# Evaluation Method of Malaysian University Website: Quality Website Using Hybrid Method

P.D.D Dominic[1)], Handaru Jati[2)]
[1)] Computer and Information Science Department
Universiti Teknologi PETRONAS Malaysia
dhanapal_d@petronas.com.my
[2)] Electronics Engineering Education Department
Yogyakarta State University Indonesia
handaru@uny.ac.id

*Abstract*—**In recent years, many people have devoted their efforts to the issue of quality of Web site. The concept of quality is consisting of many criteria: quality of service perspective, a user perspective, a content perspective or indeed a usability perspective. This research conducts some tests to measure the quality of Malaysian University website via web diagnostic tools online. We propose a methodology for determining and evaluating the best Malaysian university website based on many criteria of website quality, consists of Linear Weightage Model, Analytical Hierarchy Process, Fuzzy Analytical Hierarchy Process, and one new hybrid model. This new hybrid model has been implemented using Fuzzy Analytical Hierarchy Process (AHP) and linear weightage model to generate the weights for the criteria which are much better and guarantee more fairly preference of criteria. The result of this study confirmed most of Malaysian University websites are neglecting the performance and quality criteria. By applying Hybrid model between FAHP and LWM approach for website evaluation has resulted in significant acceleration of implementation, raised the overall effectiveness with respect to the underlying methodology and ultimately enabled more efficient and significantly equal or better procedure compared with other methods.**

***Keywords-Performance; University website; quality; web diagnostic; Hybrid model***

## I. INTRODUCTION (HEADING 1)

Website quality is a new topic in the software quality. Web based application can be used and reached more users than non web based application. The importance of website creates a demand from the users for the quality and fast delivery, unfortunately the complexities of the websites and technology which support this application make testing and quality control more difficult to handle. Automation of the testing for website quality is a new chance and a new method. Each definition of quality leads to lists of criteria about what constitutes a quality site. All of these criteria from multiple studies on Web quality to form a comprehensive tool for evaluating the quality of a Website that would serve to assess its trustworthiness explained in one research [1]. The principle was that 'if information can pass a test of quality, it is most likely to prove trustworthy and because of this belief, should have higher credibility. Thus, the challenge is to create a method that will guide the Internet user to the same finding as the commercial website evaluation tools without needed a lot of time. There are many scope of quality, and each measure will pertain to a particular website in varying degrees. Here are some of them: first factor is time, a credible site should be updated frequently. The information about latest update also should be included on the homepage. Second factor is structural, all of the parts of the website hold together and all links inside and outside the website should work well. Broken links on the webpage also are another factor that always downgrades the quality of website. Users expect each link to be valid, meaning that it leads successfully to the intended page or other resource. In the year of 2003, discovered that about one link out of every 200 disappeared each week from the Internet [2].

The third factor is content; number of the links, or link popularity is one of the off page factors that search engines are looking to determine the value of the webpage. The idea of this link popularity is that to increase the link popularity of a website, this website must have large amount of high quality content. Number of links to website improves access growth and helps to generate traffic [3]. Fourth factor is response time and latency; a website server should respond to a browser request within certain parameters, from 2003 to 2008 the average web page grew from 93.7K to over 312K [4]. Popular sites averaged 52 objects per page, 8.1 of which were ads, served from 5.7 servers [5], and object overhead now dominates the latency of most web pages [6].

The last criterion is performance. Technology continues to make a important impact in service industries and fundamentally shapes how services are delivered [7]. One of the research finding mention that website which has slow download time less attractive compare than website with faster download time [8]. In the recent time the average time of the connection speed is 5 Kbps (kilobytes per second). This facts give an implication that one web page with 40 Kb page size will be downloaded during 8 seconds. This matter in accordance with the 'eight second rule', this 8 second is a normal time for loading webpage and will not be tolerable from the user. This result are supported by many research result mentioned that mean of tolerable download time in the user

978-1-4244-6716-7/10/$26.00 ©2010 IEEE

side is 8.57 with standard deviation 5.9 seconds [9]. Another important aspect is information fit-to-task, information presented on a website is accurate and appropriate for the task at hand [10].

## II. LITERATURE REVIEW

The web site evaluation can be approached from users, web site designer/administrator or both together [11]. Website Quality Evaluation Method (QEM) for six university sites from different countries tested using this factor [12]. Web site architecture is classified into content and design [13], and each category is specified into evaluation criteria according to the characteristics and perception of a web site. Web site evaluation framework is developed to test 30 major airlines website all around the world [14]. This new framework called Airline Site Evaluation Framework (ASEF) consists of five categories: Finding, Interface, Navigation, Content, Reliability, and Technical aspects. Web site usability, design, and performance is developed using metrics and conducted a user test with them [14]. A quantitative inspector-based methodology for Web site evaluation, with a hierarchical structure called EQT4Web and the assessment method is general-purpose is Developed for cultural sites [15]. This new approach, hased on fuzzy operators, permits a sophisticated aggregation of measured atomic quality values, using linguistic criteria to express human experts' evaluations. Every webpage design has their own characteristics and this characteristic has drawbacks and benefits. There is a mechanism for measuring the effects of the webpage component toward the performance and quality of website. This mechanism will measure size, component, and time needed by the client for downloading a website. The main factor that will influences this download time are page size (bytes), number and types of component, number of server from the accessed web. Table 1 displayed a research conducted by IBM that can be used as a standard for performance measurement of quality [16].

TABLE I. STANDARD OF THE WEBSITE PERFORMANCE [17]

| Tested Factor | Quality Standard |
|---|---|
| Average server response time | < 0.5 second |
| Number of component per page | < 20 objects |
| Webpage loading time | < 30 second |
| Webpage size in byte | < 64 Kbytes |

## III. METHODOLOGY

This research is consisted of several stages, start with problem identification followed by research procedure and data collection, and ended with analysis of data. Basically our research purpose have twofold aim: 1) to propose the new methodology for evaluating the quality of Malaysia University website, 2) to determine the best Malaysia University website based on the criteria proposed in the new methodology, 3) to determine the best ranking method used to evaluate website quality. This research examined the selected Malaysian University website: Universiti Sains Malaysia, Universiti Kebangsaan Malaysia, Universiti Putra Malaysia, Universiti Utara Malaysia, and Universiti Teknologi PETRONAS. This data of quality website from Malaysian University website will be taken more than 30 trails on various occasions on the different period of time. Using website diagnostic tools and 4 methods proposed (LWM, AHP, FAHP, and hybrid method) the aim of this research will be explored. All of the data for this research was taken using PC with specification: Processor Pentium Mobile 740, using Local Area Network internet connection with average bandwidth 60 Kbps.

### *A. Web Diagnostic Tools*

We used a number of widely available web diagnostic tools online, thus we used widely available website performance tool and webpage speed analyzer online service (http://www.websiteoptimization.com). List of performance measured and reported by this service include total size, number of objects (HTML, images, CSS, scripts), and download times on a 56.6 Kbps connection, another available webpage online tools that we used are for testing quality is: http://validator.w3.org/checklink which was utilised in order to monitor broken links in the HTML code of the portals, while the W3C's HTML validator website (http://validator.w3.org) was used to validate the HTML code of the portals, this standard was set up by World Wide Web Consortium (W3C), the main international standards organization for the World Wide Web. A website tool for measuring Link popularity website (www.linkpopularity.com) is used to determine the amount and quality of links that are made to a single website from many websites, this based on the page-rank analysis. This research also conduct using accessibility software for testing whether the webpage tested already fulfill the criteria to be accessed by people with disabilities. This software has an ability to conduct an online test for webpage refer to the criteria setup by W3C-WCAG. Web Content Accessibility Guidelines (WCAG) is part of a series of Web accessibility guidelines published by the W3C's Web Accessibility Initiative. Accessibility software can be downloaded from www.tawdis.net. Testing using accessibility software consist of test for HTML code for knowing whether the webpage can be read by screen reader, and testing for knowing is there any alternative text for every single picture, animation, video, and audio in the webpage.

### *B. Sample Data*

In order to get the data for this research, we examined websites from five Asian countries. Rather than selecting any generic websites this research attempted to evaluate the website that are considered to be leaders in the area information technology implementation based on result of a survey research form Webometrics for University website. By doing such an approach it was felt that measures of 'best practices' could emerge.

### *C. Linear Weightage Model*

One of the linear weightage models is maximax. This model is very easy and mostly depending upon decision maker's judgment as they have to assign weights to the criteria

that involve in decision making process. In most cases there are some criteria considered as more important than others, such as load time, response time, traffic, page rank and broken link. Decision makers should assigned weight to each individual criterion in order to determine the relative importance of each one. These weights play a vital role in decision making process and extremely affect the final decision. After identifying all the criteria related to website selection decision, decision maker has to determine threshold for each criterion. In fact, threshold can be divided into two types, i.e. maximum and minimum. The load time, response time, markup validation number error, and broken link can be categories as "Smaller is better" and the threshold for this type of criteria must be maximum. On the other hand others criteria can be considered as "larger is better" such as traffic, page rank, frequency of update and design optimization where thresholds must be minimum.

$$\mathrm{Ws}_{\max} = \frac{\mathrm{Max} - \mathrm{website}}{\mathrm{Max} - \mathrm{Min}} \quad \textbf{(1)}$$

$$\mathrm{Ws}_{\min} = \frac{\mathrm{website} - \mathrm{min}}{\mathrm{max} - \mathrm{min}} \quad \textbf{(2)}$$

where
Ws max = specific website value that has maximum type of threshold with respect to a particular attribute/criterion.
Ws min = specific website value that has minimum type of threshold with respect to a particular attribute/criterion.
website = specific website that is considered at the time.
Max = maximum value of particular attribute/criteria among all websites
Min = minimum value of the same attribute among the whole websites.

The idea of using formula 1 and formula 2 is extremely valuable because they provide a method that enables the comparisons among decision criteria. Usually decision criteria have different units of measure so any comparisons among those criteria are not logically acceptable. By using the data normalization concepts which represented in formula 1 and formula 2, all the criteria will be having weights instead of variety of measurement units and then the comparisons can simply be made. When all values of the criteria matrix are calculated, series of calculations should be achieved by multiplying weights $W_i$ of criteria by the whole values $X_i$ within the matrix. The total score should also be calculated using formula 3 for each specific website which represents the specific websites' scores. The final decision table includes a total score for each website and the one who gains the highest score is recommended as the best website over all.

$$\text{Total Score} = \sum \mathrm{W_i X_i} / \sum \mathrm{W_i} \quad (3)$$

*D. Analytic Hierarchy Process*

Analytic Hierarchy Process (AHP) was originally designed by [18] to solve complicated multi-criteria decision problem, beside that AHP is appropriate whenever a target is obviously declared and a set of relevant criteria and alternatives are offered [19]. AHP has been proposed for determining the best website to support researcher through the decision making activity, which aims to determine the best website among pool of university website. AHP is a popular model to aggregate multiple criteria for decision making [20]. In AHP the problems are usually presented in a hierarchical structure and the decision maker is guided throughout a subsequent series of pairwise comparisons to express the relative strength of the elements in the hierarchy.

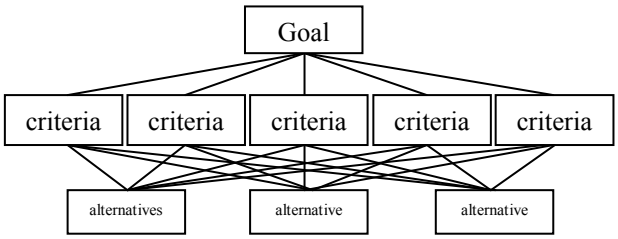


Figure 1. FAHP/AHP Model of Best Websites

Generally, AHP has the following steps:

1. Employ a pair-wise comparison approach. Fundamental scale for pair-wise comparisons developed to solve this problem [18]. The pair-wise comparison matrix A, in which the element $a_{ij}$ of the matrix is the relative importance of the $i^{th}$ factor with respect to the $j^{th}$ factor, could be calculated

$$\text{as } A = [a_{ij}] = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \cdots & 1 \end{bmatrix} \quad (4)$$

2. There are $n(n-1)/$judgments required for developing the set of matrices in step 1. Reciprocals are automatically assigned to each pair-wise comparison, where $n$ is the matrix size.

*E. Fuzzy-Analytical Hierarchy Process*

In 1965 Lotfi A. Zadeh introduced a new approach to a precise theory of approximation and vagueness based on generalization of standard set theory to fuzzy sets. Fuzzy sets and fuzzy logic are powerful mathematical tools for modeling: nature and humanity, uncertain systems in industry, and facilitators for common-sense reasoning in decision making in the absence of complete and precise information. Their role is significant when applied to complex phenomena not easily described by traditional mathematical methods, especially when the goal is to find a good approximate solution [21]. The values of fuzzy logic are ranging from 0 to 1 for showing the membership of the objects in a fuzzy set. A fuzzy number is a fuzzy quantity $M$ that represents a generalization of a real number $r$ . Intuitively, $M(x)$ should be a measure of how better $M(x)$ "approximates" $r$ . A fuzzy number $M$ is a convex normalized fuzzy set. A fuzzy number is characterized by a given interval of real numbers, each with a grade of membership between 0 and 1 [22]. A triangular fuzzy number (TFN), M is shown in Figure 2.

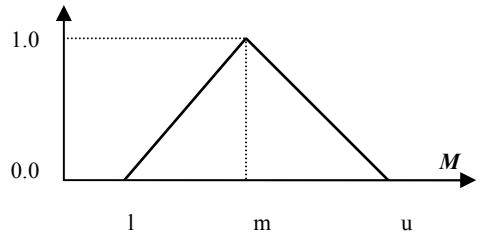


Figure 2. Triangular fuzzy number, *M*

In applications it is easy to work with TFNs because of their simple computation, and they are useful in promoting representation and information processing in a fuzzy environment. In this research implementation of TFNs in the FAHP is adopted. We have to deal with fuzzy numbers when we want to use fuzzy sets in applications. In this section, three important operations used in this research are illustrated [23]. If we define, two TFNs A and B by the triplets $A = (l_1, m_1, u_1)$ and $B = (l_2, m_2, u_2)$ . In this research the extent FAHP is used. Let $X = \{x_1, x_2, x_3, ....., x_n\}$ an object set, and $G = \{g_1, g_2, g_3, ....., g_n\}$ be a goal set. According to the method of Chang's extent analysis, each object is taken and extent analysis for each goal performed respectively. Therefore, *M* extent analysis values for each object can be obtained, with the following signs**:**

$$M^1_{gi}, M^2_{gi}, .....M^m_{gi,} \qquad i = 1,2,....,n ,$$

where $M^j_{gi}$ *(j = 1, 2, ...,M)* all are TFNs.

## IV. RESULTS AND DISCUSSION

First column in Table 2 shows the criteria of the quality website. Criteria involves in the website selection process using proposed model are load time (A), response time (B), page rank (C), frequency of update (D), traffic (E), design optimization (F), size (G), number of items (H), accessibility error (I), markup validation (J), and broken link (K).

TABLE II. TESTING RESULT FOR WEBSITES PERFORMANCE BASED ON CRITERIA

| **quality Criteria** | USM | UPM | UKM | UUM | UTP |
|---|---|---|---|---|---|
| load time | 95.51 | 85.23 | 3.59 | 12.04 | 97.58 |
| response time | 2.40 | 2.05 | 2.33 | 0.73 | 1.85 |
| page rank | 778.00 | 844.00 | 377.00 | 313.00 | 152.00 |
| frequency of update | 60.00 | 60.00 | 30.00 | 60.00 | 30.00 |
| Traffic | 185700 | 377300 | 359000 | 174600 | 90400 |
| design optimization | 29.50 | 39.00 | 30.00 | 26.50 | 63.50 |
| Size | 456135 | 38146 | 16025 | 41366 | 478578 |
| Number of items | 23.00 | 46.00 | 2.00 | 19.00 | 11.00 |
| accessibility error | 26.00 | 42.00 | 9.00 | 0.00 | 5.00 |
| markup validation | 158.00 | 234.00 | 20.00 | 2.00 | 86.00 |
| broken link | 1.00 | 19.00 | 3.00 | 0.00 | 1.00 |

Results of the websites quality test based on load time, response time, page rank, frequency of update, traffic, design optimization, size, number of items, accessibility error, markup validation, and broken link are also displayed in table 2. The data in Table 2 shows that most of the Malaysian University websites can not meet the criteria as a high quality website. Most of server response, load times, size, and number of items exceed the value standardized by IBM, except University Kebangsaan Malaysia website in load time, size, and number of items criteria. Implementation of the W3C's HTML validator highlighted that none of university website had HTML 4.01 valid entry page, most of it did not have DOCTYPE declarations. Consequences of this problem will be on the portability and development of the website. In term of broken link, four Malaysian university website or 80% of the sample have a broken link. After determining the attributes and performance results, the next step in the evaluation process is to perform a comparison of each attributes. The preference criteria matrix was obtained which compare each criterion to the others.

TABLE III. FINAL RESULT FOR E-MALAYSIAN UNIVERSITY PERFORMANCE

| Method | USM | UPM | UKM | UUM | UTP |
|---|---|---|---|---|---|
| LWM | 0.352(4) | 0.437(3) | 0.558(2) | 0.680(1) | 0.293(5) |
| AHP | 0.120(4) | 0.208(3) | 0.262(2) | 0.293(1) | 0.118(5) |
| FAHP | 0.038(5) | 0.256(3) | 0.270(2) | 0.455(1) | 0.057(4) |
| Hybrid | 0.318(4) | 0.441(3) | 0.461(2) | 0.799(1) | 0.095(5) |

Table 3 depicts the final scores of Malaysian University website based on four evaluation methods, Universiti Utara Malaysia has the highest in score in LWM, AHP, FAHP and NHM compare with the rest of university websites. Inconsistency occurred for the FAHP model, different with others three model, because for Universiti Sains Malaysia rank 5, and Universiti Teknologi PETRONAS rank 4. In order to analyze whether there is differences among the ranking composition methods, we used the Friedman test. When the null-hypothesis is rejected by the Friedman test, we can proceed with a post-hoc test to detect which differences among the methods are significant and this procedure displayed in table 4. The last step in this method is to compute the final score of each website.

TABLE IV. POST HOC TEST FOR MALAYSIAN UNIVERSITY

| | Ranked Data | | | |
|---|---|---|---|---|
| | LWM | AHP | FAHP | Hybrid |
| USM | 4 | 2 | 1 | 3 |
| UPM | 3 | 1 | 2 | 4 |
| UKM | 4 | 1 | 2 | 3 |
| UUM | 3 | 1 | 2 | 4 |
| UTP | 4 | 3 | 1 | 2 |
| $SR_i$ | 18 | 8 | 8 | 16 |

To check the ranking, note that the sum of the four rank sums is 18 + 8 + 8 + 16 = 50, and that the sum of the $c$ numbers in a row is $\frac{c(c+1)}{2}$. However, there are $r$ rows, so we must multiply the expression by r . So we have $\sum SR_i = \frac{rc(c+1)}{2} = \frac{5(4)(5)}{2} = 50$ .

Now compute the Friedman statistic

$$\chi_F^2 = \left[\frac{12}{rc(c+1)}\sum_i (SR_i^2)\right] - 3r(c+1)$$

$$= \left[\frac{12}{(5)(4)(5)}((18)^2+(8)^2+(8)^2+(16)^2\right] - 3(5)(5) = 9.96$$

In the Friedman table, the p-value for four columns and 5 rows with $\chi_F^2$ = 9.96 is 0.0185. Since all of the p-value is below $\alpha = .05$ , the null hypothesis is rejected. Since the computed FR statistic is greater than 7.815, the upper-tail critical value under the chi-square distribution having $c - 1 = 3$ degrees of freedom (Friedman Table), the null hypothesis is rejected at the 0.05 level of significant. It can be concluded that there are significant differences (as perceived by the raters) with respect to the rating produced at the four evaluation model. Naturally, we must now determine which methods are different from one another. To answer this question we use Bonferroni/Dunn's multiple comparison technique. Using this method we test $p = 12k(k-1)$ hypotheses of the form:

$H(i,j)_0$ : There is no difference in the mean average correlation coefficients between methods *i* and *j*.

$H(i,j)_1$ : There are some differences in the mean average correlation coefficients between methods *i* and *j*.

The Bonferroni t statistic is used to investigate dependent comparisons among means. This test is only good for investigating the difference between two means (i.e. it can't compare Groups LWM and AHP vs. Groups FAHP and NHM). The Bonferroni t test is the same as a normal pairwise comparison (t-test), but the critical value is different. As making many comparisons are allowed, familywise error has to be controlled by reducing the per comparison level. The overall level will be set to 0.05, and the individual per comparison levels will be equaled to 0.05 divided by the total number of possible comparisons. A total of $4C2 = 6$ different pairwise comparisons between the four means can be made. In practice, not all of these comparisons can be done, but remember that the error rate will always have to be set according to the total number of possible comparisons.

Step 1: Calculate the t' statistics.

$$t^1 = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{MS_{error}}{n} + \frac{MS_{error}}{n}}} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$$

$$SS_T = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - \frac{y^2}{N} = 150 - 125 = 25$$

$$SS_{Treatment} = \sum_{i=1}^{4} \frac{y_i^2}{n} - \frac{y^2}{N}$$

Malaysian University =

$$\frac{(18)^2+(8)^2+(8)^2+(16)^2}{5} - \frac{(50)^2}{5} = 16.6$$

$SS_E = SS_T - SS_{treatment}$

Malaysian university = = 25 – 16.6 = 8.4

$MS_{treatment} = \frac{SS_{treatment}}{a-1}$, Malaysian university = $\frac{16.6}{3} = 5.53$

$MS_E = \frac{SS_E}{[a(n-1)]}$, Malaysian university = $\frac{8.4}{[4(4)]} = 0.525$

Step 2: Determine significance of comparisons.

TABLE V. SIGNIFICANCE OF METHOD COMPARISON

| LWM vs AHP | LWM vs FAHP | LWM vs Hybrid | AHP vs FAHP | AHP vs Hybrid | FAHP vs Hybrid |
|---|---|---|---|---|---|
| $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ | $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ | $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ | $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ | $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ | $\frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{2(MS_{error})}{n}}}$ |
| $\frac{3.6-1.6}{\sqrt{\frac{2(0.525}{5}}}$ | $\frac{3.6-1.6}{\sqrt{\frac{2(0.525}{5}}}$ | $\frac{3.6-3.2}{\sqrt{\frac{2(0.525}{5}}}$ | $\frac{1.6-1.6}{\sqrt{\frac{2(0.525}{5}}}$ | $\frac{1.6-3.2}{\sqrt{\frac{2(0.525}{5}}}$ | $\frac{1.6-3.2}{\sqrt{\frac{2(0.525}{5}}}$ |
| = 4.367 | = 4.367 | = 0.873 | = 0 | =- 3.493 | = -3.493 |

Step 3. The appropriate level.

so per comparison will be: $\alpha = \frac{\alpha_{FW}}{k} = \frac{0.05}{4} = 0.0125$,

where k = number of method, df = $df MS_{error}$ = 16

Student's t tables do not contain a critical value for α=0.0125 so we have to look it up in the Dunn/Bonferroni t' table. The degrees of freedom = 16, and the number of comparison = 6. This gives a t' value: 3.008. Finally, the result for Malaysian university sector for this significant test is; LWM vs AHP: t' = 4.367 (significant), LWM vs. FAHP: t' = 4.367 (significant), LWM vs. Hybrid: t' = 0.873 (insignificant), AHP vs. FAHP : t' = 0.00 (insignificant), AHP vs. Hybrid : t' = - 3.493 (significant), FAHP vs. Hybrid : t' = - 3.493 (significant). Therefore, it can be concluded that HM ranking method is significantly better than AHP and FAHP, and Hybrid ranking method are not significantly different with LWM.

## V. CONCLUSSION

In this paper we evaluate the quality of university websites. Using a series of online diagnostic tolls, we examined many dimensions of quality, and each dimension will be measured by specific test online. Result of this study confirmed that the website presence of Malaysian university websites is

neglecting performance effort is required to meet with these criteria in the context of website design. This suggests that web developer responsible for Malaysian university websites should follow and encourage the use of recognized guidelines when designing website. To get results on the quality of a Web site, we measure sample data from five university portal and calculate response time, page size, number of item, load, mark validation, and broken link, number of link in search engine, optimization score, accessibility errors, and colorblind webpage filter test. The proposed model uses hybrid model combine Linear Weightage Model and FAHP pairwise comparisons and the measure scale to generate the weights for the criteria which are much better and guarantee more fairly preference of criteria. Limitation of this research occurred in the number of sample size and time factor, this research used limited sample size 30 data and taken during short period observation time. Future directions for this research are added criteria for evaluating websites quality, such as availability and security aspect, also from the cultural perspective, since culture has an impact upon a website. Moreover because the ultimate determinant of quality website is the users, future directions for this research also involve the objective and subjective views of the website from user's perspective.

## REFERENCES


[1] C. McInerney, "Establishing and maintaining trust in online systems," in Book Establishing and maintaining trust in online systems, *Series Establishing and maintaining trust in online systems*, Editor ed.^eds., City, 2000, pp. 257-270.

[2] F. McCowen, N. Michael, and J. Bollen, "The Availability and Persistence of Web References in DLib Magazine," in Proc. the 5th International Web Archiving Workshop and Digital Preservation (IWAW'05), 2005, pp. Pages.

[3] L. Page, R. Motwani, S. Brin, and T. Winograd, "The Anatomy of a Large-Scale Hypertextual Web Search Engine.," vol. 30, (no. 1-7), pp. 107 - 117 1998.

[4] D. Josep, P. Ana, S. Julio, Jos, and A. Gil, "A userfocused evaluation of web prefetching algorithms," *Comput. Commun.*, vol. 30, (no. 10), pp. 2213-2224, 2007.

[5] B. Krishnamurthy and C. Wills, "Cat and Mouse: Content Delivery Tradeoffs in Web Access," Editor ed.^eds., City, 2006, pp. 337 - 346

[6] J. Yuan, C. Chi, and Q. Sun, "A More Precise Model for Web Retrieval," in Book A More Precise Model for Web Retrieval, *Series A More Precise Model for Web Retrieval*, Editor ed.^eds., City: ACM, 2005, pp. 926 -927

[7] M. Durkin, "In search of the Internet-banking customer: exploring the use of decision styles," *International Journal of Bank Marketing*, vol. 22, (no. 7), pp. 484-503, 2007.

[8] J. Ramsay, A. Barbesi, and J. Preece, "Psychological investigation of long retrieval times on the World Wide Web," *Interacting with Computers*, vol. 10, (no. 1), pp.77-86, 1998.

[9] A. Bouch, A. Kuchnisky, and N. Bhatti, "Quality is in the eye of the beholder: Meeting users' requirements for Internet quality of service," in Book Quality is in the eye of the beholder: Meeting users' requirements for Internet quality of service, *Series Quality is in the eye of the beholder: Meeting users' requirements for Internet quality of service*, Editor ed.^eds., City: ACM, 2000,pp. 297 - 304

[10] E.T. Loiacono, R.T. Watson, and D.L. Goodhue, "WebQual: An Instrument for Consumer Evaluation of Web Sites," *International Journal of Electronic Commerce*, vol. 11, (no. 3), pp. 51-87, 2007.

[11] L. Olsina, D. Godoy, G.J. Lafuente, and G. Rossi, "Specifying Quality Characteristics and Attributes for Websites," in Web Engineering, vol. 2016/2001: Springer Berlin, 2001, pp. 266-278.

[12] E. Huizingh, "The Content and Design of Web Sites: An Empirical Study," *Information & Management*, vol. 37, (no. 3 ), pp. 123 - 134, 2000.

[13] G. Apostolou and A.A. Economides, "Airlines Websites Evaluation Around the World," in The Open Knowledge Society. A Computer Science and Information Systems Manifesto, vol. 19: Springer Berlin Heidelberg, 2008, pp. 611-617.

[14] W. Palmer, "Web Site Usability, Design, and Performance Metrics," *Information Systems Research*, vol. 13, (no. 2), pp. 151 - 167 2002.

[15] I. Rafikul, A. Mohiuddin, and A. Masliza Hj, "Application of Quality Function Deployment in redesigning website: a case study on TV3," *International Journal of Business Information Systems*, vol. 2, (no. 2), pp. 195-216, 2007.

[16] M. Sakthivel, S.R. Devadasan, S. Vinodh, A. Ramesh, and S. Shyamsundar, "ISO 9001:2000 based Quality Information Management Responsibility System," *International Journal of Business Information Systems*, vol. 2, (no. 2), pp. 217-237, 2007.

[17] G. Bojadziev, *Fuzzy sets, fuzzy logic, applications*: World Scientific Pub Co Inc, 1995. [18] T. Saaty, *The analytic hierarchy process: planning, priority setting, resources allocation*, London: McGraw-Hill, 1980.

[19] B. Ozden and B. Karpak, "An AHP application in vendor selection," *ISAHP*, pp. 1-21., 2005.

[20] K. Yuen and H. Lau, "Software vendor selection using fuzzy analytic hierarchy process with ISO/IEC9126," *IAENG International journal of computer science*, vol. 35, (no. 3), pp. 267-274., 2008.

[21] H. Deng, "Multicriteria analysis with fuzzy pairwise comparison," in Book Multicriteria analysis with fuzzy pairwise comparison, vol. 2, *Series Multicriteria analysis with fuzzy pairwise comparison*, Editor ed.^eds., City, 1999.

[22] Y.C. Tang and M.J. Beynon, "Application and development of a fuzzy analytic hierarchy process within a capital investment study," *Journal of Economics and Management*, vol. 1, (no. 2), pp. 207–230, 2005.

[23] D.Y. Chang, "Applications of the extent analysis method on fuzzy AHP," *European Journal of Operational Research*, vol. 95, (no. 3), pp. 649-655, 1996.

[24] H.R. Neave and P.L. Worthington, *Distribution-free tests*: Routledge London, 1989.