A Heterogeneous Multi-criteria Hierarchical Evaluation Model For Web site Services

L. Martínez[1], L.G. Pérez[1], P.J. Sánchez[1],B. Montes[2]
[1] Dept. of Computer Science, University of Jaén, 23071 - Jaén.
[2] Dept. of Psychology, University of Jaén, 23071 - Jaén. *

Abstract

Nowadays, there are a lot of websites related to the same topic that offer similar services in Internet and try to attract as many people as possible. The quality of the web services has become a critical factor in the success of these web sites. Different evaluation approaches for different types of web sites have been developed in which the users provide their opinions in a predefined numerical scale to evaluate their services. Nevertheless, the criteria used in these evaluations are related to human perceptions and not all the criteria have the same nature, therefore, the evaluation process is defined in an heterogeneous context. In this contribution, we shall propose a heterogeneous multi-criteria hierarchical evaluation model that is able to deal with aspects assessed in different domains.

Keywords: web quality, web services, evaluation, decision analysis, linguistic variables, heterogeneous information.

1 Introduction

In these days, there exist a lot of web sites competing in the same area in Internet and the quality of their services has become a critical factor for the competitiveness of the companies. In such a context, Web quality evaluation tools are necessary to filter web resources in order to avoid the bad information and services that users could receive from the web.

When we talk about the quality of a web site service, we want to show how well it meets the consumers necessities and so, it is associated with consumer satisfaction [6]. Quality can be described as conformance to requirements, while satisfaction has been defined as conformance to expectation. The ideal situation would be that there were no difference between consumer judgement of quality and experienced satisfaction. But, in fact, it is very difficult to meet all the consumers' requirements.

Due to this increasing interest in the evaluation of the services offered by the web sites we can find in the literature different models applied to specific types of web sites [5, 7]. However, most of the web evaluation models force their users to use predefined numerical scales to evaluate their services, although criteria evaluated are related to their own perceptions. Hence, the use of precise values is not very suitable, in addition, these criteria can have different nature, therefore, the users can assess them by means of heterogeneous assessments, i.e., assessed in different domains.

In this contribution, we propose a heterogeneous hierarchical quality evaluation model for general purpose web sites based on decision analysis techniques that could be specialized for specific types of web sites (e-commerce, e-bank, etc . . .). This evaluation model is user centered because it characterizes the quality of the web sites services using judgements provided by different users that surf in those web sites. Due to the fact that the aspects evaluated can have different nature, our model offers to the users the possibility of assessing their preferences using different types of information (heterogeneous context). Finally, the results of our model will be expressed with linguistic terms to make them more comprehensible by the users and/or companies. To accomplish our aims, the evaluation model will be based on decision analysis techniques and on fuzzy tools that have been used to deal with heterogeneous linguistic information [3].

Our proposal for the heterogeneous hierarchical evaluation scheme has the following steps:

1. *Evaluation framework:* this model defines an evaluation framework composed by a few numbers of quality dimensions and their respective criteria that will be evaluated by the users by means of assessments that can have different nature.
2. *Evaluation process:* we propose a hierarchical evaluation process based on two steps:

   (a)**Quality of each dimension**: in this phase we shall obtain an evaluation value for each dimension of our evaluation framework. To do so, we have to aggregate the criteria belong-

*This work has been partially supported by the Research Projects TIC 2003-03348 and TIC 2002-11942-E

ing to the dimension, but the difficulty comes from the input of heterogeneous information provided by the users because there is no standard aggregation operator for this type of information. Therefore this aggregation process consists of the following steps:

(i)*Making the information uniform:* The input information provided by the users could be expressed in different domains (numerical, interval-valued and linguistic). Therefore, to combine it, we need to unify this heterogeneous information into a unique expression domain called Basic Linguistic Term Set (BLTS) by means of the linguistic 2-tuples [1].

(ii) *Aggregation phase:* it combines the unified input assessments provided by the users using an non-weighted aggregation operator in order to obtain a collective value for each dimension.

(b) **Global Quality of the web site services**: Now, we shall aggregate the evaluation assessments obtained for each dimension of quality. In this phase, the aggregation will be carried out by means of a weighting aggregation operator, where the weights assigned to each dimension will depend on the evaluated web site. And with these weights we can annul one or several dimensions in certain types of web sites. So we can use the same framework for general purpose web sites.

This paper is structured as follows: in Section 2 we make a brief review of the 2-tuple linguistic representation model, in Section 3 we shall present our proposal for a heterogeneous evaluation model for web site services. And finally, some concluding remarks are pointed out.

2 The 2-Tuple Fuzzy Linguistic Representation Model

The 2-tuple fuzzy linguistic representation model, presented in [1], will be used in this paper to unify the heterogeneous information as in [2] . This model is based on symbolic methods and takes as the base of its representation the concept of Symbolic Translation.

Definition 1. *The Symbolic Translation of a linguistic term $s_i \in S = \{s_0, ..., s_g\}$ is a numerical value assessed in $[-.5, .5)$ that supports the "difference of information" between an amount of information $\beta \in [0, g]$ and the closest value in $\{0, ..., g\}$ that indicates the index of the closest linguistic term $s_i \in S$, being [0,g] the interval of granularity of S.*

From this concept the linguistic information is represented by means of 2-tuples (r_i, α_i), $r_i \in S$ and $\alpha_i \in [-.5, .5)$.

This model defines a set of functions between linguistic 2-tuples and numerical values.

Definition 2. *Let $S = \{s_0, ..., s_g\}$ be a linguistic term set and $\beta \in [0, g]$ a value supporting the result of a symbolic aggregation operation, then the 2-tuple that expresses the equivalent information to β is obtained with the following function:*

$$\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5)$$

$$\Delta(\beta) = (s_i, \alpha), \ with \begin{cases} s_i & i = round(\beta) \\ \alpha = \beta - i & \alpha \in [-.5, .5) \end{cases}$$

where $round(\cdot)$ is the usual round *operation, s_i has the closest index label to "β" and "α" is the value of the symbolic translation.*

Proposition 1. *Let $S = \{s_0, ..., s_g\}$ be a linguistic term set and (s_i, α) be a linguistic 2-tuple. There is always a Δ^{-1} function, such that, from a 2-tuple it returns its equivalent numerical value $\beta \in [0, g]$ in the interval of granularity of S.*

This representation model has associated a computational model that was presented in [1].

3 A Heterogeneous Multicriteria Evaluation Model For Web site Services

Here we present an user centered heterogeneous multi-criteria evaluation model for web site services, in which the users can express their opinions about the web sites by means of values assessed in different domains.

In short, we can define mathematically our problem as an evaluation process in which a set of users (experts) $E = \{e_1, \cdots, e_n\}$ will evaluate a web site, WS, providing their opinions about a set of quality dimensions, $D = \{d_1, \cdots, d_q\}$, such that each dimension, d_i, has a set of criteria, $C_i = \{c_{1i}, \cdots, c_{ti}\}$, to be evaluated . Therefore, to evaluate a web site every expert, e_k, provides his assessments about the different criteria by means of a utility vector:

$$e_k \rightarrow \{u^k_{11}, \ldots, u^k_{t1} \ldots u^k_{1q}, \ldots, u^k_{tq}\}, \ u^k_{tq} \in D^k_{tq}$$

Where u^k_{tq} is an assessment provided by the expert e_k for the criterion c_{tq} that belongs to the dimension d_q and the domain, D that can be a numerical domain N, interval-valued, I, or linguistic , L. We assume that each user may use a different domain for each criterion to evaluate the web site services according to their knowledge about the problem or the nature of the aspect assessed. Therefore, each user, e_k, can express his opinions for the criterion, c_{tq}, in the domain D^k_{tq} where $D^k_{tk} \in \{N|I|L\}$.

In the following subsections, we shall present the evaluation framework we use to evaluate the web site services and afterwards we shall present our heterogeneous hierarchical evaluation model.

3.1 Evaluation Framework

Our general purpose evaluation model shall use the following **dimensions** and *criteria*:

- **Entertainment**: it is related to amusement and pleasure contents. Its criteria are: *Interesting place to visit, pleasant browsing, entertainment and leisure, easy browsing, information diversity.*
- **Convenience**: it is related to speed, easy access to information at low cost. The criteria are: *Economy of time spent, effort spent, easy access, fast information transmission, interaction capacity, fast delivery, easy way of buying.*
- **Information reliability**: it is related to diversity, depth and actuality of information contents. And its criteria are: *Up-to-date information, information depth, search result, uncluttered web pages,easy search paths, easiness in comparing information.*
- **Security and assurance**: security perception and privacy "assurance" are known to have a big impact in user satisfaction of quality web services. Its criteria are: *Payment security, trust in supplier, privacy of purchase, data transmission assurance, privacy.*
- **Site Design**: it depends on functional and attractive elements: ease of browsing, a standard language use, interface design. Its criteria are: *Advertising contents, attractive presentation.*
- **Virtual Environment**: it is used to minimize the absence of human contact and amusement associated to shopping. Its criteria are: *Capacity of simulating reality, personal contact absence, personal-sales absence.*
- **Product Offer**: it is concerning product diversity and available brands. And its criteria are: *Easy to compare products' characteristics, diversity of product's brands, product guarantee, possibility to return.*

3.2 Evaluating web services

Our proposal to evaluate the web sites services consists of a process with the following phases:

1. Quality of each dimension.
   (a) Making the information uniform.
   (b) Aggregation phase.
2. Global Quality of the web site services.

In the next subsections we present each phase of the evaluation model in further detail.

3.2.1 Quality of each Dimension

We want to obtain a collective assessment on a dimension according to the individual opinions provided by the users regarding the different criteria. We shall aggregate the information in a two step process:

1. Making the information uniform.
2. Calculating an evaluation assessment for the dimension.

1. Making the Information Uniform: With a view to manage the information we must make it uniform, i.e., the heterogeneous information provided by the users must be transformed into a unique expression domain. We shall use fuzzy sets over a BLTS, S_T, and denoted as $F(S_T)$,

Before defining the transformation functions to unify the heterogeneous information into this BLTS, S_T, we have to decide how to choose S_T. We consider that S_T must be a linguistic term set which allows to express a quality scale easy to understand and maintain the uncertainty degree associated to each expert and the ability of discrimination to express the performance values. So in our case, we propose the following linguistic term set as, BLTS:

$$S_T = \{N, VL, L, M, H, VH, P\},$$

whose semantics has been shown in the Figure 1.

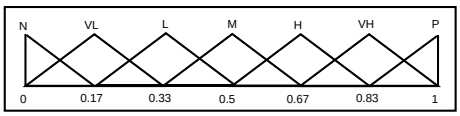

Figure 1: A Set of Seven Terms with its Semantic

Following, this process unifies the input heterogeneous information into the $F(S_T)$ in the BLTS using the next transformation functions:

1. Transforming numerical values, ϑ in $[0,1]$, into $F(S_T)$:

$$\tau_{NS_T} : [0,1] \rightarrow F(S_T)$$
$$\tau_{NS_T}(\vartheta) = \{(s_0, \gamma_0), ..., (s_g, \gamma_g)\}, s_i \in S_T \, and \, \gamma_i \in [0,1]$$
$$\gamma_i = \mu_{s_i}(\vartheta) = \begin{cases} 0, & if \, \vartheta \notin Support(\mu_{s_i}(x)) \\ \frac{\vartheta - a_i}{b_i - a_i}, & if \, a_i \leq \vartheta \leq b_i \\ 1, & if \, b_i \leq \vartheta \leq d_i \\ \frac{c_i - \vartheta}{c_i - d_i}, & if \, d_i \leq \vartheta \leq c_i \end{cases}$$

2. Transforming linguistic terms, $l_i \in S$, into $F(S_T)$:

$$\tau_{SS_T} : S \rightarrow F(S_T)$$
$$\tau_{SS_T}(l_i) = \{(s_k, \gamma_k^i) \, / \, k \in \{0, ..., g\}\}, \forall l_i \in S$$
$$\gamma_k^i = \max_y \min\{\mu_{l_i}(y), \mu_{s_k}(y)\}$$

where $\mu_{l_i}(\cdot)$ and $\mu_{s_k}(\cdot)$ are the membership functions of the fuzzy sets associated with the terms l_i and s_k, respectively.

3. Transforming interval-valued, I in $[0,1]$ into $F(S_T)$. Let $I = [\underline{i}, \overline{i}]$ be an interval in $[0,1]$. We assume that the interval has a representation, inspired in the membership function of fuzzy sets[4]:

$$\mu_I(\vartheta) = \begin{cases} 0, & if\ \vartheta < \underline{i} \\ 1, & if\ \underline{i} \leq \vartheta \leq \overline{i} \\ 0, & if\ \overline{i} < \vartheta \end{cases}$$

The transformation function is:

$$\tau_{IS_T} : I \rightarrow F(S_T)$$
$$\tau_{IS_T}(I) = \{(s_k, \gamma_k^i) \,/\, k \in \{0, ..., g\}\},$$
$$\gamma_k^i = \max_y \min\{\mu_I(y), \mu_{s_k}(y)\}$$

where $\mu_I(\cdot)$ is the membership function associated with the interval I.

Eventually, the input information is expressed by means of linguistic 2-tuples in the BLTS in order to facilitate the computation of the satisfaction assessment. To do so, we shall use the following transformation function:

Definition 3. *Let $v = \{(s_0, \alpha_0,), \ldots, (s_g, \alpha_g)\}$ be a fuzzy set over the linguistic term set $S = \{s_0, \ldots, s_g\}$. We obtain the 2-tuple equivalent in S with the following function:*

$$\chi : F(S_T) \rightarrow S_T \times [-0.5, 0.5)$$
$$\chi(F(S_T)) = \chi(\{(s_j, \gamma_j), j = 0, \ldots, g\}) = \Delta\left(\frac{\sum_{j=0}^{g} j\gamma_j}{\sum_{j=0}^{g} \gamma_j}\right) =$$
$$= \Delta(\beta) = (s, \alpha)$$

2. Calculating an evaluation assessment for each dimension: To obtain these values we apply the next phases:

1. Computing collective values for each criterion: we shall compute the collective value for the criterion (CVC), $c_{ti} \in C_i$, with a non-weighted aggregation operator, the arithmetic mean for 2-tuples [1], as:

$$CVC_{ti} = AM^*(\left(u_{ti}^k, \alpha\right),\ k = 1...n) = (u_{ti}, \alpha)$$

2. Computing an evaluation assessment for each dimension: To obtain an evaluation assessment (ED) for a dimension , d_i, we shall aggregate the collective values of its criteria by means of the arithmetic mean for 2-tuples [1] as follows:

$$ED_i = AM^*((u_{ji}, \alpha),\ j = 1...t) = (u_i, \alpha)$$

So now, we have got an evaluation assessment for each quality dimension, d_i, of our evaluation model. And we can evaluate separately each dimension to improve just certain drawbacks of our services enhancing the global evaluation of our web site services.

3.2.2 Global Quality of the web site services

In this phase, we shall obtain the global evaluation assessment, EAW, for the website services we are evaluating. In this case, we shall use a weighted aggregation vector, because depending on the specific web site different dimensions could have different importance, even some of them their value can be null. We shall apply the 2-tuple linguistic weighting average operator [1] using the weighting vector, $W = \{w_1, ..., w_q\}$, that indicates the importance of each dimension and is provided by some experts.

$EAW = W\_AM^*((u_i, \alpha), i = 1, ..., q) = (u, \alpha)$

We have obtained a global linguistic evaluation for the quality of the web site services expressed with a 2-tuple over the BLTS (linguistic evaluation scale) that is returned to the experts. This value is more comprehensible than a number or a fuzzy set.

4 Concluding Remarks

The evaluation of web site services has become a critical factor for users and companies in order to improve their commercial exchanges. However, current evaluation methods use numerical information to model users opinions, although not all the criteria used in these evaluations are of the same nature and, therefore, we have proposed a heteregeneous hierarchical evaluation model that offers the users a heterogeneous context to express these assessments giving a greater flexibility to the users that have part in the evaluation process.

References

[1] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[2] F. Herrera and L. Martínez. An approach for combining linguistic and numerical information based on 2-tuple fuzzy representation model in decision-making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(5):539–562, 2000.

[3] F. Herrera, L. Martínez, and P.J. Sánchez. Managing non-homogeneus information in group decision making. *European Journal of Operational Research*, (166:1):115–132, 2005.

[4] D. Kuchta. Fuzzy capital budgeting. *Fuzzy Sets and Systems*, (111):367–385, 2000.

[5] S. Kurnia and P. Schubert. An assessment of australian web sites in the grocery sector. pages 229–236. IADIS International Conference e-Commerce 2004.

[6] P.J.A. Nagel and W.W. Cilliers. Customer satisfaction: A comprehensive approach. *International Journal of Physical Distribution and Logistic Management*, 20(6):2–46, 1990.

[7] A. I. Torres and F. Vitorino. An e-satisfaction model - application to internet information search and purchase. pages 313–319. AIDIS International Conference e-Commerce, 2004.