Fuzz-Web: A Methodology Based on Fuzzy Logic for Assessing Web Sites

Rim Rekik and Ilhem Kallel

REGIM: REsearch Groups on Intelligent Machines, University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax, 3038, Tunisia
rim.rekik, ilhem.kallel@ieee.org

ISIMS: Higher Institute of Computer Science and Multimedia of Sfax, University of Sfax, Tunisia

Abstract: This paper presents a quality assessment methodology and model that measure the performance of dynamic websites. Called Fuzz-Web, a system that shows a comprehensive and natural manner of reasoning based on Multiple Criteria Decision Making process. We attempt so to use fuzzy logic as an intelligent technology, since the evaluation process is characterized by subjectivity and imprecision. Obviously, a phase of selecting appropriate evaluation criteria is necessary for the decision making process. Some tests realized on a set of Tunisian and foreign websites allow us to discuss the proposed reduced fuzzy method and then to validate the decision making result.

Keywords: website quality, website ranking, fuzzy logic, Fuzz-Web system.

I. Introduction

In the nineties, the web became accessible for the audience hence the approaches to assess quality have been successively elaborated. The first studies in that time were beginning to assess the quality of hypertext documents because over 50% of the hyperdocuments that had been published were unsuitable for real use according to Brown [1] who was dealing with this topic particularly and with maintainability in general in order to make solutions to survive over the years. The spread of Internet and World Wide Web applications have created new opportunities and new challenges for institutions and individuals who are either receiving or providing information on time. Web applications have been finally integrated in diverse domains such as e-commerce, education, entertainment, health, etc. One of the most important and popular branches of network applications is the development of websites which are considered as virtual showcases for the institutions. It recognizes the transition from static to dynamic, fade to attractive, incomprehensible to ergonomic, etc. and puts the user before the institution.

Today, with the large amount of data circulating on the World Wide Web, there is a growing need to help people figure out whether a website is well intentioned, truthful, or not. An even more ambitious goal is to facilitate searching valuable, accurate and appropriate information and improve websites interpretation to users on search engines.

The researchers investigated in additional ways of evaluation of online documentation in the literature that focused more directly on evaluating hypertext, hypermedia, and multimedia.

Therefore, the evaluation methodologies of quality of websites can help in ranking results for search engines. Particularly, the assessment of quality of web-products (i.e websites [15], web videos [2]) has become a recent, emerging and challenging issue.

The objectives and expected results for dynamic websites are largely dependent on the quality and information availability (content and container) and the online access efficiency. Thus, the quality of a website is a complex concept in practice and research; its measurement should be multidimensional and requires the use of techniques, such as Multi-Criteria Decision Making [3] and intelligent methods to pass up the problem of multiple conflicting criteria.

This paper reviewed the methods that aimed at evaluation of the quality of websites. We present, therefore, in section 2 the most important steps and existing challenges in this domain and its evolution from quantitative studies to qualitative ones. Section 3 explains and analyzes the proposed method. Section 4 discusses and compares the obtained results with literature ones of foreign websites. Finally, Section 5 resumes the paper and gives some perspectives about some future work.

II. Overview of website evaluation methods

The most relevant studies about evaluation methods are divided mainly into two categories: qualitative (which are informal) and quantitative (which are rather formal).

Generally, the qualitative evaluation methodologies are based on analyzing a list of features to clarify the advantages and disadvantages of an application and account them into percent. This approach is obviously attractive but only in the case of a simple problem. It’s not really very decisive in the case of complex problems. On the other side, quantitative evaluation can structure better the evaluation process relatively in a simple and accurate way. It provides global quantitative indicators which are used to find and justify an optimal decision. In fact, both of the methods are complementary and the works described below treats one or the other mode of assessment. They are based on two types, crisp and fuzzy. The crisp set approach is limited to assign numerical values either 0 or 1 score to targets. However, there
is an ambiguity to treat easily a problem in a natural way as well as the context for assessing websites, the quality result of the website can be good or not absolutely. While fuzzy sets extend crisp sets with the possibility to state variables in an interval scale.

Among the first quantitative surveys and Web domain-specific evaluation methods, in 1998, we address the study of Kirakowski et al. [4] that have worked specifically for evaluating the usability of websites since it was an important evaluation factor to improve and develop better a website. Always in the same direction and period, other studies have been conducted to learn about the usability and specifically the accuracy such as Keevil [5] work.

Quantitative strategies for websites quality evaluation have been given also by Olsina et al. [6], [7]. They have proposed a Web-site Quality Evaluation Method (QEM) tool customized to the assessment of the quality of academic websites. Particularly, in their latter academic case studies the purpose was to obtain a ranking for six internationally or regionally well recognized academic sites. They have considered for the evaluation the same high level characteristics as those prescribed in ISO 9126 standard in Table 1.

```
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Sub-characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>Understandability,</td>
</tr>
<tr>
<td></td>
<td>Learn ability,</td>
</tr>
<tr>
<td></td>
<td>Attractiveness,</td>
</tr>
<tr>
<td></td>
<td>Operability</td>
</tr>
<tr>
<td>Functionality</td>
<td>Suitability,</td>
</tr>
<tr>
<td></td>
<td>Accuracy,</td>
</tr>
<tr>
<td></td>
<td>Interoperability,</td>
</tr>
<tr>
<td></td>
<td>Security</td>
</tr>
<tr>
<td>Reliability</td>
<td>Maturity,</td>
</tr>
<tr>
<td></td>
<td>Fault tolerance,</td>
</tr>
<tr>
<td></td>
<td>Recoverability</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Time behavior,</td>
</tr>
<tr>
<td></td>
<td>Resource utilization</td>
</tr>
<tr>
<td>Portability</td>
<td>Adaptability,</td>
</tr>
<tr>
<td></td>
<td>Install ability,</td>
</tr>
<tr>
<td></td>
<td>Replace ability</td>
</tr>
<tr>
<td>Maintainability</td>
<td>Analyzability,</td>
</tr>
<tr>
<td></td>
<td>Changeability,</td>
</tr>
<tr>
<td></td>
<td>Stability</td>
</tr>
<tr>
<td></td>
<td>Testability</td>
</tr>
</tbody>
</table>
```

Table 1. ISO/IEC 9126 quality Model

However, maintainability and portability are neglected in their studies. Since those last factors have been largely addressed by researchers in the nineties [1] and have been resolved by most of them. Usually, they are measured objectively by automated procedures. Therefore, their measurement is relatively easy. Moreover, maintainability and portability are considered internal quality criteria don’t attract visitors like external quality criteria such accessibility, reliability, etc. Due to this, most of existing quality approaches of websites is based on the use of external quality criteria.

Another simple quantitative evaluation method that uses crisp values to compare between five economics faculties’ sites in northern Italy by Mich et al. [8]. They have developed a 2QC3Q model to represent weak points for each site in a radar diagram. The 2QC3Q has attributes and sub-attributes analogous to those of ISO 9126, except that it adds interrelated dimensions and considers how design affects a website. They have tested seven main dimensions, corresponding to seven loci which are: identity, content, services, location, management, usability and feasibility.

Going from crisp to fuzzy, a recent evaluation method proposed by Dominic et al. [9] determines the best Malaysia University website using website diagnostic tools and applying a New Hybrid Model (NHM) which is a process between Fuzzy Analytical Hierarchy Process (FAHP) model and Linear Weightage Model (LWM). They have considered five factors: time for site credibility, broken links for site reliability, content, response time and latency and the last criterion is performance.

A qualitative evaluation methodology based on fuzzy computing with words and characterized by its simplicity and quickness was proposed by Herrera-Videma et al. [10], [11], [12]. The evaluation dimensions concerned not only on relevancy, believability, completeness related to the quality of the information content but also concerned on site structure, broken links, multilingualism, navigation tools, user interface related to the design of the website. They have well selected criteria in order to allow users’ participation rather than many websites evaluation approaches (Olsina et al. [6], [7] work) that insist in user’s point of view but don’t supply enough means for it. Another important group decision approach for Hwang et al. [13] focuses on evaluating educational websites. It is an evaluation assistance method, takes account of user participation, able to select proper criteria using fuzzy logic techniques and deduce the ratings for each evaluated website. Indeed, the participants were not only students or internet users, but also experts’ domain or qualified Internet users. They have used else gray system theory to decide the criteria desired for the test of four well known websites in Taiwan. This domain is very earlier in Taiwan, we find else similar studies like the study of Lin [14], Huang et al. [15], etc.

The latter study [15] interested by evaluation of educational websites from the fuzzy subjective and objective perspective (FuzzSOP) which is an integrated decision model. Their concern was to gain a user’s opinion with metrics provided by Palmer [16] questionnaire with its eighteen items to assess for usability, design and performance. While the objective perspective is achieved automatically by a data mining technique and fuzzy clustering but it has a limitation to collect objective data.

We remark then researches nowadays try to combine both quantitative and qualitative methods to have more real and efficient system for evaluating websites.

**A. Synthetic overview**

To better understand the innovative features of the evaluation methods described above we propose to establish synthetic tables where we classify those methods according some inner and outer characteristics. In fact, tables below respectively Table 2 and Table 3 classify methods according to methodologies and fuzzy reasoning when emphasizing the main advantage of each method. We are especially interested in this synthesis by submitting main keys into columns to know the strategy of evaluation for each study stated above.

The composition is done by the method and its reference, the evaluated criteria column identifies the number of characteristics, sub-characteristics and attributes and the evaluated websites lists column. We noticed that the chosen evaluated websites by most of researches are well known regionally or internationally. Next, the fuzzy reasoning column classifies studies that use crisp or fuzzy evaluation and if a study uses a fuzzy reasoning, we state its advantage. The last column is for the main advantage of the listed method.
<table>
<thead>
<tr>
<th>Method and Reference</th>
<th>Evaluation Criteria</th>
<th>Evaluated websites</th>
<th>Fuzzy reasoning</th>
<th>Main advantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web-site QEM Method</td>
<td>Hundred and twenty</td>
<td>Six internationally</td>
<td>No</td>
<td>Assess the quality of academic websites to know the weaker sub-characteristics and absent attributes of the site.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2QC3Q Model Mich et al. [8]</th>
<th>Seven high level characteristics and about fifty between attributes and sub-attributes</th>
<th>Five economics faculties’ sites in northern Italy</th>
<th>No</th>
<th>Evaluate and design website quality.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- University of Trento Faculty of Economics <a href="http://www.unitn.it/economia">http://www.unitn.it/economia</a></td>
<td>- Bocconi University of Milan <a href="http://www.uni-bocconi.it">http://www.uni-bocconi.it</a></td>
<td>- Libero Istituto Universitario Carlo Cattaneo <a href="http://www.luiuc.it">http://www.luiuc.it</a></td>
<td>- University of Venice <a href="http://www.unive.it">http://www.unive.it</a></td>
<td>- University of Modena <a href="http://www.unimo.it">http://www.unimo.it</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Hybrid Model Dominic et al. [9]</th>
<th>Five high level criteria and about tens of attributes</th>
<th>Five Malaysian University website</th>
<th>No</th>
<th>Evaluate the quality of universities websites and know the most neglected criterion by Malaysian websites.</th>
</tr>
</thead>
</table>

### Table 1. Synthetic overview of quantitative methods

<table>
<thead>
<tr>
<th>Reference</th>
<th>Evaluation Criteria</th>
<th>Evaluated websites</th>
<th>Fuzzy reasoning</th>
<th>Main advantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy Linguistic Model Herrera-Videma et al. [10][11][12]</td>
<td>Four major information quality categories and about tens different evaluation dimensions</td>
<td>Websites that stores information in XML-format, e.g. scientific articles, opinion articles, etc.</td>
<td>Advantage: The generation of linguistic recommendations.</td>
<td>Evaluate the informative quality of websites. It generates the recommendations via only the visitors’ evaluation.</td>
</tr>
</tbody>
</table>

### Table 2. Synthetic overview of hybrid (quantitative and qualitative) methods
The composition of the last tables allows us so to determine and think about the proposed method how should it be and where should it be stated among those studies.

### B. Chronological evolution of proposed methods

One dynamic diagram can represent clearly the evolution of quantitative and qualitative methods throughout the time. According to the related studies described in the survey, “Fig. 1” shows this progress via a Timing Diagram.

![Figure 1. Progress of the evaluation methods via a Timing Diagram](image)

Measuring the quality of websites is becoming a real challenge across the years. In the mid-nineties, the studies were interested in the evaluation of websites in quantitative approaches. Since 2003, the studies have been oriented progressively to user viewpoint in order to evaluate in a natural and subjective way.

### C. Distribution of criteria

To better understand the innovative features of the evaluation methods described above Table 4 summarizes the different point of views in ISO 9126 Model adding some interrelations between its criteria. We have tried to make the most important sub-characteristics of websites according the mentioned works above in the survey. It seems relevant to note that each author has privileged certain quality criteria versus others. Everyone has his own vision to the criteria affecting a quality Web application; it is usually a subjective expert decision making. Some researchers are interested by quantitative evaluation with resort to online or software tools to measure criteria and others advantaged qualitative evaluation with considering the opinions of users about criteria that cannot be measured by evaluation tools. They have chosen for test websites that are well known regionally or internationally.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Sub-characteristics</th>
<th>Attributes</th>
<th>Quantitative studies - Objective</th>
<th>Hybrid studies - Objective and Subjective</th>
<th>Measurement tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability of content</td>
<td>Understandability</td>
<td></td>
<td>![6], ![7]</td>
<td>![10], ![11], ![12]</td>
<td>Webpage speed analyzer tool <a href="http://www.websiteoptimization.com/">http://www.websiteoptimization.com</a></td>
</tr>
<tr>
<td></td>
<td>Credibility</td>
<td></td>
<td>![8]</td>
<td>![13]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attractiveness</td>
<td>Download times, Total size</td>
<td>![9]</td>
<td>![14]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy</td>
<td></td>
<td>![10], ![11]</td>
<td>![13]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conciseness</td>
<td></td>
<td>![12]</td>
<td>![14]</td>
<td></td>
</tr>
<tr>
<td>Functionality</td>
<td>Relevancy</td>
<td>Link popularity</td>
<td>![6], ![7]</td>
<td>![10], ![11], ![12]</td>
<td>Search Engine Altavista <a href="http://fr.altavista.com">http://fr.altavista.com</a></td>
</tr>
<tr>
<td></td>
<td>Completeness</td>
<td></td>
<td>![9]</td>
<td>![14]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Navigability</td>
<td>Path Length</td>
<td>![10], ![11]</td>
<td>![13]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Security</td>
<td></td>
<td>![12]</td>
<td>![14]</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>Link errors</td>
<td>Broken links</td>
<td>![6], ![7]</td>
<td>![10], ![11], ![12]</td>
<td>Software tool called Xenu</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Performance</td>
<td>Correct internal URLs, Timeout URLs, Not found URLs</td>
<td>![8]</td>
<td>![13]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accessibility</td>
<td>Alternative text for images</td>
<td>![9]</td>
<td>![14]</td>
<td>An online test for webpage refer to the criteria setup by W3C-WCAG <a href="http://www.tawdis.net/">http://www.tawdis.net/</a> using WCAG 1.0 priority 1 checkpoints</td>
</tr>
</tbody>
</table>

Table 3. Criteria of quality according authors

The remaining part of this paper is composed as follows. Section 3 concerns the proposed fuzzy reduced evaluation method [17], Fuzz-Web. Section 4 deals with tests and discussion of the obtained results.

### III. Fuzz-Web: A Fuzzy Reduced Evaluation Method

Inspired by the bibliographical study and because the problem takes a multidimensional nature, we intend to use the fuzzy logic theory. So, to develop Fuzz-Web, we adopt a methodology described in “Fig. 2”.
Fuzz-Web: A Methodology Based on Fuzzy Logic for Assessing Web Sites

It is an expert based methodology based on a benchmark of institutional websites detailed in the Appendix. It contains a collection of universities and Engineering schools in Tunisia and foreign ones that have been already evaluated by some stated case studies in the survey.

First, we collect and select a set of criteria for evaluating the quality of websites. Second, we adjust the weights for chosen criteria and we determine the value of each criterion for each website based on online or software measurement tools. In the third step, the fuzzy system is designed. At the final stage, a comparison with the literature tests after getting results for the evaluated websites is necessary to validate the method.

On the other hand, the corresponding method follows the process described below.

- **First step**: The user selects and evaluates criteria for a website with the evaluation tools.
- **Second step**: The measured criteria values are as inputs of the fuzzy system to perform the fuzzy computation.
- **Third step**: Ranking the website.

“Fig. 3” else shows clearly the method process.

---

**A. A website quality characteristics and measurement tools**

We select characteristics to evaluate a website based on criteria that respect the high level criteria related to ISO 9126 Model (See Table 1). We perform the computation thanks to the attributes declared in Table 2 with the mentioned measurement tools.

1) **Usability**

A website should be usable for users so its content must be understandable, up to date, accurate and concise. Consequently, usability is a combination of factors that affect user’s satisfaction. Website design is a part of credibility which is an important element for website success. Another part of credibility is frequency of update of documents in a website. It can be given by some search engine or by software with fixing a period of control to know the frequency of update.

In our case, we concentrate to measure attractiveness sub-characteristic that plays an essential role in decision to adopt a website. According to [9], a website that makes a lot of time for loading will be less attractive than other one with faster loading. A homepage can include not only light objects as text elements but also objects with a very heavy weight such as “flash” animation or images. Thus, the nature of objects influences strongly the loading time of a web page. It is so significantly correlated and explained by its weight which proves to be a page rapidly available to users needs essentially to limit its weight. Attractiveness is computed through the proportion between web page download time and web page total size. Regarding understandability, accuracy and conciseness can be evaluated subjectively and qualitatively by an evaluation questionnaire given to users but this study is limited to measure criteria that can be evaluated in an objective and quantitative way. So, we are interested about criteria that can be measured by online or software tools.
2) Functionality

Functionality is considered as a key element of assessment according to [14].

Relevancy sub-characteristic is assessed by Web Impact Factor (WIF) developed by [18]. It is the fraction between number of links to a website and its number of indexed pages by search engine. We devote also an interest to the efficiency path length attribute used as a metric as detailed in [19].

Professional software to verify the flexibility of a site against potential errors that can be found [23].

Functionality is considered as a key element of assessment. Its testability is achieved using the average number of clicks per page thanks to navigability sub-characteristic of functionality because it needs to consistently perform its intended service without failure. Link errors were treated by most of studies as specified in [7].

Return now to the other factors of functionality that we didn’t evaluate them as accuracy and completeness of materials to use in a website because need of qualitative assessment. Besides, we should mention we ignore security sub-characteristic of functionality because it needs professional software to verify the flexibility of a site against attacks and to detect fraud or hack e.g the study of [20] shows a new security framework against web services’ XML attacks then security is a new challenge to control in World Wide Web in general.

3) Reliability

Reliability can be defined as the ability of a website to consistently perform its intended service without failure. Link errors were treated by most of studies as specified in Table 2 to measure reliability. We choose for its measurement broken links attribute that points to missing web pages as detailed in[7].

4) Efficiency

Efficiency is decomposed in performance and accessibility sub-characteristics. Both have essential role to improve the quality of a website. In the following, we define metrics that can measure them.

Precision and recall terms, defined in [21], are usually used in information retrieval domain. Precision is defined as the fraction of relevant documents retrieved by all relevant documents, while recall is defined as the fraction of relevant documents retrieved while recall is defined as the fraction of relevant documents retrieved by all relevant documents. The concept of relevant documents retrieved is replaced by found URLs crawled. We think precision can be considered as the proportion of found URLs crawled by all found URLs. On the other side, recall is considered as the proportion of found URLs crawled by all found URLs. In [22], an F-measure is interpreted as the harmonic mean of the precision and recall.

Regarding accessibility, users with disabilities or who are using assistive technologies or low-end or emerging technologies will find it impossible to access information in a website. For this reason and to ensure the website meets all current standards of accessibility, it is necessarily to satisfy this checkpoint in the evaluation. Testing is so done by looking for missing alternative text for images or graphics. It is computed by the ratio between real errors found and potential errors that can be found [23].

B. The proposed fuzzy system design

Fuzzy logic, developed by Zadeh [24], is an efficient approach dealing with problems of uncertainties and doubts in an assessment environment.

The input variables of the fuzzy system Fuzz-Web represent the six selected criteria as detailed in the previous section, namely: attractiveness, relevance, navigability, link errors, performance and accessibility. The output variable is the website evaluation.

Two linguistic terms {unsatisfactory satisfactory} are used to represent the input fuzzy sets e.g “Fig. 4” illustrates membership functions of accessibility input defined by “(1)”, while four linguistic terms represent the output {poor, average, good, excellent} defined by “(2)” and “Fig. 5”. Trapezoidal shape membership function is chosen.

\[
\mu(x) = \begin{cases} 
0, & x \leq 0.3 \\
0.1, & 0.3 < x \leq 0.5 \\
0.5, & 0.5 < x \leq 0.7 \\
1, & 0.7 < x 
\end{cases}
\]

\[
\mu_{\text{ul}}(x) = \begin{cases} 
0, & x \leq 0.3 \\
0.1, & 0.3 < x \leq 0.5 \\
0.5, & 0.5 < x \leq 0.7 \\
1, & 0.7 < x 
\end{cases}
\]

Figure 4. The membership functions of the input variable accessibility

\[
\mu(x) = \begin{cases} 
0, & x \leq 0.3 \\
0.1, & 0.3 < x \leq 0.5 \\
0.5, & 0.5 < x \leq 0.7 \\
1, & 0.7 < x 
\end{cases}
\]

\[
\mu_{\text{ul}}(x) = \begin{cases} 
0, & x \leq 0.3 \\
0.1, & 0.3 < x \leq 0.5 \\
0.5, & 0.5 < x \leq 0.7 \\
1, & 0.7 < x 
\end{cases}
\]

Figure 5. The membership functions of the output variable

We combine inputs using the logical operator AND to construct fuzzy rule base that consists of If-Then rules according to Mamdani’s inference model. A base of 64 rules is generated and Table 5 shows an example of them.

<table>
<thead>
<tr>
<th>Rule no.</th>
<th>Attractiveness</th>
<th>Relevancy</th>
<th>Navigability</th>
<th>Link errors</th>
<th>Performance</th>
<th>Accessibility</th>
<th>Evaluated weight</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>53</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>54</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0.6</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td>56</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0.6</td>
</tr>
<tr>
<td>57</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 4. Extracted examples of rules from the Rule Base of Fuzz-Web
We also adjust the weights for criteria through distributing weights for rules to highlight the importance for each criterion. An example of rule as follows:

**If** (attractiveness is satisfactory) and (relevancy is satisfactory) and (navigability is unsatisfactory) and (link errors is satisfactory) and (performance is unsatisfactory) and (accessibility is satisfactory) Then **(website is Good)**

To perform the fuzzy computation for each rule, a fuzzy set is obtained by the Minimum implication method and then they are aggregated by Maximum aggregation mechanism. The output is finally computed thanks to some defuzzification mechanisms. The centroid method is applied to convert the fuzzy set obtained from the aggregation into a crisp result.

“Fig. 6” shows the general structure of the fuzzy system.

**IV. Tests and discussion of results**

It is important to analyze different results computed for chosen criteria via online or software tools for selected websites. The **Fuzz-Web** system is implemented by Matlab fuzzy logic toolbox ([http://www.mathworks.fr/help/toolbox/fuzzy/fp243dup9.html](http://www.mathworks.fr/help/toolbox/fuzzy/fp243dup9.html)) which is an efficient tool for the conception and designing intelligent systems.

**A. Benchmark of URLs**

We choose Tunisian Engineering Schools, Tunisian universities and foreign websites for the experiment (see Appendix). Foreign websites are among those indicated above in the survey in order to compare their results with ones of Fuzz-Web system. We select three ones from the studies of Olsina et al. [6], [7] which are Stanford University (USA) classified in satisfactory category in their work, Catalunya Polytechnic University (Spain) classified also in satisfactory category and Chile University in marginal level. We take Malaysian URLs from Dominc et al. [9] ranked by order as follows with their Hybrid method: University Utara Malaysia with the highest score, University Putra Malaysia and University Sains Malaysia. Our interest was also to compare results with websites from Taiwan, we have considered some ones from the study of Huang et al. [15] classified according to their tests: Department of Computer Science of National Tsing Hua University which has as result excellent, Department of Information Management of Kainan University ranked poor, Department of Information Management of National Changhua University of education classified good and Computer Science and Engineering of Yuan Ze University with middle grade.

**B. Results per criterion**

Table 3 reveals detailed results obtained per criterion with final grades with Fuzz-Web system.

**Table 3**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Website n°</th>
<th>Attractiveness</th>
<th>Relevancy</th>
<th>Navigability</th>
<th>Link errors</th>
<th>Performance</th>
<th>Accessibility</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Website n°</td>
<td>Average</td>
<td>Good</td>
<td>Excellent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0,76</td>
<td>1,68</td>
<td>2,53</td>
<td>83,15</td>
<td>0,99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0,24</td>
<td>3,42</td>
<td>2,75</td>
<td>83,15</td>
<td>0,99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1,15</td>
<td>0,2</td>
<td>2,89</td>
<td>84,17</td>
<td>0,99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0,46</td>
<td>6,33</td>
<td>3,31</td>
<td>30,74</td>
<td>0,96</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0,78</td>
<td>0,16</td>
<td>3,09</td>
<td>94,86</td>
<td>0,99</td>
<td>0,87</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0,48</td>
<td>2,88</td>
<td>6,36</td>
<td>84,49</td>
<td>0,99</td>
<td>0</td>
<td>0,8</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0,26</td>
<td>1,59</td>
<td>4,23</td>
<td>36,33</td>
<td>0,96</td>
<td>0,77</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0,62</td>
<td>1,65</td>
<td>4,41</td>
<td>64,21</td>
<td>0,98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0,22</td>
<td>0,87</td>
<td>3,65</td>
<td>62,88</td>
<td>0,98</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0,92</td>
<td>1,98</td>
<td>2,36</td>
<td>0</td>
<td>0,91</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0,25</td>
<td>0,56</td>
<td>3,38</td>
<td>0</td>
<td>0,86</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>1,01</td>
<td>1,43</td>
<td>3,45</td>
<td>0</td>
<td>0,94</td>
<td>0,71</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>0,36</td>
<td>1,32</td>
<td>2,08</td>
<td>48,92</td>
<td>0,97</td>
<td>0,96</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>1,47</td>
<td>2,74</td>
<td>1,96</td>
<td>78,86</td>
<td>0,99</td>
<td>0</td>
<td>0,17</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0,59</td>
<td>8,19</td>
<td>2,45</td>
<td>82,33</td>
<td>0,99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>1,02</td>
<td>0,4</td>
<td>3,41</td>
<td>84,88</td>
<td>0,99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>0,23</td>
<td>0,38</td>
<td>1,91</td>
<td>54,97</td>
<td>0,76</td>
<td>0,35</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>0,24</td>
<td>2,05</td>
<td>3,35</td>
<td>48,15</td>
<td>0,97</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>0,51</td>
<td>0,28</td>
<td>2,77</td>
<td>87,47</td>
<td>0,88</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0,24</td>
<td>0,74</td>
<td>4,06</td>
<td>40,55</td>
<td>0,97</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 5. Results per criterion and grade for websites**

1) **The measurement of attractiveness**

Two principle parameters are measured:

- **Homepage download time** is calculated as the sum of the individual times of loading all the elements constituting this page (html file, images, sounds, animations, etc.).
- **Homepage total size** which is the sum of the sizes, in bytes,
of the elements constituting it. Web page download time isn’t immediately the loading time that a user can get because many browsers can multithread (charges simultaneously many elements). That is why to measure attractiveness we consider the proportion between both parameters. We consider for the test measurements for connection rate of 56kbps modem from the online tool Web page analyzer.

By considering standard for a page to load less than 30 seconds and its size less than 64 Kb, the reference for download time will be 0,47s/Kb. The range scale of this input in the Fuzz-Web system is fixed between 0 and 2.

Website of the University of Monastir (no. 14) has the bad score 1,47s/Kb for download time of homepage and the best score 0,22s/Kbytes is to Ezzitouna University website (no. 9) among Tunisian websites.

Results for attractiveness are well in general for foreign websites, the range for all of them goes between 0,21s/Kb and 0,37s/Kbytes. Globally, we remark that the foreign websites are more attractive than the Tunisian ones.

2) The measurement of relevancy
To recall relevancy is computed by “(3)”: 
\[
\text{The impact factor coefficient} = \frac{\text{number of links to a website}}{\text{number of indexed pages of that website}}
\]  
(3)

We have used Altavista search engine to ensure almost best results for those attributes.

We notice according the above results for relevancy of websites that the impact factor coefficient depends on the reputation of a website. For greater institutional websites the web impact factor (WIF) is more significant than smaller ones (10 to 0). A website with large number of indexed web pages by search engine ameliorates automatically the WIF without resort even to external links that points on it. So, the relevancy lies in the website itself with increasing its number of pages taking into account of course links from other websites. We can presume that developers of Tunisian institutional websites need supplementary efforts to improve their indexing compared with foreign ones.

3) The measurement of navigability
Navigation in a website is an important factor for its success. We realize the calculation of path length and number of nodes for each evaluated website to get average number of clicks.

The range scale for this input is taken between 0 and 10 and the reference value adopted is 2.5 according it we can classify the navigation of a website as satisfactory or not.

Virtual University of Tunis website (no. 17) has a simple navigation according to the score obtained of 1,91. However, National Engineering School of Tunis website (no. 6) has a complex navigation according to the score obtained of 6,36.

For foreign websites, Chile University website (no. 22) has the best score of 1,97 for navigability, and Department of Information Management of National Changhua of University of education (Taiwan) website (no. 29) has the bad score of 3,84.

4) The measurement of link errors
The scale for broken links indicator is between 0 and 100. If the result tends to 100 then it reveals an acceptable score.

The highest score for Tunisian evaluated websites of broken links indicator is 94,86 for National Engineering School of Sousse website (no. 5). For foreign websites, University Sains Malaysia website (no. 25) has the best satisfactory score (99,78) for link errors.

5) The measurement of performance
The range scale is into 0-1 for this criterion. Its computation is given by Xenu Software which is able to crawl whole website and generates a general report about its links. The supported attributes generated by this software are correct internal URLs, timeout URLs, server error, connection aborted, not found and forbidden access to URLs and we skipped external URLs. Those attributes make a collection and are classified into found or not found. We used so to calculate performance of website both metrics which are precision and recall given by “(4)”:

\[
\text{Precision} = \frac{\text{Number of found URLs crawled}}{\text{Total number of URLs crawled}}
\]  
(4)

\[
\text{Recall} = \frac{\text{Number of found URLs crawled}}{\text{Total number of found URLs}}
\]  

Globally, some of the evaluated websites satisfy a satisfactory measure of performance that tends to 1.

6) The measurement of accessibility
To assess accessibility we calculate the failure rate given by “(5)” and its range goes from 0 to 1.

\[
\text{Failure rate} = \frac{\text{real errors}}{\text{potential errors}}
\]  
(5)

The principle is to determine the number of potential points of failure and real errors. For example, a web page having n inline images can be seen as containing n potential points of failure. In our case, an error is an image without alternative text. Web Page Analyzer online tool can inform us about number of images in a web page. The failure rate increases with the number of real errors. The accessibility of a website is considered unsatisfactory when the failure rate increases and can reaches 1.

We remark that most of Tunisian websites have a bad accessibility which is another deficit to deal with developers. Among foreign websites, the unsatisfactory results of accessibility are for websites no. 24 and 30 respectively, the university Putra Malaysia, and Computer Science and Engineering of Yuan Ze University in Taiwan, with failure rate value equals 1.

C. Final ranking of websites
Table 3 presents also the final ranking of institutional websites with the appropriate grade. The result of defuzzification is taken as the final score to rank websites.

Based on values obtained to grade a website, we conclude that by considering for example evaluation for website of National Engineering School of Sfax, the degree of being Poor is 0, the degree of being Average is 0.69, the degree of being Good is 0.31 and the degree of being Excellent is 0. Accordingly, (Average, 0.69) is taken as the evaluation result of website (no. 3).
To better compare the results of the fuzzy reduced method for foreign websites with those of the literature review, we present comparative results in Table 4. We take the following notation to express the grade:

<table>
<thead>
<tr>
<th>Reference</th>
<th>[6][7]</th>
<th>[9]</th>
<th>[15]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Website no.</td>
<td>21 22 23</td>
<td>24 25 26</td>
<td>27 28 29 30</td>
</tr>
<tr>
<td>Literature results</td>
<td>S M S</td>
<td>0.441 0.318 0.799</td>
<td>E P G M</td>
</tr>
<tr>
<td>Fuzz-Web system results</td>
<td>E E E</td>
<td>E E E</td>
<td>E A G A</td>
</tr>
</tbody>
</table>

Table 6. Comparative table between literature results and Fuzz-Web system results

When comparing Fuzz-Web system results with Olsina et al. works [6], [7], we notice that Catalunya Polytechnic University in Spain website (no. 21) and Stanford University in USA website (no. 23) have excellent grade according Fuzz-Web system results which is the highest grade as literature results. Chile University website (no. 22) is enhanced from marginal to excellent grade through the years but it is always classified after the last websites (no. 21) and (no. 23) according “Fig. 7”.

We compare also Fuzz-Web system results with Dominic et al. work [9], we find an excellent grade for all Malaysian evaluated websites. According to “Fig. 7”, we obtain always the first rank for University Utara Malaysia website (no. 26), but in our case, University Sains Malaysia website (no. 25) is better than University Putra Malaysia website (no. 24). We can explain that those results are normal as it is mentioned above via values of measured criteria for University Sains Malaysia website which are in general satisfactory. We can mention also that the site is up-to-date.

When comparing with Huang et al. results [15], we obtain the same grades of excellent, good and average for the corresponding websites Department of Computer Science of National of Tsing Hua University in Taiwan website (no. 27), Department of Information Management of National Changhua of University of education in Taiwan website (no. 29) and Computer Science and Engineering of Yuan Ze University in Taiwan website (no. 30). The website of Department of Information Management of Kainan University in Taiwan website (no. 28) according evaluation with Fuzz-Web system has average grade and according literature results has poor grade which is between bad and middle grade. So, we report the difference to the composition of classes depending on the evaluation system structure.

The results shown above reveal the robustness of the Fuzz-Web system with the use of reduced number of criteria versus complex methods. In fact, after testing and comparing with literature results, we obtain almost the same ranking results for foreign websites. Also, concerning evaluation of Tunisian institutional websites, it was amazing as experience to know their grade and generally we notice that they have serious problems in many criteria and are not well reputed internationally. So, they should be revised by developers in order to ameliorate their quality.

Finally, a summary Table 8. below presents a briefing of the Fuzz-Web system. It can be added now to the list in Table 2. as a quantitative method.

Table 7. Summary table of the Fuzz-Web system

V. Conclusion

The assessment of website quality is a multidimensional decision making and measuring it with various tools is preferred because a single tool can’t reveal the real evaluation from different dimensions. The evaluation tools plays important role to measure different criteria and so affects the
final ranking process. We obtain sufficient results in general generated by the Fuzz-Web system comparing with literature results that reveals the right decision making and selection of different criteria. It treats only the problem of evaluation from objective quantitative way. It can be ameliorated with the subjective qualitative evaluation that leads to a new concept of assessment.

As a future work, we intend also to enlarge the benchmark and to optimize and validate the fuzzy system Fuzz-Web by adopting a training method [25].

Acknowledgement

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

References


Appendix

<table>
<thead>
<tr>
<th>Tunisian websites</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>National School of Computer Science</td>
<td><a href="http://www.ensi.rnu.tn">http://www.ensi.rnu.tn</a></td>
</tr>
<tr>
<td>National Engineering School of Monastir</td>
<td><a href="http://www.enim.rnu.tn">http://www.enim.rnu.tn</a></td>
</tr>
</tbody>
</table>
Research Group on Intelligent Machines (REGIM), National School of Engineers of Sfax.

Ilhem Kallel is active under the umbrella of IEEE; she is the founder and the chair of the “IEEE Women in Engineering” Tunisia section, since 2009; she is a member of the DIS TC in SMCS since 2010, and the chair of the IEEE CIS Tunisia Chapter for 2011. She organizes many scientific and social events and she is participating in international research cooperation projects, and in the review process of international conferences and journals. Her research interests include collective and hybrid intelligence, intelligent multirobot systems, fuzzy modeling, swarm intelligence, multiagent systems and evolving systems.

### Author Biographies

**Rim Rekik**, born in 1984, received the bachelor Degree in computer science and multimedia in 2007 from University of Sfax, the MSc graduate degree in Information Technology and Electronic Commerce, and in Information Systems and new Technologies, respectively in 2008 and 2011, from the University of Sfax, Tunisia.

She is a Phd Student on in Computing System Engineering with the Research Groups on Intelligent Machines (REGIM), National School of Engineers of Sfax.

She is an active volunteer among the ex-com of IEEE CIS Tunisia Chapter and IEEE WIE Tunisia section.

Her research field includes the Web evaluation, fuzzy modeling, and intelligent learning methods to assess web sites.

**Ilhem Kallel**, born in 1966, received the bachelor degree in computer science in 1988 from the University of Tunis, the MSc graduate degree in Information Systems and new Technologies, and PhD graduate degree in Computing System Engineering, respectively in 2003 and 2009, from the University of Sfax, Tunisia.

She is a member of the