

Modeling Web-Based Applications Quality: A Probabilistic Approach

Ghazwa Malak¹, Houari Sahraoui¹, Linda Badri², and Mourad Badri²

¹ DIRO, University of Montreal, Montreal, Qc, Canada
{rifighaz, sahraouh}@iro.umontreal.ca

² DMI, UQTR, Trois-Rivières, Qc, Canada
{linda.badri, mourad.badri}@uqtr.ca

Abstract. Quality assurance of Web-based applications is considered as a main concern. Many factors can affect their quality. Modeling and measuring these factors are by nature uncertain and subjective tasks. In addition, representing relationships between these factors is a complex task. In this paper, we propose an approach for modeling and supporting the assessment of Web-based applications quality. Our proposal is based on Bayesian Networks.

1 Introduction

Web-based applications are complex software systems that allow the user to create, publish, handle, and store data. Developing such applications in an ad hoc way may compromise their success and their viability. On the other hand, assuring their quality requires the use of sophisticated models. Nevertheless, in most cases, Web-based systems development lacks systematic approach and quality control [9].

Many authors have proposed guidelines [7], metrics and tools [6], methodologies and models [12] to assess the quality of Web sites or pages. Most of these proposals focused almost exclusively on Web applications usability aspects. In addition, several works proposed hierarchical quality models that are considered as oversimplified. Moreover, as Web-based applications are evolving software systems, they often yield uncertain and incomplete measurements [1], which major existing studies do not specifically address.

For Web applications, the quality is a multidimensional notion that involves a set of interdependent factors [3]. The majority of the studies recognize the importance of these factors, but some divergences exist for their definition, classification, and measurement [8]. Indeed, several quality factors are subjective [3]. In addition, modeling relationships that may exist between some factors is complex. From the measurement point of view, defining metrics for Web applications is difficult and often confusing [4]. These particular problems influence the objectivity of the assessment methodologies when evaluating the quality of Web applications.

To be efficient, a Web quality model must consider the inherent subjectivity, uncertainty and complexity of the quality factors, their measures and their relationships. Our objective is to propose a framework that considers specifically these properties. This can be done through probabilistic approaches, particularly using

Bayesian Networks (BNs) [11]. Indeed, according to Baldi [1], the dynamic evolution of the Web is probabilistic in nature and probability methods apply to diverse areas of Web modeling. In addition to the handling of uncertainty, BNs offer a good support for causality relationships between factors.

The rest of the paper is organized as follows: Section 2 discusses why and how to define a Bayesian-based quality model for Web applications. Section 3 illustrates the application of the approach to the evaluation of the Navigability Design criterion. Section 4 concludes the paper and gives some future work directions.

2 A Probabilistic Approach to Model Web Applications Quality

In this work, we are interested in the nonfunctional aspects of quality as defined by ISO/IEC 9126 standard [5]. Considering the criteria for Web applications quality, when looking into the related work, many limitations are reported. First, some criteria are subjective [3, 12], and optimal values are often contradictory for many of them [6]. This makes it hard to define realistic threshold values. Second, balancing criteria is important w.r.t. the variety of application domains [8]. However, although necessary, sub-criteria weight assignment adds a new subjective dimension. Finally, the same criterion can affect simultaneously several criteria [8]. These interdependences are difficult to represent in a hierarchical way. Thus, selecting a particular grouping (hierarchy) means that some relationships has to be ignored.

Consequently, we propose a framework that addresses specifically the subjectivity in criteria evaluation, the uncertainty in the determination of the threshold values, the difficulty in balancing criteria, and the representation of interdependences between criteria. Both theory and experience have shown that probabilities are powerful tools for modeling uncertainty [1]. In the context of quality models, reasoning with probabilities allows to weight criteria and to handle uncertainty issues. Moreover, using graphical representation provides a naturally interesting interface by which we can model interacting sets of criteria [10]. This convey to the adoption of the probabilistic approach of Bayesian Networks (BNs) to build a quality model for Web applications. A BN model can be used to evaluate, predict and possibly optimize decisions when assessing Web applications.

Building a BN for a quality model can be done in two stages. The first one consists in constructing the graph structure. Criteria are considered as random variables and represent the nodes of the BN. Criteria affecting the same criterion should be independent variables. The second step deal with the definition of node probability tables (NPTs) for each node in the graph. A conditional probability function models the uncertain relationship between each node and its parents [11]. As usual, probability tables are built using a mixture of empirical data and expert judgments.

In our previous works [8], we attempted to collect Web applications quality criteria proposed by several authors. We extended the obtained list, refined it by applying GQM paradigm [2] and classified hierarchically retained criteria on the basis of the characteristics and sub-characteristics definitions in ISO 9126 standard [5]. Then, to apply the probabilistic approach, we represented the criteria gathered hierarchically in the form of a Bayesian Network. Considering the great number of criteria and sub criteria gathered, the resulting BN model is large and complex.

However, as explained in [11], a BN can be built starting from semantically meaningful units called network “fragments”. A fragment is a set of related random variables that could be constructed and reasoned about separately from other fragments. They should formally respect the syntax and semantics of BNs. However, their use decreases the complexity when dealing with large models. Although our ultimate objective is to propose a comprehensive BN model for Web applications quality, we concentrate in this paper on the definition of navigability design fragment (at the page level). This will illustrate our approach.

3 Evaluation of the “Navigability Design” Criterion

Several works recognize the navigability design as an important quality criterion for Web applications [7, 12]. For some authors, the navigability design is a criterion of functionality [12], for others it characterizes usability [6, 7]. Authors propose many design elements, control points, directives, and guidelines to ensure the quality of navigability design. According to many definitions [7, 15, 16], navigability design in a Web application can be determined by: “the facility, for a given user, to recognize his position in the application and to locate and link, within a suitable time, required information. This can be done via the effective use of hyper links towards the destination pages”. This criterion can be also assessed at the page level and be determined by the presence of some design elements and mechanisms that improve the navigability design.

3.1 Network Structure Construction

For the rest of this section, let’s consider the *NavigabilityDesignP* as the variable representing the quality of the navigability design criterion at a Web page level, *Locate* as the variable representing the facility, for a given user, to know exactly in which page of the application he is and localize required information within the page, *Access* as the variable representing the facility, for a given user, to access to the required information in the destination page from the selected page, and *Revisit* as the variable representing the possibility, for a given user, to return to the selected page within a suitable time.

NavigabilityDesignP, *Locate*, *Access*, and *Revisit* are represented by four nodes (Fig. 1). The node *NavigabilityDesignP* is defined in terms of the three others. The direction of the edges indicates the direction of the definition/influence between criteria [11]. As the definitional relation is by nature uncertain, this structure will be completed later by probabilistic functions to state the degree to which parent nodes define the child node [11].

The obtained first structure is recursively refined. The same process is followed to construct the sub networks for *Locate*, *Access*, and *Revisit* nodes. Let’s now discuss the *Locate* fragment. To ensure a good identification of “where I am”, the presence of many design elements (as shown in fig. 2) can help user to recognize his position [7, 12]. Also, a good indication about the destination will help the user to determine his goal with less error [7]. Moreover, the presence of many navigational mechanisms supports the user in his information retrieval [7, 15]. Subsequently, the relationship

that exists between *Locate* and its parents is causal and not definitional as we have seen in figure 1. However, with 9 parents, a reorganization of this network is essential to avoid the combinatory explosion during the preparation of the probability tables. For that, we propose to group some nodes together when it is possible. The introduction of new nodes (meaningful or synthetic) gathering some parents' nodes and decreasing their number helps in the definition of probability tables. According to many definitions [7, 15] we can add the synthetic nodes by grouping some other nodes when the presence of each one of these nodes is independent from the other (fig. 3).

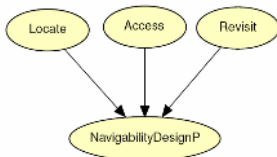


Fig. 1. BN sub network of NavigabilityDesignP criterion

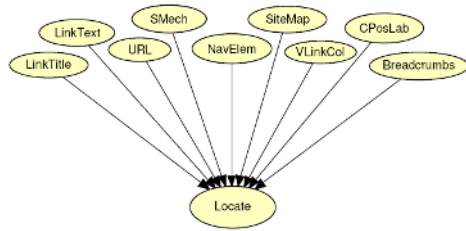


Fig. 2. BN sub network of Locate sub criterion

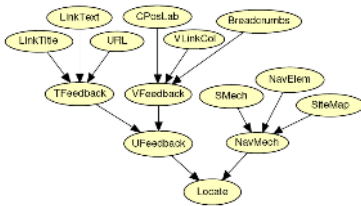


Fig. 3. Final 'Locate' BN

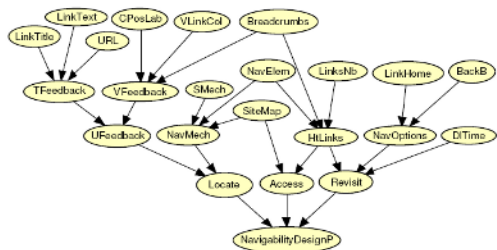


Fig. 4. Final Navigability design BN at Web page level

After constructing the sub networks for “Access” and “Revisit” nodes, all the fragments are put together to obtain the BN of the Navigability design at a page level (Fig. 4).

3.2 Defining the Probability Tables for Navigability Design Nodes

Now, to define the probability tables for navigability design nodes, we consider two types of nodes in the BN of figure 4: intermediate and input nodes. Intermediate nodes of the BN are defined/influenced by their parents such as nodes *Locate* or *NavigabilityDesignP*. These nodes are not directly measurable and their probability distribution is determined by expert judgments. Accordingly, for each node C_c that has possible values $\{V_{c1}, \dots, V_{ck}, \dots, V_{cn}\}$ and has parents $\{C_{p1}, \dots, C_{pi}, \dots, C_{pm}\}$ with

possible values $\{V_{cij}, \dots, V_{cij}, \dots, V_{cij}\}$, we need to define a table that gives the probability for all the possible value combinations

$$P(V_{ck} | V_{p1j}, \dots, V_{pmj})$$

In an initial phase, the probability tables can be defined using expert judgments. They can be adjusted using automatic learning from data samples or from processed cases.

Input nodes of the BN are criteria considered as measurable variables that do not have parents. For the majority, these criteria take binary values (present or not). According to various studies [7, 15], their presence is recommended and contributes to improve the quality of the navigability design. The other input variables have measurable numerical values. As the number of possible values can be infinite, we have to transform them into discrete variables with a limited number of values. This is done to ease the definition of probabilities. According to [13], this transformation can be achieved using fuzzy logic. Indeed, the fuzzification process takes the different criterion values and replaces them with a set of functions that represent the degree of membership of each value to different fuzzy labels (usually, “High”, “Medium”, and “Low”).

More concretely, the process of transforming crisp values into probabilities of criterion labels is defined as follows. First, we measure the criterion value for a large number of Web pages. We apply then a fuzzy clustering algorithm on them specifying the number of classes (2 or 3 classes). Figures 5 and 6 are examples of obtained clusters for the criteria *DITime* and *LinksNb* with respectively 2 and 3 classes. The third step consists in defining cluster boundaries using an approximation method (drawing intersecting lines segments tangent to the curves of clusters as shown in figures 5 and 6). Finally, when using the BN for assessing the quality of a particular page, the measure of the input criteria is transformed into a set of probabilities corresponding each to a label/class. Note that as explained by Thomas in [14], the membership function degrees can be used as probabilities with the condition that both the fuzzy clustering algorithm and the approximation method preserve the condition that the sum of the membership degrees is always equal to 1.

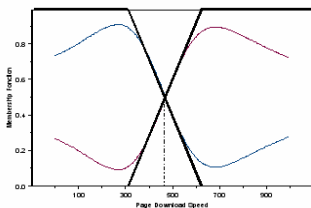


Fig. 5. Fuzzy clusters of “DITime”

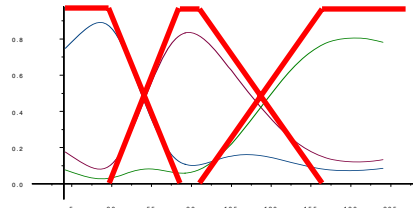


Fig. 6. Fuzzy clusters of “LinksNb”

3.3 Application Example and Preliminary Evaluation

The following example shows the feasibility of our approach using the partial BN of NavigabilityDesignP of Fig. 4. Values for input nodes are directly measured for a chosen page (in our example <http://www.pbs.org/>). In the present application context,

probability value calculated for a measured criterion is known as “evidence”. This evidence propagates through the BN via the definitional/causal links, resulting in updated probabilities for other criteria. Although the quality of NavigabilityDesignP is not known with certainty, it is evaluated in our case as good with a probability of 76.70 % good (fig.7).

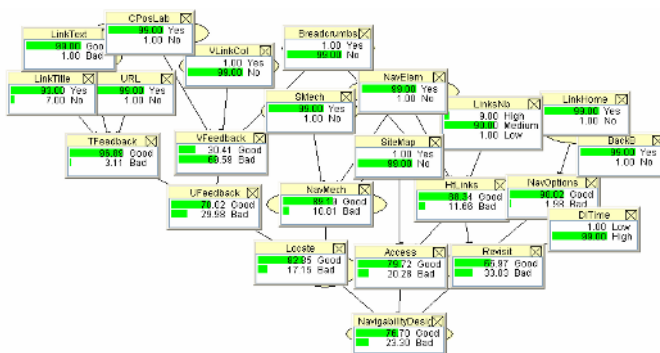


Fig. 7. State of BN probabilities showing the Navigability design quality for the PBS page

In a first experimentation to evaluate our approach, we selected two groups of web pages. The first group is composed of web pages recognized for their good quality (from *Top 100 of Webby Awards*) while the second one contains pages with poor quality (from *Worst of the Web or Webby Worthy*). We applied our BN on each page of each group and collected the probability of “good” navigability design. The idea of the evaluation is that our assessment must not contradict the status of each group.

As shown in table 1, we obtained a good score for the pages of the first group (*Webby Awards*) and low scores for the ones of the second group (*Worst of the Web or Webby worthy*). This score give a first indication that the selected and evaluated criteria are relevant.

During this experience, we used the ability of BNs to generate “What if” scenarios, i.e. giving a fixed value for one or more variables. Then, when considering high speed connection, the quality of the Navigability Design has significantly increased for many pages.

Table 1. Partial results for a rapid validation

Web applications	Navigability design quality at page level	Navigability design quality at page level if High Speed Connection	Web applications	Navigability design quality at page level	Navigability design quality at page level if High Speed Connection
Winner of Webby Awards	85.44 %	85.44 %	The Worst of the Web or The Webby Worthy	59.24 %	67.84 %
	85.09 %	90.68 %		51.22 %	58.26 %
	80.96 %	89.40 %		74.33 %	74.33 %
	84.83 %	92.81 %		57.88 %	57.88 %

4 Conclusions and Future Work

This paper proposes a general framework that uses Bayesian networks to support the quantitative assessment of Web applications quality. We illustrate our methodology for a BN “fragment”, corresponding to the Navigability Design criterion, to demonstrate the feasibility of our model. A BN graph was constructed for the considered criterion. Many experiences, involving different Web pages, were conducted using this BN. A rapid validation of the proposed approach demonstrated its relevancy. The proposed framework is extensible and adaptable. It can be re-used for specific cases to assess a particular criterion, a super criterion, a sub characteristic, a characteristic or the whole quality of a Web application. This can be done for one page, for specific pages or for all the Web application. We believe that our proposal is for web quality evaluation a good alternative to the classical hierarchical models. As future work, we plan to extend the defined network to cover more quality characteristics, to complete a global evaluation for Web applications usability, and to validate empirically the proposed model using a large scale controlled experience.

References

1. Baldi, P., Frasconi, P., Smyth, P.: *Modeling the Internet and the Web; Probabilistic Methods and Algorithms*. Wiley (2003)
2. Basili, V.R., Caldiera, G., Rombach, H.D.: *The Goal Question Metric Approach*. (1996)
3. Brajnik, G.: *Towards Valid Quality Models for Websites*. Proceedings of the 7th Conference on Human Factors and the Web (2001)
4. Calero, C.; Ruiz, J.; Piattini, M.: *A Web Metrics Survey Using WQM*. Proceedings of the International Conference on Web Engineering (2004)
5. ISO/IEC (2001) ISO/IEC 9126: *Quality characteristics and Guidelines for their use* (2001)
6. Ivory, M.: *An empirical foundation for automated Web interface evaluation*. Thesis (2001)
7. Koyani, S. J., Bailey, R. W., and Nall, J. R.: *Research-Based Web Design & Usability Guidelines*. National Institutes of Health (2003)
8. Malak G., Badri L., Badri M., Sahraoui H.: *Towards a Multidimensional Model for Web-Based Applications Quality Assessment*. Proc. of the fifth I. C. E-Commerce and Web Technologies (EC-Web'04), Spain, LNCS Vol. 3182. Springer-Verlag, (2004) 316-327
9. Murugesan, S., Deshpande, Y., Hansen, S., Ginige, A.: *Web Engineering: A New Discipline for Development of Web-based Systems*
10. Naïm, P., Wuillemin, P.H., Leray, P., Pourret, O., Becker, A.: *Réseaux Bayésiens*. (2004)
11. Neil, M., Fenton, N.E., Nielsen, L.: *Building large-scale Bayesian Networks*. The Knowledge Engineering Review, 15(3). (2000) 257-284
12. Olsina, L. Rossi, G.: *Measuring Web Application Quality with WebQEM*. IEEE MultiMedia, Vol. 9, No. 4 (2002)
13. Sahraoui, H., Boukadoum, M., Chawiche, H. M., Mai, G. Serhani, M. A, *A fuzzy logic framework to improve the performance and interpretation of rule-based quality prediction models for object-oriented software*. The 26th (COMPSAC'02). Oxford (2002)
14. Thomas, S.F.: *Possibilistic uncertainty and statistical inference*. ORSA/TIMS Meeting. Houston, Texas (1981)
15. *W3C Recommendation 5-May-1999, Web Content Accessibility Guidelines 1.0*.
16. Zhang, Y., Zhu, H., Greenwood, S.: *Website Complexity Metrics for Measuring Navigability*. The Fourth International Conference on Quality Software (QSIC'04) (2004)