Information and Software Technology 51 (2009) 599–609

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Customizing ISO 9126 quality model for evaluation of B2B applications

Behshid Behkamal [a,*], Mohsen Kahani [b], Mohammad Kazem Akbari [c]

[a] Information Technology, Ferdowsi University of Mashhad, Mashhad, Iran
[b] Ferdowsi University of Mashhad, Mashhad, Iran
[c] Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 17 December 2007
Received in revised form 20 July 2008
Accepted 8 August 2008
Available online 26 September 2008

Keywords:
Quality model
Software evaluation
Business to business application
Customizing method
ISO 9126

ABSTRACT

A software quality model acts as a framework for the evaluation of attributes of an application that contribute to the software quality. In this paper, a quality model is presented for evaluation of B2B applications. First, the most well-known quality models are studied, and reasons for using ISO 9126 quality model as the basis are discussed. This model, then, is customized in accordance with special characteristics of B2B applications. The customization is done by extracting the quality factors from web applications and B2B e-commerce applications, weighting these factors from the viewpoints of both developers and end users, and adding them to the model. Finally, as a case study, ISACO portal is evaluated by the proposed model.

1. Introduction

Software is being used in an increasingly wide variety of application areas, and its correct operation is often critical for business success. Developing or selecting high quality software products is therefore of prime importance. Comprehensive specification and evaluation of software product quality is a key factor in ensuring adequate quality. This can be achieved by defining appropriate quality characteristics, taking account of the purpose of usage of the software product [1]. It is important that every relevant software product quality characteristic is specified and evaluated, whenever possible using validated or widely accepted metrics.

There are essentially two approaches that can be followed to ensure product quality, one being assurance of the process by which the product is developed, and the other being the evaluation of the quality of the end product [2].

To evaluate the quality of the end product, a set of quality characteristics that describe the product and form the basis for the evaluation is required. This set of characteristics and the relationships between them is quality model [3] (which provides the basis for specifying quality requirements and evaluating quality). Examples of uses of a quality model are to:

- Validate the completeness of a requirements definition.
- Identify software requirements.
- Identify software design and testing objectives.
- Identify acceptance criteria for a completed software product.

In this paper, we intend to customize a quality model (ISO 0126) to identify acceptance criteria and evaluate a particular application domain; B2B application.

The rest of this paper is structured as follows: To choose a base model, first the existing quality models are reviewed and compared. In Section 2, B2B electronic commerce is defined. Then, the need for customization is discussed, and in Section 5, our approach will be presented in details. In next section, the proposed model will be applied for evaluation of a case study. Finally, a conclusion and future works are given in Section 7. The research method used to extend the model is descriptive-analytical method based on library studies.

2. Review of quality models

The state of the art in software technology does not yet present a well established and widely accepted description scheme for assessing the quality of software products. Much work has been done since about 1976 by a number of individuals to define a software quality framework. According to ISO 9126-1 [4] quality is defined as a set of features and characteristics of product or service that bears on its ability to satisfy the stated or implied needs. A quality model is defined as the set of characteristics and the relationships between them, which provide the basis for specifying quality requirements and evaluating the quality.

* Corresponding author. Tel.: +98 9155027726.
E-mail addresses: behkamal@gmail.com (B. Behkamal), kahani@um.ac.ir (M. Kahani), akbari@ce.aut.ac.ir (M.K. Akbari).

0950-5849/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.infsof.2008.08.001

In this section, some of the most standard and well-known quality models are briefly discussed, focussing on their strength and weaknesses.

2.1. McCall model

McCall's model for software quality combines eleven criteria around product operations, product revisions, and product transitions [1]. The main idea behind McCall's model is to assess the relationships among external quality factors and product quality criteria. It was developed by the US air-force electronic system decision (ESD), the Rome air development center (RADC), and general electric (GE), with the aim of improving the quality of software products.

One of the major contributions of the McCall model is the relationship between quality characteristics and metrics, although there has been criticism that not all metrics are objective. One aspect not considered directly by this model is the functionality of software products.

2.2. Boehm model

Boehm added some characteristics to McCall's model with emphasis on the maintainability of software product [3]. Also, this model includes considerations involved in the evaluation of a software product with respect to the utility of the program. The Boehm model is similar to the McCall model in that it represents a hierarchical structure of characteristics, each of which contributes to the total quality.

Boehm's notion includes users' needs, as McCall's does. However, Boehm's model contains only a diagram without any suggestion about measuring the quality characteristics.

2.3. FURPS model

The FURPS model [5] proposed by Robert Grady and Hewlett-Packard Co. decomposes characteristics in two different categories of requirements:

- Functional requirements (F): Defined by input and expected output.
- Non-functional requirements (URPS): Usability, reliability, performance, supportability.

One disadvantage of the FURPS model is that it fails to take into account the software product's portability.

2.4. Dromey model

Dromey's model seeks to increase understanding of the relationship between the attributes (characteristics) and the sub-attributes (sub-characteristics) of quality [6]. The layers of this model are defined as high-level attributes and subordinate attributes. The main idea to create this new model was to obtain a model broad enough to work for different systems.

Dromey recognizes that the evaluation is different for each product and a more dynamic idea for modelling of the process is needed. This model suffers from lack of criteria for measurement of software quality.

2.5. BBN model

The Bayesian belief network (BBN) is a special category of graphical models, where the nodes represent variables and the directed arrows represent the relation between the nodes [7,8]. As a quality model, it can be represented in a hierarchical structure. The root of the tree is the node Quality and it is connected to quality characteristics nodes. Furthermore, each quality characteristic node is connected to the corresponding quality sub-characteristics.

This model can represent and manipulate complex models that might never be implemented using conventional methods; however, because of the lack of the criteria, it can not be used for evaluation of a software product.

2.6. Star model

The software quality Star model is a conceptual model for presenting different perspectives of software quality [5]. Although this model considers different viewpoint of quality, but as with the BBN model, it suffers from the lack of the criteria.

2.7. ISO model

Other quality models have been proposed and submitted for use. While studies were useful, they have also caused confusion because of the many quality aspects offered. Thus, the need for one standard model was felt. It is for this reason that the ISO/IEC JTC1 began to develop the required consensus and encourage standardization world-wide. First considerations originated in 1978, and in 1985 the development of ISO/IEC 9126 was started.

The ISO 9126 [4] is part of the ISO 9000 standard, which is the most important standard for quality assurance. In this model, the totality of software product quality attributes is classified in a hierarchical tree structure of characteristics and sub characteristics. The highest level of this structure consists of the quality characteristics and the lowest level consists of the software quality criteria. The model specifies six characteristics including *Functionality*, *Reliability*, *Usability*, *Efficiency*, *Maintainability* and *Portability*; which are further divided into 21 sub characteristics. These sub characteristics are manifested externally when the software is used as part of a computer system, and are the result of internal software attributes.

The characteristics defined are applicable to every kind of software, including computer programs and data contained in firmware and provide consistent terminology for software product quality. They also provide a framework for making trade-offs between software product capabilities [4].

To compare the aforementioned model, their main features are presented in Table 1, showing the strengths and weaknesses of each model.

As discussed before, the ISO model seems to be more complete than the others, and is free of shortcomings of others. The most important characteristics of ISO model are: hierarchical structure, having criteria for evaluation, comprehensive expressions and terms, simple and accurate definitions, and one-to-many relationship between various layers of the model. Therefore, ISO model was chosen as the base model and was customized to make it more suitable for the evaluation of a particular application domain, which is a B2B (business to business) application.

3. B2B electronic commerce

As organizations become aware of the strategic importance of e-commerce, they will also become aware of the need for good-quality applications to achieve success in this rapidly growing e-commerce market.

Since industrial and academic interest in Business to Business electronic commerce is still evolving, any definition of what is or is not included in e-commerce is bound to be controversial. Zwass [9] uses a broad definition; the electronic commerce involves sharing business information, maintaining business relationships and

Table 1
Comparison of software quality models

Quality model	Structure	Number of levels	Relationship	Disadvantages	Advantages
McCall [14]	Hierarchical	Two	Many to many	Overlapping of components	Having evaluation criteria
Boehm [4]	Hierarchical	Three	Many to many	Lack of criteria	Including factors related to hardware
FURPS [22]	Hierarchical	Tow	One to many	Not considering portability	Separating functional and non-functional requirements
Dromey [12]	Hierarchical	Tow	One to many	Incomprehensiveness	Applicable to different systems
ISO [20]	Hierarchical	Three	One to many	Generality	– Having evaluation criteria – Separating internal and external quality
Star [22]	Non-hierarchical	–	Many to many	Lack of criteria	Considering different viewpoints
BBN [31]	Non-hierarchical	–	Many to many	Lack of criteria	Having weighted quality factors

conducting business transactions by means of telecommunication networks.

Turban [10] defines electronic commerce as the process of buying, selling, or exchanging products, services, and information via computer networks. To narrow this definition, B2B can be thought of as an e-commerce model in which all of the participants are organizations [10].

B2B e-commerce is rapidly changing the way enterprises conduct business. So, effective deployment of e-commerce applications can streamline procurement processes and improve general B2B communications, potentially resulting in large cost reductions [11]. As a result, it is essential for businesses to use appropriate applications to perform their B2B transactions.

4. Need to customize ISO model for B2B applications

In this section, we support our claim for customization of ISO quality model by reviewing the most common assessment methods of the software quality, some problems, and solutions. Then, we propose our approach to software quality evaluation, which solves most of the problems.

4.1. Review of software quality assessment methods

There are many methods for evaluating the quality of software [12,13]. In this section, three main methods of software quality assessment are introduced and their weaknesses for evaluating the quality of B2B applications are discussed.

4.1.1. Expert review

In an expert review, specialist(s) evaluate if aspects of a given software product follow established quality principles, known as heuristics [14]. The review is best done by someone who is skilled in quality issues and has some familiarity with the product domain. An expert review may be part of a larger evaluation of product quality. There are many flavors of heuristic reviews including exhaustive, principled, and scenario-based. For example, a scenario-based review requires clear customer profiles and task descriptions. So, the reviewer steps through the task scenario, keeping a log of issues and drafts a report that includes problems and recommended improvements.

Expert review is popular since it is much quicker and cheaper to carry out than other methods; but often evaluation of one person can not be expanded as acceptable evaluation for other people, because different people have different views on quality. To solve this problem, experts should be categorized into different levels and play different roles, such as users and developers. By considering their needs for software quality, tailored quality models will be developed for each group.

4.1.2. Software metrics

A good mechanism for assessing the quality of a software product is the use of metrics [15]. Software metrics are defined as standard of measurement, used to judge the software attributes, such as complexity; in an objective manner. Subjective measurement of quality comes from human estimation.

There are two methods for quality evaluating by software metrics: Black Box and White Box. For both methods, some metrics are defined and quantitative checklists are developed. In black box evaluation, different inputs are given to the software products and corresponding outputs are audited, without attention to what happens in the source code. This method is used when it is not possible to access the source code.

In white box method, metrics are just computing internal attributes of software such as line of code, cohesion, coupling, and complexity. However, acceptable quality in code evaluation does not guarantee performance and quality of software in execution. Also, with different views of quality, it is hard to find a numerical value for quality which is acceptable to all people.

4.1.3. Quality model

As discussed earlier, quality model is defined as the set of characteristics and the relationships between them; but it only explains the relationship between quality factors and sub-factors without considering their value. However, not all sub-characteristics equally are affecting characteristic. Also, these values are not equal for different software products. For example, in ISO quality model [4], *Security* and *Interoperability* are two sub-factors related to *Functionality*. If the quality of an e-payment system is assessed by the ISO model and the value of *Security* and *Interoperability* are respectively computed as S and I, then which of the following results will be the value of *Functionality*? 'S + I' or '2/3 S +1/3 I' or ....

To address this problem, for different software products, the relation and impact of each characteristics and sub-characteristics should be distinguished. Quality models can be more useful for evaluation of different applications by using weighted quality characteristics and sub-characteristics.

In addition, in the proposed model, all of these methods are used. The model is based on a quality model and customized for a particular software product (B2B) by assigning weights to the quality factors by experts. Also, software metrics are used to design an evaluation questionnaire. In the next section, the proposed approach will be explained in more details.

5. Our approach for customizing ISO quality model

In order to address the aforementioned problem, the software quality evaluation model contains five steps as follows:

Step 1: Choosing ISO quality model as a basis; after studying some well-known quality model, the ISO 9126 model was chosen as a basis.

Step 2: Identifying quality characteristics of a particular application. In this phase, first, quality attributes of B2B applications are identified through studying web applications

and B2B applications. Then by comparing these quality characteristics with quality factors of ISO 9126 model, the attributes that do not exists in this model will be added in the appropriate level in the hierarchical structure of the model.

Step 3: Choosing a group of software expert familiar with B2B e-commerce systems. This group should assign the value of all quality characteristics from the viewpoints of both users' and developers' of B2B applications.

Step 4: Assigning weights to the quality factors and sub-factors by experts.

Step 5: Developing the quality criteria.

In the following sections, each step will be explained in more details.

5.1. Choosing ISO quality model as a basis

ISO 9126 quality model, as defined in Section 2, specifies six characteristics including *Functionality*, *Reliability*, *Usability*, *Efficiency*, *Maintainability* and *Portability*; which are further divided into 21 sub characteristics. An overview of the defined characteristics and sub characteristics of this model is shown in Fig. 1 and the quality characteristics and sub characteristics of ISO model are defined below.

5.1.1. Functionality

Functionality is "a set of attributes that bear on the existence of a set of functions and their specified properties; the functions are those that satisfy stated or implied needs" [4: part 6.1] This factor consists of five sub-factors: *Suitability* means that the functionality of the application fits the needs of a user to fulfill a certain task without overwhelming the user. *Accuracy*means that the results or application's behavior is correct. *Interoperability* means that the application is able to interact with the specified systems. *Compliance* means that the application is compliant with standards, conventions or regulations in laws and similar prescriptions. *Security* is the ability to prevent unauthorized access, whether accidental or deliberate, to programs and data.

5.1.2. Reliability

Reliability is "a set of attributes that bear on the capability of software to maintain its level of performance under stated conditions for a stated period of time" [4: part 6.2] Reliability consists of three sub-factors: *Maturity* is the frequency of software faults. *Fault tolerance* is the ability of software to deal with of software faults or infringement of its specified interface. *Recoverability* is the capability to recover data affected in case of a failure and measured by the time and effort needed for it.

5.1.3. Usability

Usability is "a set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users" [4: part 6.3]. The quality factor consists of three sub-factors: *Understandability* describes the users' effort for recognizing the logical concept of an application and the applicability of that logical concept. *Learn ability* is the users' effort for learning the application as for example, operation control, input and output and finally *Operability* – the users' effort for operation and operation control.

5.1.4. Efficiency

Efficiency is "a set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions" [4: part 6.4]. The efficiency of the *time* and *resource* behavior is distinguished. The time

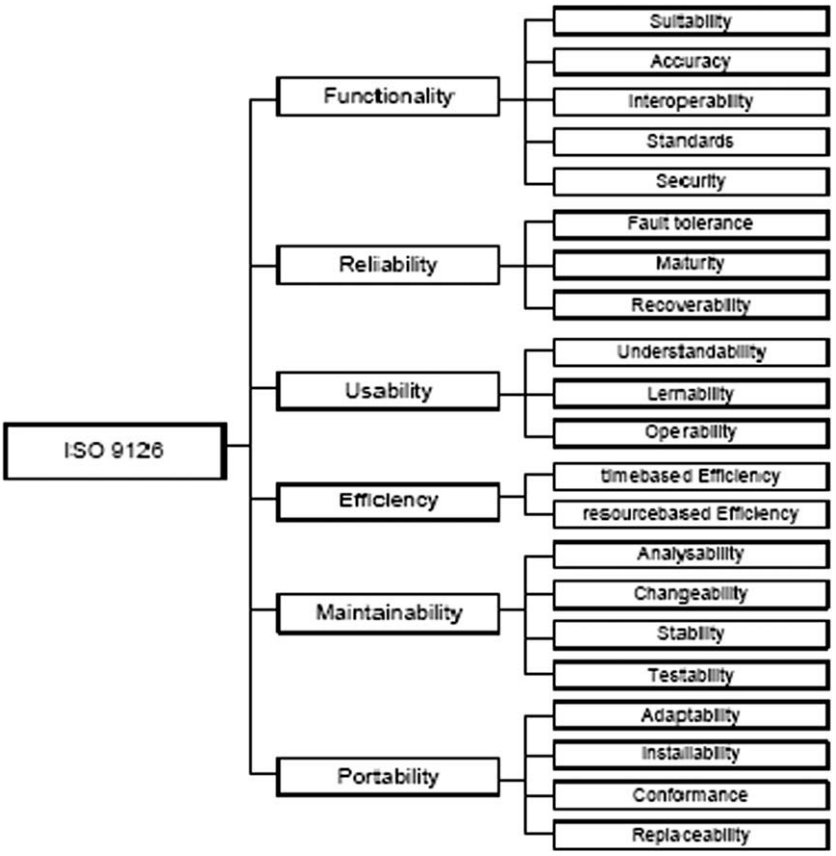

Fig. 1. The ISO 9126 quality model attribute tree.

behavior describes for instance processing times and throughput rates while resource behavior means the amount of resources used and the duration of use.

5.1.5. Maintainability

Maintainability is "a set of attributes that bear on the effort needed to make specified modifications" [4: part 6.5]. The factor consists of four sub-factors: *Analyzability*, the effort needed for the diagnosis of deficiencies or failures and for the identification of parts to be modified. *Changeability* is effort needed for modification, fault removal or for environmental change. *Stability* is the tolerance of the application towards unexpected effects of modifications. *Testability* is the effort for validating modification.

5.1.6. Portability

Portability is "a set of attributes that bear on the ability of software to be transferred from one environment to another" [4: part 6.6]. The factor consists of four sub-factors: *Adaptability* is the opportunity for adapting the application to different environments without additional effort. *Install ability* is the effort for installing the software. *Conformance* means the conformance of an application to standards or conventions relating to portability. *Replace ability* is the opportunity and effort to use an application as a replacement for another application.

Although ISO model has many advantages; but the quality factors of ISO model are too broad and general. Also, it only explains the relationship between quality factors and sub-factors, without considering the value of each relationship. So, it should be customized to become appropriate for evaluation of a particular domain application (i.e. B2B applications). The process of customization will be discussed later.

5.2. Identifying quality characteristics of B2B applications

Web applications have certain attributes which are different from most generic software systems [16]. The medium of e-commerce is the world wide web and most of e-commerce transactions perform on the basis of web. Since organizations interact through web, it is evidence that e-commerce quality is related to the quality of the web pages and web services.

In the other hand, from the architecture perspective, all e-commerce systems consist of multiple web documents. In order to develop a more effective e-commerce system, we need to know critical design factors that may substantially influence the performance of web applications [17–19].

So, the quality of e-commerce applications should follow the same principles as web applications quality. Here, to identify quality attributes of B2B applications, initially, the quality characteristics of web applications have been ranked. Then, some researches investigating the most well-known B2B applications have been studied to rank quality characteristics of these systems.

5.2.1. Web application quality

Web applications possess common quality characteristics, and B2B applications are not exceptions. An important quality characteristic for all web applications is the traceability of transactions [20], which has not been directly referred to in the ISO model. Also, considering the web nature, which requires the system availability for the users' access as 7×24, one of the web application quality needs is their availability [21]. This requirement should be added to the ISO model as a quality characteristic. There have been several researches dealing with some of these quality aspects [22–25]. Since it is not possible to explore all of them here, the quality factors common in most of them, have been summarized and ranked in Table 2.

Table 2
Ranking of web applications quality factors

Quality factor	Rank
Efficiency	1
Security	2
Usability	3
Traceability	4
Availability	5
Scalability	6
Functionality	7
Customizability	8
Recoverability	9
Consistency (Data)	10

5.2.2. The quality of B2B Applications

Most of the discussions surrounding quality metrics for e-commerce systems focus on usability, which has been covered extensively in the literature [26,27]. This is because usability has many faces; site navigation, ease of use, general layout or site topology, information presentation, product ordering and many other factors also fall under the usability umbrella. This quality factor is not discussed in more details, here, because usability is one of the main quality factors of ISO model and was explained in Section 2.7.3.

Companies producing software have presented various B2B solutions containing different components and characteristics [28]. The B2B solutions of six well-known vendors including Sieble, Oracle, People Soft, IBM, Microsoft, and SAP are investigated by many researchers [29–31]. The quality characteristics of those systems have been extracted and ranked based on the frequency of referring to them in Table 3.

5.2.3. Adding the identified quality characteristics to the model

In the previous section, the quality characteristics of B2B applications were determined. Having in mind the quality characteristics of the ISO model, four quality characteristics, *traceability*, *availability*, *customizability*, and *navigability*, are identified that do not exist in the ISO model. These characteristics have the highest ranking and significance in B2B applications. The definitions of these factors are presented here:

- *Traceability*: It refers to the software capability for exploring the correctness of the information processing in different stages of a process. This characteristic is merely used in the web application and particularly in B2B applications; it is used for tracing the system operations such as order placement or payment.
- *Availability*: It refers to the extent to which the software is available for users whenever the system is required. Availability is of great importance in B2B applications.
- *Customizability*: It refers to the software capability in accordance with the users' need and the increase in their satisfaction when using the software. This characteristic has been added into our model. Examples of this criterion in the B2B applications could be multi-language or the multi-currency.

Table 3
Comparison of quality factors of B2B applications

Quality factor	Total points	Quality factor	Total points
Security	5	Intractability	2
Scalability	5	Functionality	2
Efficiency	5	Customizability	2
Accessibility	5	Supportability	1
Traceability	3	Open source platform	1
Reliability	3	Portability	1
Integrity	3	Generality	1
Usability	2	Changeability	1
Manageability	2	Compatibility	1

- *Navigability*: This characteristic refers to the ease of users' quick and efficient access to information. This characteristic can be measured through criteria such as having alphabetical and subject-based table of contents, site map, and the ways of accessing to information through maps (vertical, horizontal and hybrid).

As mentioned before, the ISO model has three levels of quality factors, sub-factors, and criteria. These B2B characteristics should be place in the appropriate level in the hierarchical structure of the ISO model. Since quality characteristics of the first level of ISO model are so comprehensive, these new characteristics should be placed in the second level. According to the definitions of these quality characteristics, *traceability* is the sub-factor of *functionality*, and *availability* is related to *reliability*. Accordingly, *navigability* and *customizability* are placed as sub-factors of *usability*.

5.3. Choosing a group of software experts

Different perspectives of quality can be considered, such as user view, developer view, manager view, etc. For example, from the system user's point of view, *usability* of the system is the most important one; while from the viewpoints of developers and maintainers, *analysis ability* and *maintainability* are the most important ones.

The overall quality of a product can be expressed by a combination of different views. However, it must be taken into account that no software product can satisfy all of the stakeholders' needs at the same time. In our context, the users' and developers' viewpoints will be considered.

So, after a lot of efforts for selecting knowledgeable and skilled experts in software engineering, we chose twenty persons including ten Iranian and ten International experts from different universities. They all are professors with PhD degree and specialists of software quality assessment who are familiar with or have good experiences using e-commerce systems. The questionnaires were sent them via email. A total of 17 responses were received, giving a return rate of 85%. Two responses were invalid, because of the conflict in ranking of some quality characteristics in the first part of the questionnaire. Thus, the remaining 15 questionnaires formed the basis of the analysis. Evaluators were asked to fill out the questionnaire in one of the two viewpoints, users and developers of B2B applications. The 50% of evaluators responded from end user' viewpoint; while the remaining 50% did it as developers.

5.4. Assigning weights to the quality factors and sub-factors by experts

Given the fact that there are some trade-offs between software product capabilities and it is not possible to satisfy all of the software requirements at the same time, it is essential to determine the weights of all quality aspects in any software product.

There are many ways to determine the weights of the quality characteristics that are discussed in the literature [32–34]. As our approach is a combination of the *mutual comparison method* and *analytical hierarchical process*, these methods are briefly explained here; and then, the proposed method will be discussed.

5.4.1. Analytical hierarchy process (AHP)

One of the most appropriate methods for weighing the criteria is analytical hierarchy process (AHP) [35,36] method. This method is a strong managerial tool for multi-criteria decision making. The AHP decision problem is structured hierarchically at different levels; each level consists of a finite number of decision elements. The top level of the hierarchy represents the overall goal, while the lowest level is composed of all possible alternatives. One or more intermediate levels embody the decision criteria and sub-criteria. The relative importance of decision elements (weight of criteria and scores of alternatives) is assessed indirectly from comparison judgment during the assessment of local priorities. The decision-maker is required to provide his/her preference by comparing all criteria, sub-criteria and alternatives with respect to the upper value element. The last step of AHP aggregates all local priorities by a simple weighted sum.

5.4.2. Mutual comparison method

The relative importance between some attributes can be derived on a ratio scale by mutual comparisons of the quality attributes [37,38]. The set of all such comparisons can be represented in a square matrix with its elements denote the strength of preference of an attribute in the left column over an attribute in the top row. The judgment reflects the answers to two questions: which of the attributes is more important, and how strongly. This is done using 5 scales ('1' denotes the equal importance and '5' denotes the extreme importance). For a set of 'n' quality attributes one needs $n\,(n-1)/2$ comparisons because there are 'n' 1's on the diagonal due to comparison of attributes with themselves and for the remaining judgments half are reciprocals.

5.4.3. Our approach

We design a questionnaire for deriving the weights of the quality factors that has two parts. In the first part, quality factors of the first level in the model are weighted, and the values of quality sub-factors are assigned in the second part (A sample questionnaire is presented in the appendix).

Our approach to derive the weights of quality factors is a combination of the two methods of mutual comparison and analytical hierarchical process. The weights of the six quality factors of model is obtained through mutual comparison method; while the relative importance of 25 quality sub-factors are obtained by AHP.

After collecting questionnaires, the summary of the processes done for weighting of the quality factors and sub-factors is as follows:

Phase 1:

Based on the relative importance of six quality factors, the square matrix for mutual comparison has been formed in Table 4, where 'F', 'R', 'U', 'E', 'M', and 'P' stand for "Functionality", "Reliability", "Usability", "Efficiency", "Maintainability", and "Portability", respectively. An example of a comparison statement would be 'usability is very strongly more important than portability', yielding the value 4 in the corresponding cell, of the comparison matrix.

As seen in Table 4, the '1' on the main diagonal of the matrix is the result of comparing each quality factor with itself. Generally, according to the definition of mutual comparison matrix, the following relations exist in matrix A:

$$A(i,j) = n \Rightarrow A(j,i) = 1/n \quad i \# j$$
$$A(i,j) = A(j,i) = 1 \quad i = j$$

Phase 2: After solving the equation $A - \lambda I = 0$, the Eigen values are obtained (in this equation, 'λ' is the Eigen value of matrix 'A', and 'I' is 6×6 identify matrix).

Table 4
Mutual comparison of matrix 'A'

	F	R	U	E	M	P
F	1	1	3	2	3	3
R	1	1	3	2	4	5
U	1/3	1/3	1	1/3	3	4
E	1/2	1/2	3	1	4	4
M	1/3	1/4	1/3	1/4	1	4
P	1/3	1/5	1/4	1/4	1/4	1

Phase 3: The weights of the attributes are now obtained by solving the Eigen value problem related to the comparison matrix and setting the weight vector equal to the Eigen vector corresponding to the largest Eigen value (λ_{max}). The result is shown in Table 5.

Phase 4: The normalized weight vector of the above matrix is obtained as:

$$(F,\ R,\ U,\ E,\ M,\ P) = (0.26, 0.29, 0.12, 0.21, 0.08, 0.04)$$

These four phases have been repeated for each of the collected questionnaires. After calculating the weights of the first-level quality factors of the model, the weights of sub-factors can be determined.

Phase 5: At the second part of the questionnaires, Respondents were asked to determine the relative importance of the quality sub-factors under the common quality factor by assigning a number, without duplication. This number was used to calculate the mean score of each sub-factor and determine the overall relative importance of the factors.

After analyzing the results, the weights of quality factors and sub-factors of model, determined from both of users' and developers' viewpoints, are shown in Table 6. For example, weights of *usability* are 0.14 and 0.22 from developers' and users' viewpoints, respectively. Also, the relative importance of security is 0.19 from user's viewpoint; while the absolute weight of the security is 0.57 (0.30*0.19=0.57) which was obtained though multiplying relative weight of *security* by the weight of *functionality*.

Table 5
Eigen vector and weight vector of matrix 'A'

Eigen value of λ max	Normalized weight vector
0.5620	0.2614
0.6177	0.2874
0.2611	0.1215
0.4435	0.2064
0.1687	0.0785
0.0964	0.0448

5.5. Development of quality criteria

Using the literature in the field of software metrics [21,39,18,19] and criteria of the ISO model [4], we obtained a list of 44 criteria for evaluation of B2B applications. These criteria were added to the third level of the proposed model. To better show the proposed model, the first and the second levels of the model have been shown in Table 6, along with the weights of each quality factors; and the relationships between quality factors and quality criteria are presented in Table 7.

6. Applying the proposed model to a case study

To show that the proposed model can be used in practice; it was applied to evaluate a B2B portal. The case study of our work is the ISACO portal [40]. This portal facilitates online communication of ISACO Company with its suppliers and distributors. So, first the company and its business partners are introduced briefly, followed by the evaluation of the B2B application.

6.1. ISACO portal

Iran Khodro Co. (IKCo) is the largest car manufacturing company in Iran, with more than 60% of the car market share in the country. ISACO is one of the IKCo partners responsible for production, distribution, import and export of automobile parts. ISACO takes advantage of a wide network of distributors, including 822 stores, 674 representatives and 946 retailers and has an integrated portal for online communication with its partners, both inside and outside of the country.

ISACO portal includes an intranet and a web site and both of them are connected to a database. So, this integration of all back office and front office systems makes it different from similar systems.

Table 6
First and second levels of proposed model with their weights

Quality factor	Developers' viewpoint			Users' viewpoint		
	Weight	Quality sub factor	Relative weight	Weight	Quality sub factor	Relative weight
Functionality	0.25	Suitability	0.18	0.30	Suitability	0.23
		Accuracy	0.19		Accuracy	0.19
		Interoperability	0.17		Interoperability	0.19
		Security	0.25		Security	0.19
		Traceability	0.21		Traceability	0.20
Reliability	0.28	Maturity	0.23	0.19	Maturity	0.30
		Fault Tolerance	0.26		Fault Tolerance	0.18
		Recoverability	0.21		Recoverability	0.18
		Availability	0.30		Availability	0.34
Usability	0.14	Understandability	0.16	0.24	Understandability	0.20
		Learn ability	0.17		Learn ability	0.20
		Operability	0.20		Operability	0.14
		Attractiveness	0.13		Attractiveness	0.17
		Customizability	0.19		Customizability	0.17
		Navigability	0.15		Navigability	0.12
Efficiency	0.19	Time behavior	0.49	0.18	Time behavior	0.63
		Resource utilization	0.51		Resource utilization	0.37
Maintainability	0.08	Analyzability	0.25	0.05	Analyzability	0.28
		Changeability	0.29		Changeability	0.25
		Stability	0.24		Stability	0.24
		Testability	0.22		Testability	0.23
Portability	0.06	Adaptability	0.31	0.04	Adaptability	0.23
		Install ability	0.24		Install ability	0.24
		Co-existence	0.21		Co-existence	0.27
		Replace ability	0.24		Replace ability	0.26

Table 7
Relationships between quality factors and quality criteria

No.	Quality factors	Functionality	Reliability	Usability	Efficiency	Maintainability	Portability
	Criteria						
1	Access audit ability	*		*			
2	Access controllability	*					
3	Audit ability	*				*	
4	Audit trail capability					*	
5	Clarity	*		*			
6	Communication Commonality	*					
7	Completeness of description			*		*	
8	Completeness	*					
9	Complexity					*	
10	Computational Accuracy	*					
11	Conciseness						
12	Correctness	*					
13	Data Commonality	*					
14	Degradability		*				
15	Demonstration accessibility			*			
16	Ease of installation						*
17	Effectiveness of help System in use			*			
18	Effectiveness of the user documentation			*			
19	Expendability						*
20	Failure avoidance		*				
21	Globalization			*			
22	Hardware environmental adaptability						*
23	Hardware Independence						*
24	I/O devices Utilization				*		
25	Incorrect operation avoidance		*				
26	Instrumentation		*				
27	Mean recovery time		*				
28	Memory utilization				*		
29	Modification complexity					*	
30	Modularity					*	
31	Organizational environment adaptability						*
32	Response time				*		
33	Restart ability		*				
34	Restorability		*				
35	Reusability					*	
36	Robustness		*				
37	Self-Description Documentation					*	
38	Simplicity			*			
39	Software Independence						*
40	Throughput				*		
41	Trust ability		*				
42	Turnaround Time				*		
43	Uniformity	*		*			
44	Withdraw ability						*

6.2. Evaluation of ISACO portal

In this section, the evaluation of ISACO portal using the customized quality model is discussed. This whole process can be broken down into four steps, as is explained here.

6.2.1. Design and completion of the evaluation checklist

Initially, based on the previous researches [41,42], a checklist was developed, including 68 questions for measuring 25 sub-factors of the proposed model. There are two types of questions: The first part should be answered by the system developers and is mostly related to the quality factors such as maintainability, portability, and reliability. The second part includes questions designed to extract the users' viewpoints. They mostly concern with the quality factors such as functionality, usability and efficiency.

For answering the first group of questions, ten developers of ISACO Portal have been requested to complete the checklist, and the second group of questions is answered by a couple of the skilled user of this system, including the first author.

6.2.2. Calculating the values of the quality factors

Based on the results of the checklist, the value of each sub-factor is obtained. The results represent the quantities of the quality factors in ISACO portal. Table 8 shows the raw values of the quality factors of the model. For example, the value of *functionality* (3.67) is obtained from the average of five values of sub-factors under functionality (3.2, 3.4, 4, 4 and 3.75).

6.2.3. Calculating the values of the model's components

The values of the sub quality factors are calculated from users' and the developers' viewpoints. In this section, the values of the six quality factors of the model are calculated based on the values of the corresponding sub factors and their weights. For example '22.94' for final value of functionality is obtained from normalizing the multiplication of the weight of Functionality (0.25) by its value (3.67). Table 9 briefly represents the final value of quality factors of the model.

6.2.4. Calculating the final quality of system

In the former steps, the values of the first and second levels of the model were calculated. In this section, the final quality of ISACO portal is calculated based on the six values of the main quality specifications.

As shown in Table 9, the final quality of the ISACO portal is 75/90 from the users' viewpoint and 74/96 from the developers'. The 1% difference denotes the high accuracy of the model, considering

Table 8
Numerical values of model's components for ISACO portal

First level of model	Value (out of 4)	Second level of model	Value (out of 4)
Functionality	3.67	Suitability	3.20
		Accuracy	3.40
		Interoperability	4.00
		Security	4.00
		Traceability	3.75
Reliability	3.12	Maturity	2.80
		Fault tolerance	2.33
		Recoverability	4.00
		Availability	3.33
Usability	2.91	Understandability	2.50
		Learn ability	2.00
		Operability	2.60
		Attractiveness	4.00
		Customizability	3.00
		Navigability	3.33
Efficiency	2.25	Time behavior	2.50
		Resource utilization	2.00
Maintainability	2.75	Analyzability	2.75
		Changeability	2.25
		Stability	2.67
		Testability	3.33
Portability	2.54	Adaptability	2.50
		Install ability	2.25
		Co-existence	1.75
		Replace ability	3.67

Table 9
Final quality of system from both viewpoints

Quality factors	Developers' viewpoint		Users' viewpoint	
	Weight	Final value	Weight	Final value
Functionality	0.25	22.94	0.30	27.52
Reliability	0.28	21.84	0.19	14.82
Usability	0.14	10.18	0.24	17.46
Efficiency	0.19	10.69	0.18	10.12
Maintainability	0.08	5.50	0.05	3.43
Portability	0.06	3.81	0.04	2.54
Total quality	74.96		75.90	

the high difference between the weight of values of the quality factors from the users' and developers' viewpoint.

7. Evaluation and discussion

One important stage in each engineering process is evaluation [16]. There are various methods for evaluating the quality, each of which has different advantages and disadvantages. So, in this section, the proposed model will be evaluated by "Metrics Evaluation method". Using this method allows better understanding of the model efficiency for evaluating software products.

Since there are not many exact criteria to be referred to in this area; according to previous works, some available criteria for evaluating a quality model are gathered and then, on the basis of each criterion, the model is evaluated [43,13]. Some of these criteria are related to the existing models [14]; while the others are about characteristics that a desirable quality model should have. In this section, four criteria have been chosen including comprehensiveness, understandability and accuracy.

7.1. Comprehensiveness

One of the criteria used to assess quality models is comprehensiveness. This criterion looks for the inclusion of all quality factors and characteristics of the software systems in the area of study. Since the proposed model is provided for assessing B2B applications, the comprehensiveness of the model is studied from two aspects including web documents architecture and stakeholders' viewpoints.

7.1.1. Comprehensiveness in term of web document architecture

As it was mentioned in Section 5.2, from the architecture perspective, all e-commerce systems consist of multiple web documents. So, in this section, the quality factors of B2B applications are presented based on the layers of the web document architecture. Web document architecture has four layers of content, structure, interaction, and presentation [17]. Table 10 shows the quality factors required for the success of B2B applications based on the layers of this architecture. As shown, this model contains all the quality factors required for various layers of web document architecture.

7.1.2. Comprehensiveness in terms of considering different viewpoints

As it was mentioned before, the quality of software is not the same from different stakeholders' viewpoints, such as owners, developers, and users. To evaluate the success of a system, it is necessary to consider the satisfaction of all of these stakeholders. Here, the weights of the quality factors of the model have been determined according to the viewpoints of the developers and users, who are among the most important stakeholders. This allows a more accurate evaluation of quality in e-commerce applications. Therefore, the final quality of the application is compared and studied from these two points of view.

7.2. Understandability

Another criterion for evaluation of quality models from the standpoint of software engineering is clarity and understandability [26]. According to this criterion, the structures and components of the model should be clear and unambiguous. Ambiguity in the model results in an incorrect interpretation of the relationships among model components and error in using the model for software evaluation. As mentioned earlier, among the remarkable factors of the ISO quality model are hierarchical structure, use of common expressions and titles, presentation of clear and accurate definitions of model components, and one-to-many relationship between various layers of the model. Since the structure of the proposed model is based on the ISO model, it has all of these factors, and hence, has a higher level of clarity and understandability in comparison with other models.

7.3. Accuracy

As mentioned before, the final quality of the case study is 75.90 from the users' viewpoints and 74.96 from that of the developers' ones. The 1% difference denotes the high accuracy of the model, considering the difference between the weights of values of the quality factors from the users' and developers' viewpoint.

Also, by evaluation of ISACO portal by ISO model, the total quality calculated as '70.83' as it is shown in Table 11.

Table 10
Relation between quality factors and web document architecture

Layer	Quality factors	Layer	Quality factors
Content	Reliability Accuracy	Interaction	Interoperability Accessibility Efficiency
Structure	Navigability Traceability Compatibility Efficiency	Presentation	Attractiveness Compatibility Customizability Understandability

Table 11
Quality evaluation of ISACO by ISO model

First level of ISO Model	Value (out of 4)	Second level of ISO Model	Value (out of 4)
Functionality	3.65	Suitability	3.20
		Accuracy	3.40
		Interoperability	4.00
		Security	4.00
Reliability	3.04	Maturity	2.80
		Fault Tolerance	2.33
		Recoverability	4.00
Usability	2.77	Understandability	2.50
		Learn ability	2.00
		Operability	2.60
		Attractiveness	4.00
Efficiency	2.25	Time behavior	2.50
		Resource utilization	2.00
Maintainability	2.75	Analyzability	2.75
		Changeability	2.25
		Stability	2.67
		Testability	3.33
Portability	2.54	Adaptability	2.50
		Install ability	2.25
		Co-existence	1.75
		Replace ability	3.67
Total quality	70.83		

It is determined that customized model is more accurate than base model for these reasons:

(1) There are not some quality characteristics in the ISO model, including customizability, traceability, navigability and availability. So, the values of these quality factors are not considered in the total quality value calculated by the ISO model.
(2) Because of the lack of weighed quality factors in the ISO model, there is not different between quality assessment from users' and developers' viewpoints.

8. Conclusion and future works

In this paper, ISO 9126 quality model was extended for the evaluation of B2B applications. First, by studying web applications and B2B solutions, the most important common quality factors were extracted and ranked. After adding the extracted quality factors to the ISO model, the weight of each quality factor of the model is determined from the viewpoints of developers and end users. At the end, a case study is evaluated by proposed model.

In general, the contributions of this paper can be summarized into three areas: presenting a method for customizing a general quality model for evaluation of a particular domain; considering two main viewpoints in quality assessment of software products and presenting a method for weighting the quality factors.

Due to the novelty of e-commerce systems, there are many research directions in which this work can be expanded. For instance, it is possible to study the quality of the production process and system development in addition to the evaluation of the product. Therefore, one way to continue the research is taking into account the quality factors during system development processes.

On the other hand, subsequent research on this field may be focused on defining quantitative criteria for the quality factors presented in this paper, so that by using such criteria, the quality factors, and eventually, the final quality of the software can be quantitatively measured. Also, given the fact that the main goal of software quality evaluation is taking the system into its final success, and in e-commerce systems this success is affected by environmental factors (such as technical and legal infrastructures), the availability of the required backgrounds has a tremendous effect on the successful execution of various phases of transactions in the digital space. Therefore, focusing on environmental infrastructures as another dimension of the model may be an appropriate research direction.

Acknowledgement

We would like to thank the Software Development Department of ISACO Company for invaluable assistance in completing the questionnaire of our study.

Appendix. The Questionnaire

Dear Sir/Madam,

This questionnaire has been designed to weighing quality factors of B2B applications. In my research, I have identified six quality factors: *Functionality*, *Reliability, Usability, Efficiency, Maintainability* and *Portability*. These quality factors are details into quality attributed in the second level of model. Your opinion helps me determine the degree of importance of them.

This questionnaire has been divided into two parts; the goal of the first part is to obtain the strength of preference of quality factors using *Mutual Comparison* method. In the second part, the degree of the importance of quality attributes under the common quality factors will be obtained.

Thanks to you for your attention and time
Behshid Behkamal

Part I

– In each row of the following Table 12, which of the quality factors is more important with respect to the B2B applications? And how strongly?Please check the related box and use the 1–5 scale for your idea.(1: Equal importance – 5: Extreme importance)

Part II

– In your opinion, what is the relative importance of the quality attributes under the common quality factors?

Please identify your answer by percentage for each of quality attributes under common factor (See Table 13).

Table 12
Mutual comparison of quality factors

First quality factor	Second quality factor	Strength of preference (1–5)
Functionality	Reliability	
Maintainability	Portability	
Usability	Efficiency	
Efficiency	Maintainability	
Reliability	Portability	
Functionality	Usability	
Reliability	Efficiency	
Efficiency	Maintainability	
Portability	Reliability	
Functionality	Efficiency	
Usability	Maintainability	
Portability	Usability	
Maintainability	Functionality	
Reliability	Usability	
Functionality	Portability	

Table 13
Relative importance of quality sub-factors

		Importance (%)
1. Attributes of functionality		
1.1	Suitability	
1.2	Accuracy	
1.3	Interoperability	
1.4	Security	
1.5	Trace ability	
2. Attributes of reliability		
2.1	Maturity	
2.2	Fault tolerance	
2.3	Recoverability	
2.4	Accessability	
3. Attributes of usability		
3.1	Understandability	
3.2	Learn ability	
3.3	Operability	
3.4	Attractiveness	
3.5	Navigability	
3.6	Customizeability	
4. Attributes of efficiency		
4.1	Time behavior	
4.2	Resource utilization	
5. Attributes of maintainability		
5.1	Analyzability	
5.2	Changeability	
5.3	Stability	
5.4	Testability	
6. Attributes of portability		
6.1	Adaptability	
6.2	Install ability	
6.3	Co-existence	
6.4	Replace ability	

References

[1] R. Fitzpatrick, Software quality definitions and strategic issues, Technical Paper, Staffordshire University, 1996.
[2] R. Pressman, Software Engineering, a Practitioner's Approach, fifth ed., McGraw-Hill, 2000.
[3] P. Botella, X. Burgues, J.P. Carvallo, X. Franch, C. Quer, Using Quality Models for Assessing COTS Selection, in: Proceeding of MPEC'04 and ICSE'04, 2004.
[4] ISO/IEC 9126-1, Software engineering – product quality – Part 1: Quality Model, first ed.: 2001-06-15.
[5] K. Khosravi, Y. Gueheneuc, A Quality Model for Design Patterns, Technical Report 1249, University of Montreal, 2004.
[6] R.G. Dromey, A model for software product quality, IEEE Transactions on Software Engineering 21 (2) (1995).
[7] A. Stefani, M. Xenos, D. Stavrinoudis, Modelling E-Commerce Systems' Quality with Belief Networks, International Symposium on Virtual Environments, Human-Computer Interfaces, and Measurement Systems, Switzerland, 2003.
[8] A. Stefani, D. Stavrinoudis, M. Xenos, Experimental based tool calibration used for assessing the quality of e-commerce systems, in: Proceedings of the First IEEE International Conference on E-Business and Telecommunication Networks, Portugal, 2004, pp. 26–32.
[9] V. Zwass, Electronic commerce: structures and issues, International Journal of Software Quality 1 (1) (1996).
[10] E. Turban et al., Electronic Commerce: a Managerial Perspective, Prentice Hall, 2004.
[11] E. Murray, D. Jennex, L. Amoroso, e-Business and technology issues for developing economies, Electronic Journal on Information Systems in Developing Countries 10 (2002) 1–14.
[12] A. Alvaro, E. Almeida, S. Meira, Component quality assurance: towards a software component certification process, in: Proceeding of Information Reuse and Integration Conference, IEEE International Conference, 2007, pp. 134–139.
[13] M. Morisio, I. Stamelos, A. Tsoukias, A new method to evaluate software artifacts against predefined profiles, in: Proceeding of SSEKE'02, ACM, Italy, 2002.
[14] T. Rosqvist, M. Koskela, H. Harju, Software quality evaluation based on expert judgement, Software Quality Journal 11 (2003) 39–55.
[15] M. Lorenz, J. Kidd, Object-Oriented Software Metrics, Prentice Hall, 1994.
[16] E. Dustin, J. Rashka, D. McDiarmid, Quality Web Systems, Addison-Wesley, 2001.
[17] J. Kim, J. Lee, Critical design factors for successful e-commerce systems, Behavioral and Informational Technology 21 (3) (2002) 185–199.
[18] D. Selz, P. Schubert, Web assessment: a model for the evaluation and the assessment of successful electronic commerce applications, EM-Electronic Markets 7 (1997) 46–48.
[19] C. Liu, K. Arnett, Exploring the factors associated with web site success in the context of electronic commerce, Information and Management 38 (2000) 23–33.
[20] E. Georgiadou, A generic, multilayered, customizable, software quality model, Software Quality Journal 11 (2003) 313–323.
[21] C. Calero, J. Ruiz, M. Piattini, A web metrics survey using WQM, in: Proceeding of ICWE04, 2005.
[22] R. Hower, Beyond broken links, internet systems, DBMS Magazine, 1997.
[23] D. Meyerhoff, D. Huberty, Testing Web-Based Home Banking Applications: Software Quality, Testing, and Tools, Springer, 2001. pp. 200–208.
[24] H.Q. Nguyen, Testing Applications on the Web: Test Planning for Internet-Based Systems, Wiley, 2001.
[25] A. Rudolf, R. Pirker, E-business testing-user perceptions and performance issues, in: Proceeding of the First Asia-Pacific Conference On Quality Software, IEEE, China, 2000.
[26] M.F. Bretoa, J.M. Troya, A. Vallecillo, Measuring the usability of software components, Journal of Systems and Software, 79, Elsevier, 2006, pp. 427–439.
[27] J. Nielson, Usability Engineering, American Press, 1993.
[28] J.I. Kima, D.L. Shunkb, Matching indirect procurement process with different B2B e-procurement systems, Computers in Industry Journal 53 (2004) 153–164.
[29] B. Khorami, B. Sohrabi, M. Akbari, H.R. Jalali, Studying of e-Business solutions, Institute For Trade Studies and Research Publishing, Iran, 2005.
[30] B. Mahadevan, Emerging market mechanisms in business-to-business e-commerce: a framework, International Conference on Advances in Infrastructure for e-Business, e-Education, e-Science, and e-Medicine on the Internet, Italy, 2002.
[31] O. Signore, Towards a quality model for websites, in: Proceeding of ACM International Conference, Poland, 2005.
[32] D. Han, I. Han, Prioritization and selection of intellectual capital measurement indicators using AHP for the mobile telecommunications industry, Expert Systems with Applications Journal 26 (2004) 519–527.
[33] M.J. Hilla, R. Braatenb, S.M. Veitchc, B.G. Leesd, S. Sharma, Multi-criteria decision analysis in spatial decision support, Environmental Modelling and Software Journal 20 (2005) 955–976.
[34] A. Salo, A. Punkka, Rank inclusion in criteria hierarchies, European Journal of Operational Research 163 (2005) 338–356.
[35] P. Davoli, F. Mazzoni, E. Corradini, Quality assessment of cultural websites with fuzzy operation, Journal of Computer Information Systems (Fall) (2005) 44–57.
[36] A. Pekin, G. Ozkan, O. Eski, U.Karaarslan, G. Ertek, K. Kilic, Application of the AHP for Selection of Forecasting Software, Fifth International Symposium on Intelligent Manufacturing Systems (IMS'06), Turkey 2006.
[37] J. Cochran, H. Chen, Multi-criteria selection of object-oriented simulation software for production system analysis, Computers and Operations research Journal, 32, Elsevier, 2005, pp. 153–168.
[38] F. Tiryaki, M. Ahlatcioglu, Fuzzy stock selection using new fuzzy ranking and weighting algorithm, Applied Mathematics and Computation Journal, 170, Elsevier, 2005, pp. 144–157.
[39] C. Calero, J. Ruiz, M. Piattini, Classifying web metrics using the web quality model, Online Information Review Journal 29 (3) (2005) 645–661.
[40] ISACO Portal is available at: <www.isaco.ir>.
[41] S. Barnes, R. Vidgen, Assessing the quality of auction web sites, in: Proceeding of 34th HICSS01, Hawaii, 2001.
[42] L. Mikhailov, P. Tsvetinov, Evaluation of services using a fuzzy analytic hierarchy process, Applied soft computing Journal, 5, Elsevier, 2004, pp. 23–33.
[43] G. Malak, L. Badri, M. Badri, H. Sahraoui, Towards a multidimensional model for web-based applications quality assessment, in: Computer Science Journal, Springer, 2004, pp. 316–327. Issue 3182.