how to implement them. Since metrics are crucial source of information in decision making, web metrics are used to estimate the quality of the web sites. Properly designed metrics provide an objective appraisal of a product or process. So it is necessary to continuously evaluate the websites and then make improvements over those evaluations.**

This paper focuses on the quantitative analysis of the web page attributes and how these attributes contribute to the website quality. For this work, nine quantitative web measures are identified for each websites. These metrics are computed using an automated Metrics Analyzer tool developed in java programming language. We Collect evaluation data from different categories of Pixel awards 2010 and 2011 and classify the websites into good or bad on the basis of selected web page attributes. In order to analyze the results, we have used logistic regression technique on the dataset. The analysis result provides an empirical foundation for website design guidelines and also suggests which metrics can be most important for evaluation.

Keywords—*Metrics; Web metrics; Website; Website design; Web page; Website quality;*

I. INTRODUCTION

Internet and World Wide Web has been grown rapidly in their scope and extent of use affecting all aspects of our lives. Developing a website does not end with putting necessary information, media and software. After designing of the website, we have to test and evaluate the website for customer satisfaction. Evaluation gives good quality to our website. We can attract more users if our website is of good quality. There is currently much debate about what constitutes good web site design [1][2]. Many detailed usability guidelines have been developed for both general user interfaces and for web page design [3][4]. However, designers have historically experienced difficulties following design guidelines. Guidelines are often stated at such a high level that it is unclear how to operationalize them. Furthermore, there is no general agreement about which web design guidelines are correct. The problem with expert guidelines is that they are often based upon assumptions that are difficult to substantiate. The suggestions they provide are frequently general in nature, requiring the web designer to correctly interpret how to implement them. Even factual studies may provide misleading interpretations of their results. Because of this evaluators of software applications and websites strive to make sure that their software is up to the quality standards relative to others.

Characteristics of every website can be classified into wide range of types. For example, there are several characteristics that are related to performance (e.g., websites processing speed, and the speed of executing transactions etc,) characteristics related to usability (e.g., number of distribution of objects, colours, etc), and many other characteristics.

The quality of a website is represented in terms of the amount of information presented in one page, the relevancy of information provided, the user-friendliness of website and many other properties.

Website attributes can be also classified into two major types: internal and external attributes. Internal attributes are those characteristics that we can directly measure while external attributes are those we like to measure or know. Internal attributes can be used to help us know the external one. However, usually the relation is not simple and direct between internal and external attributes. For example, the number of links in a web page is a size internal metric that we can directly measure. Such internal metric has relation with several external metrics such as size, complexity, quality, etc.

Web site engineering metrics are mainly derived from software metrics, hyper media and Human computer interaction. The intersection of all the three metrics will give the website engineering metrics [5]. For proper designing of websites, website engineer need to understand the subset of metrics on which the goodness of website design metrics depends. These attributes can be used to predict the usability of websites.

In this paper we present some attributes related to web page measures and calculate the values of web page attributes with the help of an automated tool. We applied logistic regression technique over the metric estimations calculated for the collection of websites and find the subset of metrics out of nine metrics to capture the criteria of goodness of websites. We report the results of empirical analyses of the page-level elements on a large collection of expert reviewed web sites.

The rest of the paper is organized as follows. Section II provides the related research. Section III provides an overview of web page measures. In Section IV describes our research methodology and we end paper with our conclusions and future work in Section V.

II. RELATED WORK

A lot of existing work has been done on evaluating web page quality, but most quantitative methods for evaluating web sites focus on statistical analysis of usage patterns in server[6][7][8]. The analysis based on such data is quite uncertain since web server logs provide incomplete traces of user behaviour, and because timing estimates may be skewed by network latencies. The above work focuses more on navigation history; Server logs are problematic because they only track unique navigational events (e.g., do not capture use of back button) and thus are hard to understand because of caching. Another method for evaluating web pages of user interest automatically investigates various factors in a user's browsing behaviour such as number of scrolls, form input, search text etc [9].

Other approaches were inspection-based that rely on assessing static HTML according to a number of pre-determined guidelines, such as whether all graphics contain ALT attributes that can be read by screen readers [10]. For example, WebSAT (Web Static Analyzer Tool) is used to check the accessibility issues (i.e., support for users with disabilities), forms use, download speed, maintainability, navigation and readability of Web pages. There are many other techniques that compare quantitative web page attributes such as the number of links or graphics to thresholds [11]. However, there are no clear thresholds established for a wider class of quantitative Web page measures.

Simulation has also been used for web site quality evaluation. For example, a simulation approach has been developed for generating navigation paths for a site based on content similarity among pages, server log data, and linking structure [6]. The simulation models hypothetical users who are traversing the site from described start pages, making use of information "scent" (i.e., common keywords between the user's goal and linked pages content) to make decisions related to navigation. The approach does not consider the impact of various web page attributes, such as the amount of text or layout of links.

The quality of a website can be defined in terms of functional as well as non-functional properties. K. M. Khan [12] has derived the non-functional attributes such as reliability, usability, efficiency, security and assessed them. The work done in [12] adopts a Goal-Question-Metric (GQM) approach to derive quality metrics. It defines the goals that are needed to be measured, then it develops the questions derived from goals that are required to determine if the goals are fulfilled, and finally, their measurements are the answers of the questions which are known as metrics.

III. BACKGROUND

Our study aims at determining the effect of various webpage attributes on the goodness of the web pages. The foremost task is to select the web page metrics for evaluation. Classification of web page measures include:-

Page composition metrics:-The example of this metrics are No. of words, Body Text words, Words in page title, Total number of links etc.

Page formatting metrics: - They comprise of Font size, Font style, Screen coverage etc.

Overall page quality or assessment metrics: - Example of these metrics is Information quality, Image quality, Link Quality etc.

From a list of 42 web page attributes associated with effective design and usability [13], we developed an automated tool to compute the nine metrics that we focus on in this study (see Table I).

TABLE I. METRICS SELECTED FOR THIS WORK

	Metrics	Description
1	Word Count	Total Words on a page
2	Body Text Words	Words that are body vs. display text
3	Page Size	Total bytes of the page and images
4	Table Count	Number of tables present on the webpage
5	Graphics Count	Total images on a page
6	Division Count	Total divisions on a page
7	List Count	Lists on a page
8	Number of Links	Links on a page
9	Page title length	Words in page title

The description of the attributes calculated by the tool is:

1) Word Count

Total number of words on a page is taken. This attribute is calculated by counting total number of words on the page. Special characters such as & / are also considered as words.

2) Body text words

This metric counts the number of words in the body Vs display text. In this, we calculate the words that are part of body and the words that are part of display text separately. The words can be calculated by simply counting the number of words falling in body.

3) Page size

It refers to the total size of the web page and can be found in properties option of the web page.

4) Table Count

This metric gives the number of tables used in making a web page.

5) Graphic Count

This refers to the total number of images on a page. And can be calculated by counting the total number of images present on the page. It has been analyzed that usable and good quality pages contain more images which contribute to the larger page size.

6) Division Count

This metric can be calculated by analyzing the number of divisions in a web page.

7) List Count

This metric can be calculated by counting total number of ordered and unordered list present on a web page.

8) Link Count

These are the total number of links on a web page and can be calculated by counting the number of links present on the web page.

9) Page title length

These refer to the word in the page title and can be calculated by counting the total number of words in the page title.

IV. RESEARCH METHODOLOGY

A. Data Collection

This study computes quantitative web page attributes (e.g., number of words, list, and division) from web pages that was evaluated for 2010-2011 pixel Awards. The Pixel Awards are the website award, annually honouring compelling sites that have shown excellence in web design and development. The websites placed in 24 categories are judged on the basis of creative and technical blend of impeccable graphic design, artistry, technological expertise, and a powerful, stimulating user experience [14]. We used A1 Website Download which is configured to crawl 0-level i.e. homepages and 1-level pages from each site; thus, we collected homepages and 1-level web pages for each website.

There are 2 awards given in each category, one is chosen by judges as winner, and another is People's Champ Winner. We have considered the winner websites in all the categories as good and all the other nominee websites as bad. Thus, we have 31 good classified and 59 bad classified web pages in dataset of year 2010, 33 good and 53 bad classified websites in year 2011.Table II shows the number of good and bad classified websites in 2 year data of Pixel awards 2010-2011.

TABLE II. CATEGORISATION OF WEBSITES

Category	Websites 2010	Websites 2011
Good	31	33
Bad	59	53

B. Automated Website Evaluation

We have developed a Metrics Analyzer, a JAVA based automated metrics analyzing tool that calculates 9 metrics to determine which of these attributes are correlated with the goodness and usability of the web page. Fig.1 shows the workflow of the methodology.
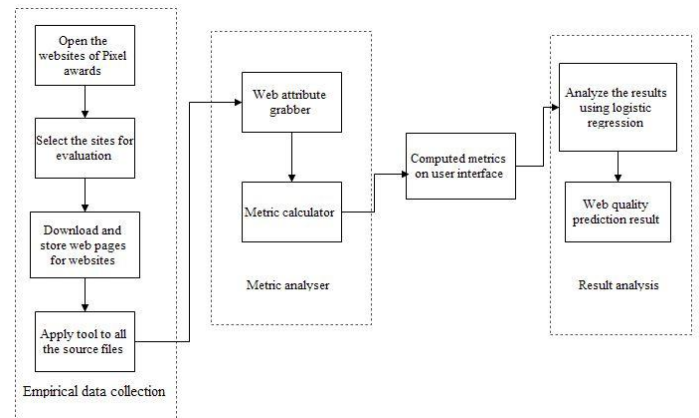

Fig. 1. Automated Website Evaluation Workflow

The idea is to automatically collect information about the web pages that gives an idea of the flavour of the document. This method demonstrates a way of rating web pages automatically for information content. From the above mentioned tool we can calculate different web attributes. We can also save the result for future use.

C. Data Analysis

In this section we describe the methodology used to analyze the metrics data computed for web sites. We use Logistic Regression technique to analyze the data.

Logistic Regression:-Logistic regression is the common technique that is widely used to analyze data. It is used to predict dependent variable from a set of independent variables. In our study the dependent variable is good/bad and the independent variables are web metrics. Logistic regression is of two types (1) Univariate logistic regression and (2) Multivariate logistic regression.

Univariate logistic regression is a statistical method that formulates a mathematical model depicting relationship between the dependent variable and each independent variable.

Multivariate logistic regression is used to construct a prediction model for goodness of design of web sites.

The multivariate logistic regression formula can be defined as follows:-

$$\text{Prob}(X_1, X_2 \ldots X_n) = \frac{e^{(A_0 + A_1 X_1 + \ldots + A_n X_n)}}{1 + e^{(A_0 + A_1 X_1 + \ldots + A_n X_n)}}$$

In LR, two stepwise selection methods, forward selection and backward elimination can be used [15]. Stepwise variable entry examines the variable that is selected one at a time for entry at each step. This is a forward stepwise procedure. The backward elimination method includes all independent

variables in the model. Variables are deleted one at a time from the model until stopping criteria are fulfilled.

D. Descriptive Statistics

Descriptive statistics provide the simple summaries about the sample dataset and quantitatively describe its main features. Each table [III-IV] show minimum, maximum, mean and standard deviation for all metrics considered in this study.

TABLE III. DESCRIPTIVE STATISTICS OF DATASET 2010

Metrics	Min	Max	Mean	Standard deviation
Table count	0	570	14.58	65.74
Link count	0	10619	809.32	1518.97
Graphic count	0	1806	168.18	304.88
List count	0	735	94.56	154.08
Page title length	0	333	62.04	85.23
Division count	0	4581	574.20	921.21
Word count	1	48964	6042.39	9167.66
Body text words	0	117720	7005.63	14855.95
Page size	1	1817	276.53	410.78

TABLE IV. DESCRIPTIVE STATISTICS OF DATASET 2011

Metrics	Min	Max	Mean	Standard deviation
Table count	0	539	9.38	58.23
Link count	1	3576	632.03	804.71
Graphic count	0	1450	180.88	271.85
List count	0	862	84.59	134.26
Page title length	0	323	56.50	61.39
Division count	1	7254	538.76	938.17
Word count	1	28205	5202.37	5932.24
Body text words	0	27882	4868.35	5798.27
Page size	2	1108	252.43	282.80

E. Evaluation of the Model

We used logistic regression to discriminate good from bad pages. This technique is suitable where we have one dependent variable and many independent variables. We have one dependent variable named good or bad and independent variables are the selected metrics of pixel awards. Table V shows the web page prediction of logistic regression for the 2 years data.

TABLE V. WEBSITE PREDICTION OF LOGISTIC REGRESSION

Parameter	2010	2011
Number of good websites correctly predicted	24	20
Number of bad websites correctly predicted	54	46

There are 5 variables selected in dataset of 2010 where the independent ones are Link Count, List Count, Word Count and Page Size. 6 significant variables which include Link Count,

List Count, Page title Length, Word Count and Body Text Words as independent ones are selected in dataset of 2011.

Link count, List count, Word count is attributed as significant metric in the two datasets which makes it important to be considered by the designers. It shows that number of links, lists and words should be appropriate to enhance the quality of a website.

V. CONCLUSION AND FUTURE WORK

The importance of web metrics has dramatically increased due to extremely fast growth in Internet technology. The web metrics is directly related to the purpose of the good website design. Improving the quality of web content is imperative for further web development, because good measuring sticks are needed for web sites. We have quantitatively evaluated the relationship of web metrics and quality of website using logistic regression analysis.

The evaluation data is based on pixel awards 2010-2011. Pixel Awards are the website award that honours the websites showing great design and development. However, the judging criteria for such websites are so broad, that it is not possible for the designers to understand the criteria and improve their websites. Thus, this work can be applied to provide the designers with the important metrics that must be considered for the website design. The websites which are found to be bad will need extra attention and can be improved further.

We plan to carry our research for all the levels of web pages in the website and to propose some more web page metrics that are more effective for the website design. This will simplify the work of website engineer and improve the quality of the websites.

REFERENCES

[1] Jakob Nielsen. Designing Web Usability: The Practice of Simplicity. New Riders Publishing, Indianapolis, IN,2000.

[2] Julie Ratner, Eric M. Grose, and Chris Forsythe. Characterization and assessment of HTML style guides. In Proceedings of ACM CHI 96Conference on Human Factors in Computing Systems, volume 2, pages 115–116, 1996.

[3] Tim Comber. Building usable web pages: An HCI perspective.In Roger Debreceny and Allan Ellis, editors, Proceedings of the First Australian World Wide Web Conference AusWeb'95, pages 119–124. Norsearch, Ballina, 1995.

[4] Patrick J. Lynch and Sarah Horton. Web Style Guide: Basic Design Principles for Creating Web Sites. Yale University Press, 1999.

[5] K.K Aggarwal, Yogesh Singh, "Software Engineering", 3rd edition, New Age Publication, India , 2008.

[6] E. H. Chi, P. Pirolli and J. Pitkow, "The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a web site," In Proceedings of ACM CHI 00 Conference on Conference on Human Factors in Computing Systems, 2000.

[7] M. C. Drott, "Using web server logs to improve site design," In ACM 16th International Conference on Systems Documentation, pages 43–50, 1998.

[8] R. Fuller and J. J. De Graaff, "Measuring user motivation from server log files," In Proceedings of the Human Factors and the Web 2 Conference, 1996.

[9]   G. Velayathan and S. Yamada, "Behavior-Based Web Page Evaluation," In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT Workshops),2006.

[10] J. Scholtz, S. Laskowski and L. Downey, "Developing usability tools and techniques for designing and testing web sites," In Proceedings of the 4th Conference on Human Factors & the Web, 1998.

[11] H. Thimbleby, "Gentler: A tool for systematic web authoring," In International Journal of Human-Computer Studies, 47, 139-168, 1997.

[12] K. M. Khan, "Assessing Quality of Web Based System," IEEE/ACS International Conference on Computer Systems and Applications, AICCSA, Page(s): 763 – 769, 2008.

[13] M. Y. Ivory, R. Sinha, and M. A. Hearst, "Preliminary findings on quantitative measures for distinguishing highly rated information-centric web pages," In Proceedings of the 6th Conference on Human Factors and the Web, 2000.

[14] Pixel Awards | Web Awards Competition (2006). Available from http://www.pixelawards.com.

[15] Mayers, J.H and Forgy E.W. (1963). The Development of numerical credit evaluation systems. Journal of the American Statistical Association, Vol.58 Issue 303 (Sept) pp 799–806.