

The Automatic Evaluation of Website Metrics and State

Izzat Alsmadi, Yarmouk University, Jordan

ABSTRACT

This paper focuses on studying website structural and related metrics that can be used as indicators of the complexity of the website and predict maintainability requirements. The second goal of the study is to evaluate possible correlations between structural metrics and popularity (particularly in-links) metrics. Examples of some of the structural metrics evaluated in this paper include: size, complexity, and speed of page loading. While results showed that structural metrics are not good indicators of websites' popularity, they may influence indirectly the popularity through their impact on the performance or the usability of those websites. A method is developed to evaluate the state of the website automatically and evaluate any change in that state. The study points to certain requirements that educational or higher institutes' websites should have. Those websites should combine somewhat conflicting requirements of: high performance, particularly web page loading and speed of transaction, reliability; current, correct and up to-date information, navigability, visibility and popularity where website information should be visible internally and externally and should be easily indexed and searched for.

Keywords: Computer Science, Education, Information Retrieval, Inlinks, Navigability, Outlinks, Testing, Web Metrics, Website Complexity

INTRODUCTION

Evaluators of software applications and websites strive to make sure that their software is up to the quality standards relative to others. They used metric tools and methods to be able to get the software characteristics and compare them with known standards. In order to make the approach plausible, those attributes should be gathered automatically through tools. Web applications have some characteristics that make their maintenance expensive. This includes: heterogeneity, speed of evolution, and dynamic code generation (Ghosheh, Black, & Qaddour, 2008).

The characteristics of every software or website can be classified into a wide range of types or categories. For example, there are several characteristics that are related to performance such as websites processing speed, and the speed of executing transactions. There are also several attributes related to reliability such as number of errors in pages, in scripts, the percentage of time that the website is running or available, etc. In some cases, some of those characteristics may not be fully measured unless the website is operational. This may include operational quality results from four characteristics: effectiveness, productivity, safety, and satisfaction. Such attributes can only be measured during the operating environments of the software.

DOI: 10.4018/jwlts.2010100101

Software or website attributes can be also classified into two major types: internal and external attributes. Internal attributes are those characteristics that can directly be measured while external attributes are goals that cannot be measured directly. Internal attributes are indicators to the external ones. For example, the number of links in a web page is the internal size metric is one that can be directly measured. Such internal metric has relations with several external metrics such as website size, complexity, quality, etc.

An ISO standard (ISO/IEC-9126-1, 2001) defines 6 major quality characteristics. Those are: functionality, reliability, usability, efficiency, maintainability, and portability. This quality model describes each quality characteristic of a software product by further sub-characteristics that elaborate each characteristic. The Goal Question Metric (GQM) approach provides a three-step framework for a quality model: (1) list the major goals of the empirical study; (2) derive from each goal the questions that must be answered to determine if the goals have been met; (3) decide what must be measured in order to be able to answer the questions adequately (i.e. the definition of the metric) (Basili & Weiss, 1984).

Websites' evaluation can be distinguished from typical software evaluation. In typical software applications, for example, traffic and usage metrics may not be as much as important as they are in a website evaluation. Website users' behavior can be studied through eye tracking, log files, studying users' sessions, etc. However, one of the challenges in this evaluation is that there are some user related attributes that are hard to automatically collect through metrics such as loyalty, satisfaction, understandability, etc.

Usability, ease of use, user-friendliness, and ease of learning is a set of related metrics to indicate the ability of the user to communicate easily with the website and understand its functionalities with minimum effort. It can be measured through different ways such as the time it takes a user to perform a task or different

tasks using the website (usually in comparison with other websites).

Productivity metrics try to evaluate the amount of resources consumed through using the website. Examples of those metrics include: time, memory, and some other attributes that relate it to the level of effectiveness (i.e. number of tasks achieved in a certain time) gained by the user. Indicators of productivity can include: time needed to complete a task, the parts of the goals reached within a unit of time, cost-effectiveness of tasks, and the amount of time it takes the user to perform certain actions.

The focus of this paper is on studying the web structural that can be gathered automatically and their relation with traffic metrics. The research hypothesis is to study the amount of correlation (if there is) between structural metrics such as links, documents, forms, etc with traffic metrics (mainly back or inlinks). The original motivation to the research is that initial investigation of top websites in terms of popularity shows that they vary widely in terms of their structural metrics and nature. One possible explanation is that the domain of the website may impact its size nature. Accordingly, in this study, several websites from several business domains are selected to study if there are any consistencies in the structure and popularity metrics based on the business domain.

There are many business domains such as universities or higher institutions that depend heavily on their websites as their outlet to their students and the world. Those websites include information going in and out and are extensively used for announcements, lectures, students' registration, courses, assignments, faculty members' pages, etc. The structural, navigability and popularity metrics studied in this research are somewhat at odds where focusing on building a website with heavy structure of images, docs, multimedia, etc may impact its navigability and popularity. On the other hand a website with many missing, incorrect or out of-date links may affect its visibility or navigability attributes.

15 more pages are available in the full version of this document, which may be purchased using the "Purchase" button on the product's webpage:

www.irma-international.org/article/automatic-evaluation-website-metrics-state/52596/

Related Content

Towards Automated Specifications of Scenarios in Enhanced Learning Technology

A. Ruis Gavidia, M.A. Sicilia Urban, E. Garcia-Barriocanal, and G. Macarro Plazuelos (2008). *International Journal of Web-Based Learning and Teaching Technologies* (pp. 68-77).

www.irma-international.org/article/towards-automated-specifications-scenarios-enhanced/3001/

Preparing Online Instructors: Beyond Using the Technology

(2010). *Web-Based Education: Concepts, Methodologies, Tools and Applications* (3 Volumes) (pp. 277-292).

www.irma-international.org/chapter/preparing-online-instructors/41346/

Content Areas: Syllabus, Notes, Lesson Plans, and Documents

Lisa Dawley (2007). *The Tools for Successful Online Teaching* (pp. 24-49).

www.irma-international.org/chapter/content-areas-syllabus-notes-lesson/30411/

The eLogBook Framework: Sustaining Interaction, Collaboration, and Learning in Laboratory-oriented CoPs

Yassin Rekik, Denis Gillet, Sandy El Helou, and Christophe Salzmann (2007). *International Journal of Web-Based Learning and Teaching Technologies* (pp. 61-76).

www.irma-international.org/article/elogbook-framework-sustaining-interaction-collaboration/2988/

ICT, Knowledge Construction, and Evolution: Subject, Community, and Society

Antonio Cartelli (2006). *Teaching in the Knowledge Society: New Skills and Instruments for Teachers* (pp. 33-53).

www.irma-international.org/chapter/ict-knowledge-construction-evolution/30068/