Web Metrics

Akanksha, Akansha Rakheja and Ajay Singh

ABSTRACT:

Web Engineering is the application of systematic, disciplined and quantifiable approaches to development, operation, and maintenance of Web-based applications. Web page metrics is one of the key elements in measuring various attributes of web site. Metrics gives the concrete values to the attributes of web sites which may be used to compare different web pages. In this paper, a quantitative evaluation strategy to assess the quality of Web sites and applications (WebApps) is discussed. We give an overview of the WebQEM (Web Quality Evaluation Method) and its supporting tool. Also, we aim to investigate 15 metrics proposed by various researchers. We present the findings of quantitative analysis of web page attributes and how these attributes are calculated. The result of this paper can be used in quantitative studies in web site designing.

Keywords: Web Page Quality, Quantitative Evaluation, WebQEM, empirical validation, limitations.

1. INTRODUCTION

The science of webometrics (also cybermetrics) tries to measure the World Wide Web to get knowledge about the number and types of hyperlinks, structure of the World Wide Web and usage patterns. 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 bibliometricand informetric approaches.[7]" A second definition of webometrics has also been introduced, "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, which emphasises the development of applied methods for use in the wider social sciences. The purpose of this alternative definition was to help publicise appropriate methods outside of the information science discipline rather than to replace the original definition within information science.

Akanksha, Akansha Rakheja and Ajay Singh are with Information Technology (IT), Dronacharya College of Engineering, Gurgaon, India, Emails: akanksha_50@yahoo.com, akansha_aug23@yahoo.com, crazyajaysingh@gamil.com

A key element of any web site engineering process is metrics. Web metrics are used to better understand the attributes of the web page we create. But, most important, we use web metrics to assess the quality of the web engineered product or the process to build it. Since metrics are crucial source of information for decision making, a large number of web metrics have been proposed in the last decade to compare the structural quality of a web page. The Web is playing a central role in diverse application domains such as business, education, industry, and entertainment. As a consequence, there are increasing concerns about the ways in which WebApps are developed and the degree of quality delivered. Thus, there are compelling reasons for a systematic and disciplined use of engineering methods and tools for developing and evaluating Web sites and applications [1]. We need sound evaluation methods for obtaining reliable information about the product's quality. These methods should identify which attributes and characteristics should be used to obtain meaningful indicators for assuring specific evaluation goals given a user viewpoint. In the Web Quality Evaluation Method (QEM) [2, 3] and some aspects of its supporting tool, WebQEM_Tool [4], we show that, by using the methodology for assessment purposes, we can give recommendations both by controlling quality requirements in new Web development projects and by evaluating requirements in operational phases.

Also, as the number of web metrics available in the literature is large, it become tedious process to understand the computation of these metrics and draw conclusion and inference from them. Thus, properly defined metrics is used for predictions in various phases of web development process. For proper designing of websites, we need to understand the subset of metrics on which the goodness of website design metrics depends. In this paper we present some attributes related to web page metrics and calculate the values of web attributes with the help of an automated tool. This tool is developed in JSP and calculates about 15 web page metrics with great accuracy.

To meet the above objective following steps are taken:

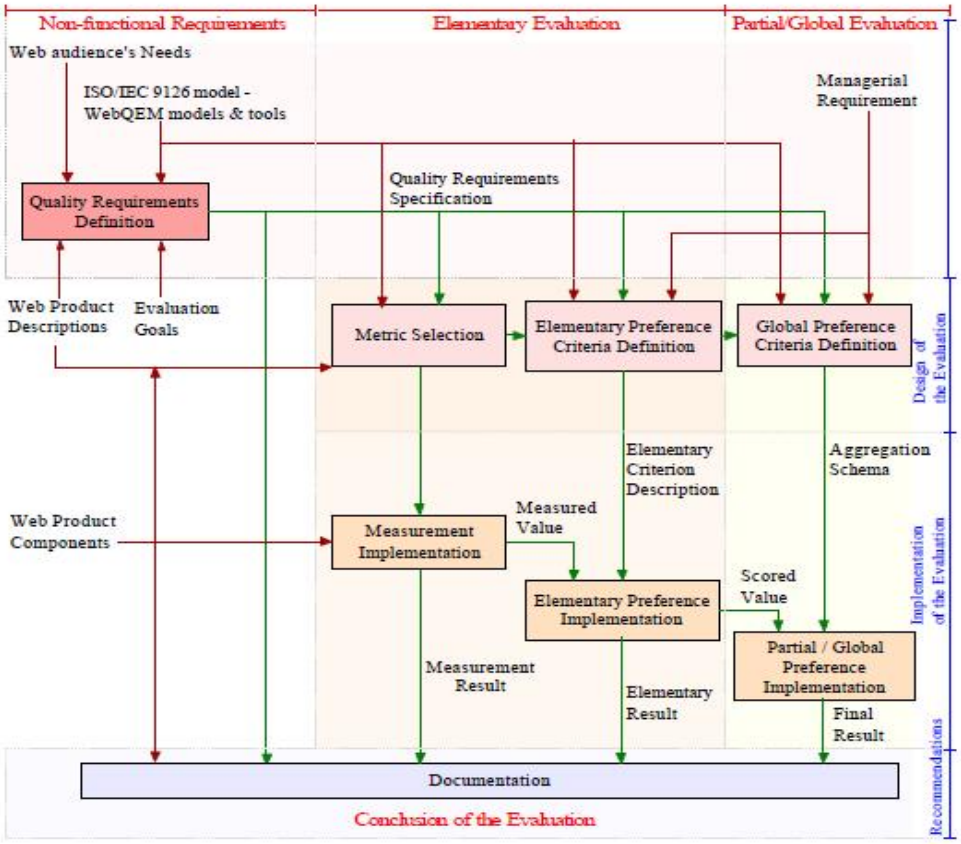
- Set of 15 metrics is first identified and their values are computed for 514 different web sites (2007-2010) webby awards data.
- The interpretations are drawn to find the subset of attributes which are related to goodness of website design. Further, these attributes can be used to assess the data into good sites and bad sites.

2. EVALUATION PROCESS IN THE WebQEM METHODOLOGY

The WebQEM process steps are grouped in the following four major technical phases:

1. Quality Requirements Definition and Specification;

2. Elementary Evaluation (both Design and Implementation stages);
3. Global Evaluation (both Design and Implementation stages);
4. Conclusion of the Evaluation (regarding Recommendations).



2.1 Quality Requirements Definition and Specification.

In this phase, evaluators must clarify the evaluation goals and the intended user viewpoint. They should select a quality model, for instance, the ISO-prescribed characteristics in addition to attributes customized to the Web domain. The relative importance of these components should be identified considering the WebApps audience and the extent of the coverage required. Regarding the user profile, at least three abstract evaluation views of quality may be defined, i.e., visitors, developers and managers views. For example, the visitor category can be decomposed in general and expert visitor subcategories. Thus, taking into accounts the domain and product descriptions, the agreed goals, and the selected user view (i.e., the explicit and implicit user needs), characteristics, sub characteristics and attributes should be specified in a quality requirement tree. In the end of this phase, a quality requirement specification document is produced.

2.2 The Elementary Evaluation.

In this phase, two major stages are defined as depicted in Fig. 1: The design and the implementation of the elementary evaluation. For each measurable attribute Ai from the requirement tree, we can associate a variable Xi, which will take a numerical value from a direct or indirect metric. However, the value of this metric will not represent the level of satisfaction of this elementary requirement at all. For that reason, it is necessary to define an elementary criterion function that will result afterwards in an elementary indicator or preference value. In the implementation stage, the selected metrics are applied to the Web application as shown in Fig. 1. Some values can be measured observationally, while others can be obtained automatically by using computerized tools.

2.3 The Global Evaluation Phase.

Again, two major stages are defined: The design and the implementation of the partial/global quality evaluation. In the design stage, aggregation criteria and a scoring model should be selected. The goal of quantitative aggregation and scoring models is to make the evaluation process well structured, accurate, and comprehensible
by evaluators. There are at least two types of models: for example those based on linear additive scoring models [5], and those based on nonlinear multi-criteria scoring models [6] where different attributes and characteristics relationships can be designed. In both cases, the relative importance of indicators is considered by means of weights. For example, if our procedure is based on a linear additive scoring model

the aggregation and computing of partial/global indicators or preferences (P/GP), considering relatives weights (W) is based on the following formula:

$P/GP = (W_1 EP_1 + W_2 EP_2 + ... + W_m EP_m)$;

such that if the elementary preference (EP) is in the unitary interval range the following is held:

2.4 The Conclusion of the Evaluation.

In this phase, the documentation of Web product components, the specification of quality requirements, metrics, criteria, elementary and final results are recorded. In the end, the strengths and weaknesses of the assessed product with regard to established goals and user viewpoint can be analyzed and understood by requesters and evaluators. Recommendations can be suggested and justified.

2.5 Automating the Process using WebQEM_Tool.

The evaluation and comparison processes require both methodological and technological support. We have developed a Web-based tool [16] that supports the administration of evaluation projects. It allows editing and relating non-functional requirements. For instance, in our E-commerce case study, more than ninety attributes have intervened [15]. Then, by means of the automatic or manual edition of elementary indicators, WebQEM_Tool allows to aggregate the elements in order to yield a schema and calculate a global quality indicator for each site. This enables evaluators to assess and compare the quality of Web products. The WebQEM_Tool relies on a Web-based hyper document model that supports traceability of evaluation aspects. The results of an evaluation are shown through linked pages with textual, tabular and graphical information, which are dynamically generated from tables stored in the data layer.

3. EMPIRICAL VALIDATION METHOD

3.1 Description of Web Page Metrics

Although, several researchers proposed many metrics for web page, out of those we identified only 15 metrics for our study. There are 42 web page metrics and classification of those is given below:-

- 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 are Information quality, Image quality, Link Quality etc.

The description of the parameters used in this study is given below:-

1) Number of words: 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 metrics counts the number of words in the body Vs display text (i.e. Headers). In this, we calculate the words that are part of body and the words that are part of display text that is header separately. The words can be calculated by simply counting the number of words falling in body and number of words falling in header.

3) Number of links: 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.

4) Embedded links: Links embedded in text on a page. These are the links embedded in the running text on the web page.

5) Wrapped links: Links that spans in multiple lines. These are the links which take more than one lines and can be calculated by counting the number of links that spans in multiple lines.

6) Within page links: These are the links to other area of the same page. This can be calculated by counting the number of links that links to other area of the same page. Example in some sites have top bottom.

7) Number of !'s :Exclamations points on a page can calculated by counting total number of ! marks on a page.

8) Page title length: These refer to the words in the page title and can be calculated by counting the total no of words in the page title.

9) Number of graphics: These refer to the total number of images on a page. And can be calculated by counting the total number of images present on the page.

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

11) Number of list: This metrics can be calculated by counting total number of ordered and unordered list present on a web page.

12) Number of tables: This metrics gives the answer of the question .How many number of tables is used in making a web page?

13) Frames: This metrics can be calculated by analyzing whether a web page contains frames or not.

14) Text emphasis: This metric can be calculated by analyzing the web page and counting the total number of words which are in bold, italics and capital.

3.2 Tool Description

To automate the study of web page metrics we develop a tool for calculating 15 web page attributes. We use JSP for this purpose. JSP technology is one of the most powerful, easy to use and fundamental tools in a Web site developer's toolbox. JSP technology combines HTML and XML with Java servlet (server application extension) and JavaBeans technologies to create a highly productive environment for developing and deploying reliable, interactive, high performance, platform-independent web sites. JSP technology facilitates creation of dynamic content on the server. It is part of the Java platform's integrated solution for server-side programming which provides a portable alternative to other server-side technologies, such as CGI. JSP technology integrates numerous Java application technologies, such as Java servlet, JavaBeans, JDBC, and Enterprise JavaBeans. It also separates information presentation from application logic and fosters a reusable-component model of programming [9]. From the above mentioned tool we can calculate different web attributes. We can select all the attributes or select some of the above list. We can also save the result for further use.

3.3 Data Analysis

In this section we describe the methodology used to analyze the metrics data computed for 514 web sites. We use Logistic Regression to analyze the data. Logistic Regression:-LR 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. LR is of two types (1) Univariate LR and (2) Multivariate LR. Univariate LR is a statistical method that formulates a mathematical model depicting relationship between the dependent variable and each independent variable. Multivariate LR is used to construct a prediction model for goodness of design of web sites. The multivariate LR formula can be defined as follows:-

$$\text{Prob}(X1, X2 \ldots Xn) = e^{(A0+A1X1+\ldots+AnXn)} / 1 + e^{(A0+A1X1+\ldots+AnXn)}$$

In LR, two stepwise selection methods, forward selection and backward elimination can be used [10]. 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.

3.4 Result Analysis

We employed statistical techniques to describe the nature of the data. Table [1] presented in the following subsection show min, max, mean and SD for all metrics considered in this study.

Descriptive Statistics				
Metrics	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Words in page title	0	17	6.72	3.86
Body text words	1	34523	2942.06	4757.21
Total words on page	85	35443	3343.40	4861.28
Total number of Links	0	844	117.01	178.52
Number of!'s	2	417	43.04	67.36
Number of graphics	0	230	25.97	38.78
Page Size	930	429870	44581.33	63574.64
Total embedded links	0	239	11.80	30.63
Number of Lists	0	197	11.16	23.60
Number of frames	0	0	0	0
Number of tables	0	66	2.15	7.58

We used Logistic Regression to discriminate good from bad pages. This technique is suitable where we have one dependent variable and many independent variables. As in our study, we have one dependent variable named good/bad and independent variables are the whole web metrics of webby awards. We built predictive model for identifying good pages. This model predicts the goodness of the website design based on the subset of metrics which we get from the Logistic regression technique. The description of the model is based on webby awards of the year 2010. In this model we get many metrics which leads to a bad design like body text words, number of!'s, Page size, Number of tables and within page links. If we have higher values of these metrics we will get a bad design. For predicting model we use backward elimination method of Logistic Regression.

4. LIMITATIONS OF WEB METRICS

The managers admitted that there is a multitude of data that could potentially be used to improve the website. However, overload of information is a concern. For web metrics to be useful the managers try to relate the metrics to the objectives for the website. The objectives are stated as, for example, customer self help, customer service or sales. Also, the lack of standards in the industry was mentioned as a concern. The topic is so new that standards or benchmarks are not available, and even the nomenclature is not standardized. Details and processes still tend to be confusing. Also, the people involved with website operations need to be trained in the use of web metrics, to ensure that they understand what kind of metrics are needed and what kind of queries managers need to extract meaningful data. The managers expressed a need for more background information about the users. For example, managers would like to know whether a visitor is an IT professional, a student or a CEO. This information can be collected through surveys, and assists in improving the website. One example of the usefulness of survey data concerns the homepage. Data revealed that of 200,000 homepage visitors, less than 10% actually read the feature story. The managers were trying to figure out whether the visitors believed the story, whether it made an impression, and whether visitors found it insightful. Survey data supplied that type of information, allowing the company to determine what made a compelling story. Web metrics were then used to evaluate the success of the implemented changes. For the future, managers envision a greater use of surveys to investigate the visitors' thoughts and impressions, e.g., how they rate a story, the quality of the technical information, and the writing itself. In summary, new problems that arose regarding the use of web metrics included the large amount of available data and metrics, the problem of selecting meaningful measures,

understanding of data collection limitations, and the need to understand underlying assumptions. The following section explains how those limitations were addressed.

4.1 Overcoming the limitations

There are multiple ways to address the limitations that were discussed in the previous section.

- *Determination of needs:* Since standards are not established yet and the application of web metrics is relatively new, it is critical to clearly define information needs before selecting the appropriate web metrics to ensure their usefulness for website optimization. The interviewed managers expressed the lesson learned: that it is critical to set clear objectives before web metrics can be useful for site optimization. Different measures are needed for different site optimization efforts. For example, one of the managers explained that it would be very useful to know the most bookmarked pages to determine redesign efforts or clean out content. Also, it is important to know which pages are never visited, which is hard to derive from the data, as they tend to reflect what people did look at, not what they did not look at.

- *Multiple sources:* Managers consistently mentioned that to really understand online customer behaviour, the data should be complemented with other data such as demographics, internal (sales and customer contact) data, or attitudinal information. For example, usability lab testing is conducted with actual customers, monitoring how they fulfil particular tasks on the website. This feedback is mostly solicited in the early stages of the web design development when web designers evaluate layouts, terminology or desired features. As opposed to metrics, the lab tests explore the 'why' of user behaviour. Data from lab usability tests and the data complement each other: The metrics might identify a problem but cannot explain the 'why'; lab tests will add valuable information. Likewise, improvements to the site, made based on qualitative feedback, can be tested with analytics. Since the various sources are imperfect (e.g. self-selection problems with surveys and the artificial environment of usability tests) managers consider an integrated approach the most appropriate one. Combining data sources results in a steep learning curve for web managers as the individual applications such as web metrics or usability tests are relatively new topics, and their combination has not been standardized. In addition to using multiple sources of information, aggregation of data across systems and combining transactional and web systems is considered to be really important

5. CONCLUSION

In this article, we have discussed the main ideas behind a quantitative evaluation methodology for Web sites and applications. WebQEM can be employed in assessing and comparing quality requirements in the operative phase of Web sites and applications as well as in early phases of Web development projects. By using the methodology, we can discover either absent attributes, absent sub characteristics, or requirements poorly implemented. If we get justifiable and traceable indicators, i.e., global, partial, and elemental ones, we have an objective basis to give recommendations for improvements. WebQEM can be used to assess diverse application domains according to different user views and evaluation goals. It should be noticed that the definition and the specification of quality requirements are essential activities in the evaluation process. In the empirical evaluation, the research is to capture quality of web sites. As E-business is emerging and websites are not just medium for communication, but they are also a product of providing services. Therefore imparting quality, security and reliability to web sites are very important. We empirically validate the relationship of web metrics and quality of websites using logistic regression technique. Any site that is submitted for the award is examined by three judges on six criteria. It is unclear and unknown how the experts rates the website but hope we present a way towards a new methodology for creating empirically justified recommendation for designing a good web sites. In this paper we present the attributes which, if have higher value can lead to a bad design. From the above attributes we also find profile of good pages.

REFERENCES

[1] Murugesan, S.; Deshpande, Y.; Hansen, S.; Ginige, A., 1999, "*Web Engineering: A New Discipline for Development of Web-based Systems*", Proceed. of ICSE (Int'l Conference on Software Engineering) Workshop on Web Engineering, Los Angeles, US, pp. 1-9.

[2]Olsina L., Apr-2000, "*Quantitative Methodology for Evaluation and Comparison of Web Site Quality*", Doctoral Thesis, (in Spanish), Ciencias Exactas School, UNLP, La Plata, Argentina.

[3]Olsina, L.; Lafuente, G.J.; Rossi, G., 2000, "*E-commerce Site Evaluation: a Case Study*", LNCS 1875 of Springer, 1st International Conference on Electronic Commerce and Web Technologies, EC-Web 2000, London, UK, pp. 239-252.

[4]Olsina L., Papa, M.F., Souto, M.E., Rossi, G.; 2001, *"Providing Automated Support for the Web Quality Evaluation Methodology"*, Proceed. of the Fourth Workshop on Web Engineering, at the 10th International WWW Conference, Hong Kong, pp. 1-11.

[5] *Gilb, T., 1976, Software Metrics, Chartwell-Bratt, Cambridge, MA.*

[6] Dujmovic, J.J., 1996, *"A Method for Evaluation and Selection of Complex Hardware and Software Systems",* The 22nd Int'l Conference for the Resource Management and Performance Evaluation of Enterprise CS. CMG 96 Proceedings, Vol. 1, pp. 368-378.

[7] http://en.wikipedia.org/wiki/Webometrics

[8] 13.Olsina, L., Godoy, D; Lafuente, G.J; Rossi, G.; 1999, *"Assessing the Quality of Academic Websites: a Case Study "*, In: New Review of Hypermedia and Multimedia (NRHM) Journal, Taylor Graham Publishers, UK/USA Vol. 5, pp. 81-103.

[9] Olsina L., Apr-2000, "*Quantitative Methodology for Evaluation and Comparison of Web Site Quality*", Doctoral Thesis, (in Spanish), Ciencias Exactas School, UNLP, La Plata, Argentina.

[10] Olsina, L.; Lafuente, G.J.; Rossi, G., 2000, "*E-commerce Site Evaluation: a Case Study*", LNCS 1875 of Springer, 1st International Conference on Electronic Commerce and Web Technologies, EC-Web 2000, London, UK, pp. 239-252.