A Multicriteria Approach for Assessing the Quality of Information Technology Support Services

André Luis Policani Freitas
North Fluminense State University – Production Engineering Laboratory
Av. Alberto Lamego, 2000 – Campos dos Goytacazes/RJ, Brasil – CEP 28013-620

Cláudia Márcia Ribeiro Machado Albernaz
North Fluminense State University – Production Engineering Laboratory
Av. Alberto Lamego, 2000 – Campos dos Goytacazes/RJ, Brasil – CEP 28013-620

Abstract

Information Technology Support Services (ITSS) have played an important role to the companies which wish for more efficiency and competitive advantages. For doing so, the starting point is to identify the real stage the services are on by assessing the quality of services that have been provided. Recent studies show that an effective service analysis of the services produced by IT department should take into account how users perceive IT services. This work presents an alternative approach for helping managers of IT service companies to investigate and solve users' non-satisfaction problems. Using Multiple Criteria Decision Aid (MCDA) methods (ELECTRE TRI and Non-weighted Average), individual IT users' satisfaction degree concerning several criteria were used to assess and to classify the quality of IT support services. A case study was conducted in order to analyze the application of the proposed approach to the assessment of quality of IT support service provided by a Federal Education Institute in Brazil.

Keywords: Information Technology Support Services; ELECTRE TRI; user satisfaction

1. Introduction

Over the course of the last two decades, the emergence of new information technologies (IT) is playing an increasingly important role in the worldwide economy. The importance of using of modern information technology service is widely known and it has changed the way that companies conduct their daily operations.

These services include those that enable communication and collaboration (i.e. email, desktop videoconferencing, instant messaging), data capture (i.e. Internet-based data entry systems, business intelligence, customer portals), processing (i.e. order processing, invoicing, contract management, account management), storage (i.e. data centers and databases with information about customers, inventories, assets, etc.), access (i.e. *ad hoc* queries, report writing), and analysis (i.e. analytics, modeling) (PEPPARD, 2003).

The traditional performance measures, which are mostly financial, have been used to assess the performance of IT departments. Unfortunately, these traditional measures were found to have some serious shortcomings when used to measure the service performance of IT departments (Watson et al, 1993; Kang and Bradley, 2002).

On the other hand, an effective service analysis of the services produced by IT division for other organizational divisions, or IT client divisions, should take into consideration how these clients perceive IT services (Roses et al, 2009). The concept of service quality as a measure of performance originated from the field of marketing, which proposes that there is a need for organizations to understand and measure customer expectation (Kang and Bradley, 2002). In this context, Costa *et al* (2007) reported that one of the most commonly used procedures to evaluate and classify the quality of services is by measuring the user's degree of satisfaction concerning a set of relevant criteria. The concept of using service quality to measure the performance of IT service providers is due to the acknowledgement that the service performance can be determined by the perception of users of IT services.

Furthermore, customers actually not only demand quality products obtained from mature processes, but they also require quality in the services they receive. In recent years, while companies have been deploying their software development processes, there has been an on-going demand for better IT services (Mesquida et al, 2012).

Despite the existence of some best practices in sector management (for example, ITIL and COBIT), such practices do not explicitly inform how the process of assessing the quality of IT services should be performed, or which dimensions and criteria should be taken into account. For example, the ITIL IS Service Management framework has become an established platform for designing and managing IST services in an increasing number of organizations, but it focuses solely on the services around IT operations, such as security management, service reporting, availability and contingency management, release management and incident management, which are only part of the overall portfolio of IST services (PEPPARD, 2003).

This work presents an alternative approach for helping managers of IT service companies to investigate and solve users' non-satisfaction problems. Using Multiple Criteria Decision Aid (MCDA) methods (ELECTRE TRI and Non-weighted Average), varying individual IT users' satisfaction degree concerning several criteria were used to assess and to classify the quality of IT support services. A case study was conducted in order to analyze the application of the proposed approach to the assessment of quality of IT support service provided by a Federal Education Institute in Brazil.

2. Information Technology Support Services

Technical support is a service that the Information System department provides to the users (Bharati and Berg, 2003). However, according to Cronin and Taylor (1992), one of the most complex actions concerns in how to assess service quality in a proper and accurate way. In general, managers and administrators of services companies need to identify some important issues, such as: (a) what elements (or dimensions) of a particular service best define its quality; (b) what elements (or dimensions) of a particular service with which the customers are most (or less) satisfied, and (c) the implications (or effects) of service quality and customer satisfaction in the purchase intentions.

In spite of the numerous considerable published works on the service quality field all over the world, there are still some lacks of consensus between the managers and marketing researches concerning (Freitas & Costa, 2012):

a) the real meaning of "service quality". The only existent consensus is that service quality is still an elusive and abstract construct that is difficult to define and measure (Parasuraman et. al., 1985, 1988; Carman, 1990; Cronin and Taylor, 1992). Specially, these difficulties can be associated to the following aspects:

   - the enormous variety of services (heterogeneity) and the strong relationship with human behaviors and attitudes make difficult the evaluation and the standardization tasks;
   - services are consumed almost simultaneously as they are provided, which make more difficult or incapacitating the fails detection and correction before they occur;
   - to measure service quality, conventional measurement tools are supported on cardinal or ordinal scales. Most of the criticism about scale based on measurement is that scores do not necessarily represent user preference. This is because evaluators have to internally convert their preference to scores and the conversion may introduce distortion of the preference being captured (Tsaur et al., 2002), and;
   - since service quality consists of not only tangibles criteria (attributes), but also intangible/subjective criteria, the evaluation process is often characterized by uncertainties and imprecision which are increased when multiple evaluators are involved.

b) the comprehension of the relationship between the service quality *vs.* customer satisfaction constructs (see Bitner, 1990; Bolton and Drew, 1991a,b; Cronin and Taylor, 1992; Rust and Oliver, 1994) and such implications as customer fidelity, positive word of mouth, and (re)purchase intentions. In this aspect, service quality has been initially identified as an antecedent to satisfaction (Cronin and Taylor, 1992; Parasuraman et al, 1988). On the other hand, some researchers (Bolton and Drew, 1991a,b; Bitner, 1990) suggested that satisfaction was an antecedent of service quality. Finally, some researchers suggested that neither satisfaction nor service quality might be an antecedent to the each other (Dabholkar, 1995; McAlexander et al., 1994; Cronin and Taylor, 1992), and;

c) the most adequate way to measure service quality. In this context, the SERVQUAL scale (Parasuraman et al., 1985, 1988) has been the dominant and traditional technique to measure service quality. Supported on the gap theory, the SERVQUAL scale suggests that service quality should be measured by the difference (or gap) between customer's expectations about some specific service and the actual service performance (i.e., the disconfirmation performance). On the other hand, Cronin and Taylor (1992) argued that if "service quality is considered similar to an attitude" (as proposed by Parasuraman et al, 1985, 1988), service quality measure could be better represented by an attitude-based conceptualization. Therefore, they suggested that the expectations scale (SERVQUAL scale) should be discarded in favor of a performance-only measure of service quality, that they named SERVPERF.

Since the satisfaction construct can be related to the positive (or negative) performance of a service company (e.g. Bolton and Drew, 1994; Rust and Oliver, 1994; Oliver, 1997), therefore there exists a great dependence between service quality and customer satisfaction, and an increase in one is likely to lead an increase in another (e.g. Sureshchandar et al., 2002). In addition to the emphasize strategies centering only in the provision of quality services, managers of service organizations need to consider all of the possible determinants that have greatest positive influence on the satisfaction of their current and potential customers.

Although services generally have the same characteristics, IT services have peculiarities that make their evaluation even more subjective. According to Peppard (2003), IT services are more or less intangible, i.e., they are generally something one cannot touch or feel, although they may be associated with something physical, such as the provision of information through a computer screen or personal digital assistant. While an IT service may have a predominantly physical outcome, for example the delivery and installation of a PC or the provision of a cable for a network connection, there are other services that may be totally intangible, such as advice and support from a help desk, IT training, consultancy, systems design, or upgrading server software. The information handling services provided via applications are also intangible but they do require a physical platform in order to exist.

Many IT services are produced and consumed simultaneously. For instance, support from a helpdesk is generally provided and utilized immediately. The consequence of this is that a *bad service* cannot be perceived and avoided before it has been received by the user. A bad experience can impact the perception that the user will have the next time he uses the same service. Then, the need arises to evaluate service quality concerning the user perspective.

Nowadays, the usage of IT best practices is becoming more and more common. Several studies have focused on the adoption of IT Service Management (ITSM) as well as a specific service oriented best practices (Marrone and Kolbe, 2011). The IT Governance Institute (2011) estimates that ISO 20000 or ITIL is the external framework most frequently mentioned as a basis for an enterprise's Governance of Enterprise IT (GEIT) approach with 28%. The second most commonly cited external framework or standard on which an enterprise bases its GEIT approach is ISO 17799/ISO 27000, the Information Security Framework or other security standards with 21.1%. Control Objectives for Information and related Technology (CobiT) is being used by 12 percent of the respondent enterprises.

These practices in IT management clearly specify the need to ensure quality services and this requires an assessment closer to the reality of the organization. However, the guidelines of these models do not show how this evaluation should be done, much less than the dimensions and criteria that should be used in the evaluation process.

Due to the importance of measuring the quality of IT support services, this work presents an alternative approach for helping managers of IT service organizations. Supported on the Multiple Criteria Decision Aid (MCDA) methods, the main objective of this approach is to be specific to assess the quality of IT support services by means of the measurement of the users' satisfaction degree concerning several criteria but, simultaneously suitable to be applicable to any kind of company which has an IT infrastructure.

3. MCDA and Multicriteria sorting problem

Multi-criteria Decision AID (MCDA) is a branch of Operations Research that aims to give the decision maker some tools in order to enable him to advance in solving a decision problem where several - often contradictory - criteria and points of view should be taken into account (VINCKE, 1992).

Decision-aiding (DA) seeks to put science in the service of shedding light on managerial decisions and/or guiding complex decision-making process within organized systems. As a corollary, because DA helps to construct, and not simply to describe, it should give pride of place to a dynamic approach facilitating easy insertion of DA practitioners into the decision-making process. In some cases, DA can thus contribute to legitimating the final decision (ROY, 1999).

According to Vincke (1992), when dealing with a multicriteria problem is that there does not exist, in general, any action or solution which is the best simultaneously in all criteria. Therefore, the term "optimization" should no longer be used in such a context because, contrary to the classical techniques of Operations Research, multicriteria methods do not objectively provide optimal solutions. The solution should be satisfactory, subject to a mathematical logic, with quality and acceptable cost. Therefore the use of the term "aid" seems to be more convenient.

When dealing with this type of problem is that there does not exist, in general, any action or solution which is the best simultaneously in all criteria. Therefore, the term "optimization" should no longer be used in such a context because, contrary to the classical techniques of Operations Research, multicriteria methods do not objectively provide optimal solutions. The solution should be satisfactory, subject to a mathematical logic, with quality and acceptable cost. Therefore the use of the term "aid" seems to be more convenient.

The multiple criteria sorting/classification problem is a decision problem which besides evaluating a finite set $\underline{\boldsymbol{X}} = \{X_1, ..., X_n\}$ of alternatives (or actions) concerning a set of criteria, it also requires to assign these alternatives to one of the predefined categories (C_1, C_2, …, C_t) (Roy, 1985). The assignment of an alternative X_i to one of the categories should rely on the comparison of the intrinsic value of X_i to specific reference points that define each category (and not on the comparison of X_i to other alternatives). Roughly speaking, this kind of problem occurs when service companies are evaluated concerning a set of criteria and, according to their performances, these companies are qualified into some specific categories.

Even though both classification and sorting terms purposes the assignment of alternatives into categories, they represent two slightly distinct problems (Zopounidis and Doumpos, 2002). Classification problems concern to the case where the categories are defined in a nominal way (Fig. 1a). On the other hand, sorting problems concern to the case where the categories are defined in an ordered way starting from those including the best alternatives to those including the worst ones (Fig. 1b). In this case, C_1 will denote the category that contains the most preferred alternatives, while C_t will denote the category containing the least preferred alternatives.

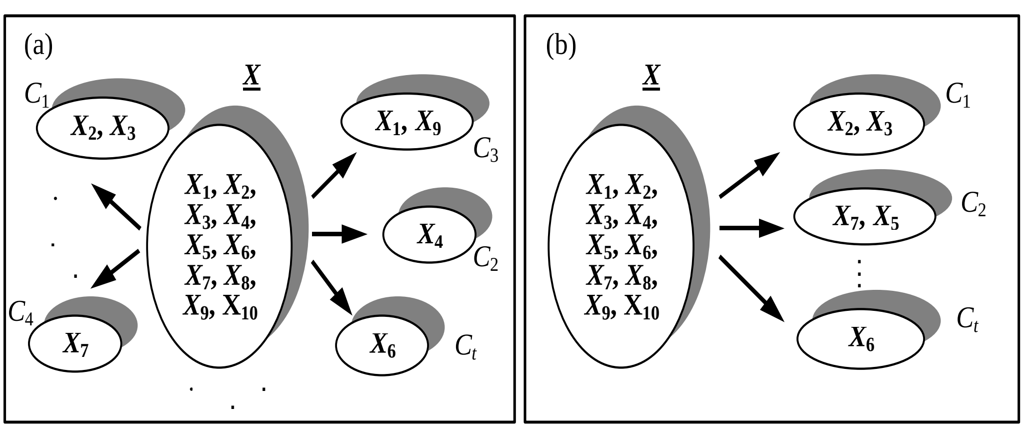

Figure 1: The multiple criteria classification/sorting problems

However, Mousseau et al. (2000) distinguish the Multiple Criteria Sorting Problems (MCSP) into two groups: *nominal* MCSP, where the predefined categories are not ordered (Fig. 1a) and *ordinal* MCSP, where the categories are necessarily ordered (Fig. 1b). Additionally, MCSP differ from clustering methods, where the categories are unknown a priori (the categories result from the analysis).

On the other hand, Léger and Martel (2002) established that the classification methods are divided into two groups: the assignment methods and the "clustering methods". The assignment methods are associated to the sorting problematic and they are based on the notion of supervised learning. Under supervised learning, previous information is available and it serves to calibrate the parameters of the method. Thus, the method is able to assign a general alternative to the most appropriate category. On the other hand, the clustering methods consist in regrouping alternatives into a restricted number of categories. These methods are based on the notion of non-supervised learning. In this case, the alternatives are known, but no information about their belonging to the categories is available.

Considering the aforementioned definitions, there are still contradictions between the researchers concerning the concepts of *classification* and *sorting*. Thus, a just preventive recommended attitude is to be careful with the use of the term "classification". In this sense, the term *classification* will be used in this work in a more generalist view (as proposed by Léger and Martel, 2002). However, the main objective is to assess the service quality classification through the ordinal MCPS analysis.

3.1 Sorting and classification methods

Several methods, techniques and models have been developed for studying classification/sorting problems and its applications in a variety of research fields, including marketing, finance, medicine, pattern recognition, education, etc. Such variety of real-world applications has been a great motivation for the researchers in developing methodologies for constructing classification/sorting models.

In particular, the development of multidimensional classification models started on the linear discriminant analysis (Fischer, 1936) that was later extended to the quadratic form by Smith (1947). The most recent research in developing classification/sorting models and methods is focused on operations research and artificial intelligence techniques. Methodologies such as neural networks, machine learning, rough sets, fuzzy sets and MCDA are considered by researchers both at the theoretical and practical levels (ZOPOUNIDIS & DOUMPOS, 2002).

The Weighted Average method and the ELECTRE TRI method (YU, 1992) are some of the most used MCDA methods for dealing with multicriteria classification problems.

3.1.1 The Weighted Average Method

The Weighted Average method is one of the elementary multicriteria methods. According to Vincke (1992), elementary methods are those that immediately come to mind when decision maker is confronted with a multicriteria aggregation problem. They are commonly used in practice. However, due to their simplicity, some important aspects of the problem may be not revealed. The weighted average method involves the construction of a global preference structure as represented in (1):

$$\begin{cases} X_1 P X_2 \text{ if and only if } \sum_{j=1}^{n} p_j g_j(X_1) > \sum_{j=1}^{n} p_j g_j(X_2) \\ X_1 I X_2 \text{ if only if } \sum_{j=1}^{n} p_j g_j(X_1) = \sum_{j=1}^{n} p_j g_j(X_2) \end{cases} \quad \textbf{(1)}$$

In such a case, 'weights' p_j represent substitution rates between criteria. By using this method, the decision maker assumes that all criteria must be expressed in identical units and that the differences between values on different criteria can compensate each other. The preference relations X_1PX_2 (X_1 *is preferred to* X_2) e X_1IX_2 (X_1 and X_2 are indifferent) are obtained by comparing the weighted average values of each alternative. The alternative of greater weighted average value must be chosen. A particular situation occurs when the decision maker does not want or does not consider himself able to define the importance of criteria.

In such case, only the average of the points of each alternative is computed (Non-weighted average method). The alternative of greater average value must be chosen.

For classification purposes, in both situations assignment rules are generally used to compare the average value of each alternative to reference points that defines the frontiers of the predefined categories. Table 1 shows the reference points and the assignment rules to assign a generic alternative X_1 into one of the predefined categories.

Table 1: – Reference Points

Categories	Reference points
C_1	$\sum_{j=1}^{n} p_j g_j(X_1) > 5{,}0$
C_2	$4{,}0 < \sum_{j=1}^{n} p_j g_j(X_1) \leq 5{,}0$
⋮	⋮
C_p	$1{,}0 < \sum_{j=1}^{n} p_j g_j(X_1) \leq 2{,}0$
C_{p+1}	$\sum_{j=1}^{n} p_j g_j(X_1) < 1{,}0$

3.1.2 The Electre Tri Method

In the MCDA context, one of the most widely used criteria aggregation models is the outranking relation model. Firstly defined by Roy (1985), the outranking concept can be schematized as follows: an alternative X_1 outranks an alternative X_2 ($X_1\mathbf{S}X_2$) if, given the information about the preferences of the decision maker, there are sufficient arguments to affirm that X_1 is at least as good as X_2 an there is no really important reason to refuse this assertion.

On the contrary to the classical multicriteria aggregation models, the outranking approach is based on preference models including incomparabilities and it does not impose any transitivity properties. As reported by Vincke (1992), incomparability can be interpreted in a significant way: when two alternatives are incomparable, it points out the conflicts or lack of information and invites the analyst and the decision maker to insert (*if they want to do so*) into some aspects of the problem.

Several outranking methods can be found in the literature, differing from each other in the way they formalize this definition (Vincke, *op. cit*.). TRICHOM (Moscarola, 1977), the N-TOMIC methods (Massaglia and Ostanello, 1991), Perny's method (Perny, 1998), PROAFTN (Belacel, 2000), and more recently, the TRINOMFC method (Léger and Martel, 2002) and the IRIS method (Dias and Mousseau, 2003) are some of these methods. However, the ELECTRE TRI method (Yu, 1992) is the most widely used MCDA sorting method based on the outranking relations approach (ZOPOUNIDIS & DOUMPOS, 2002).

In ELECTRE TRI categories are ordered. Let $\underline{\mathbf{X}}$ =$\{X_1, ..., X_m\}$ denote a set of alternatives (or actions) evaluated concerning a set of n criteria $g_1, g_2, \ldots, g_n$, and $C = \{C_1, C_2, \ldots, C_h, \ldots C_k\}$ denote the set of ordered categories from the worst or lowest (C_1) to the best or highest (C_k). Each category must be characterized by a lower and an upper profile. The assignment of an alternative X_i into a certain category C_h results from the comparison of X_i to the profiles defining the lower and the upper limits of the categories; b_h being the upper limit of category C_h and the lower limit of category C_{h+1} ($h = 1, 2, \ldots, k$).

ELECTRE TRI builds an outranking relation S, i.e., validates or invalidates the assertion X_iSb_h (and b_hSX_i), whose meaning is "X_i is at least as good as b_h". The indifference and preference thresholds ($q_j(b_h)$ and $p_j(b_h)$) constitute the intra-criterion preferential information. They account for the imprecise nature of the evaluations $g_j(X_i)$. $q_j(b_h)$ specifies the largest difference $g_j(X_i)$ - $g_j(b_h)$ that preserves indifference between X_i and b_h on criterion g_j; $p_j(b_h)$ represents the smallest difference $g_j(X_i)$ - $g_j(b_h)$ compatible with a preference in favor of X_i on criterion g_j.

At the comprehensive level of preferences, in order to validate the assertion X_iSb_h (or b_hSX_i), two conditions should be verified:

- concordance: for an outranking X_iSb_h (or b_hSX_i) to be accepted, a "sufficient" majority of criteria should be in favour of this assertion,

- non-discordance: when the concordance condition holds, none of the criteria in the minority should oppose to the assertion X_iSb_h (or b_hSX_i) in a "too strong way".

Two types of inter-criteria preference parameters intervene in the construction of S:

- the set of weight-importance coefficients (k_1, k_2, ..., k_n) is used in the concordance test when computing the relative importance of the coalitions of criteria being in favour of the assertion X_iSb_h,

- the set of veto thresholds ($v_1(b_h)$, $v_2(b_h)$, ..., $v_n(b_h)$), $\forall\, h$ (h = 1, 2, …, k) nis used in the discordance test. $v_j(b_h)$ represents the smallest difference $g_j(b_h)$ - $g_j(X_i)$ incompatible with the assertion X_iSb_h.

Generally, the outranking relation is constructed through the following steps (see details in Yu (1992) and Mousseau, Figueira and Naux (2001):

(i) to compute the partial concordance indices $c_j(X_i, b_h)$ and $c_j(b_h, X_i)$ which respectively expresses the extension that "X_i outranks the profile b_h" and "the profile b_h outranks X_i", considering the criterion g_j.

(ii) to compute the global concordance indices $C(X_i, b_h)$ and $C(b_h, X_i)$ which respectively expresses the extension that "X_i outranks the profile b_h" and "the profile b_h outranks X_i", considering all criteria.

(iii) to compute the partial discordance indices $d_j(X_i, b_h)$ and $d_j(b_h, X_i)$ which respectively expresses to which extend the criterion g_j is opposed to the statement "X_i outranks the profile b_h" and "the profile b_h outranks X_i".

(iv) to compute the outranking relation based on the indices $\sigma(X_i, b_h)$ and $\sigma(b_h, X_i)$ which respectively expresses the extension that "X_i outranks the profile b_h" and that "the profile b_h outranks X_i", according to the global concordance indices and the discordance indices.

(v) to define a λ-cutting level (λ∈[0.5, 1]) in order to determine the preference relation between X_i and b_h:

- X_i is preferred to a profile b_h (X_i **P** b_h) if $\sigma(X_i, b_h) \geq \lambda$ and $\sigma(b_h, X_i) < \lambda$;
- X_i and b_h are indifferent (X_i **I** b_h) if $\sigma(X_i, b_h) \geq \lambda$ and $\sigma(b_{h\,r}, X_i) \geq \lambda$;
- X_i and b_h are incomparable (X_i**R** b_h) if $\sigma(X_i, b_h) < \lambda$ and $\sigma(b_h, X_i) < \lambda$, and
- b_h is preferred to X_i (b_h **P** X_i) if $\sigma(X_i, b_h) < \lambda$ and $\sigma(b_h, X_i) \geq \lambda$.

Two assignment procedures are available:

The pessimistic procedure: such procedure compares X_i successively to b_i (i = k, k-1, …,0). b_h being the first profile such that X_i **S** b_h, this procedure assign X_i to category C_{h+1}. On the other hand, the optimistic procedure compares X_i successively to b_i, (i = 1, 2, …, k+1). b_h being the first profile such that b_h **P** X_i, this procedure assign X_i to category C_h.

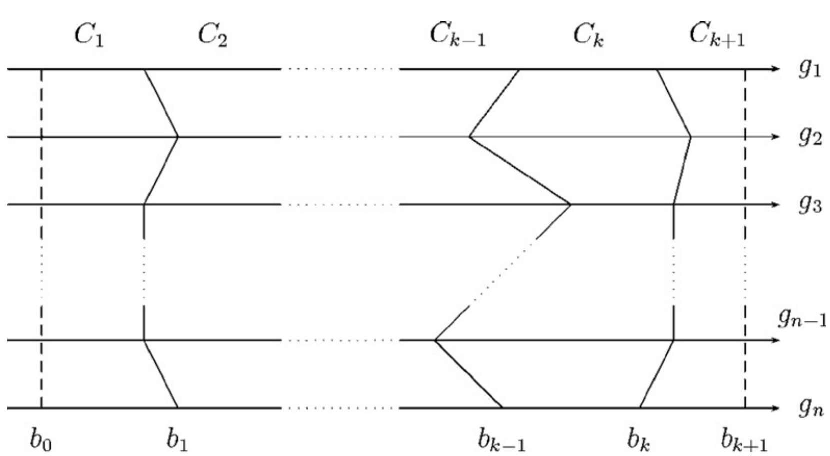

Figure 2. Definition of categories through limit profiles (Mousseau, Slowinski and Zielniewikz, 2000)

The description and understanding of the ELECTRE TRI classification algorithm require extra effort, especially by the fact that this method is based on recent concepts of fuzzy logic. Despite this, the understanding and modeling problems by means of ELECTRE TRI excuse the decision maker from the detailed description of the classification algorithm (COSTA & FREITAS, 2005).

4. Case Study

The study investigates the application of the proposed model in the evaluation of support services provided by the IT department located in the main campus of a Federal Education Institute. Such Institute currently has about 8,000 students and it offers courses of different levels: High School, Technical, Technology, College and Postgraduate. Services are performed online support (help desk) and *in loco*, highlighting the following activities: replacement of components, install/restore applications and support for physical and logical network.

The set of evaluators was composed of faculty, staff and students who used the IT support service during a specific period of analysis. The Dimensions and criteria were supported on service quality literature and on the Moments of Truth of the IT support Service Cycle. According to Albrecht and Bradford (1998), each Moment of Truth is the moment when the customer comes into contact with any aspect of the organization (staff, facilities, telephone / fax, etc.). And, according to the contact he can form his opinion about the quality of service. Upon receiving a service, the client goes through a sequence of *Moments of Truth,* called *Service Cycle.* Thus, through the understanding of this cycle and the moments of truth, any faults that occur can be more easily identified and, by taking corrective/preventive actions, these faults can be avoided in order to provide a better service (See Table 2).

The evaluators were not IT experts. Then we suppose they were not able to make technically reliable judgments regarding the importance of the criteria. More than this, it’s strongly recommended the elicitation of the importance of the criteria if the evaluators are really able to do so. Under this circumstance, the non-weighted average method was considered and all criteria had the same importance in order to use ELECTRE TRI.

However, since the evaluators made their judgments concerning the experience with the IT support services they received, such circumstance do not represent a restriction to the use of the proposed approach. It is noteworthy this approach allows the user to score the 'NA (Not Available)' option if he does not have experienced the service or he does not want to evaluate the service on some criterion.

Table 2 - Dimensions of Quality and criteria used in the study

Dimensions	*Criteria*
D1- Tangibility	$Cr_{1.}$ Appropriateness and usefulness of equipment and tools used by IT technician.
	$Cr_{2.}$ Appropriateness and timeliness of the software used by IT technician.
	$Cr_{3.}$ Physical facilities of the IT service sector.
	$Cr_{4.}$ The procedure to request a service.
	$Cr_{5.}$The appearance of the professionals.
	Cr6. Number of employees for a quick service.
D2 - Reliability	$Cr_{7.}$ Provision of the service as promised.
	$Cr_{8.}$ The service correctly performed at first time.
	$Cr_{9.}$ Professional interest in solving problems.
	$Cr_{10.}$ Accuracy of information provided about the service provided.
	$Cr_{11.}$ Written records about the entire service.
	$Cr_{12.}$ Compliance with the promised time to perform the service.
D3 - Responsiveness	$Cr_{13.}$ Information about when the service will be done.
	Cr_{14}. Immediate care of the request.
	$Cr_{15.}$ Availability of technicians to serve requests.
	Cr_{16}.Waiting time to be served.
	Cr_{17}.Speed of service.
	Cr_{18}.Runtime service.
D4 - Warranty	Cr_{19}. Confidence inspired by the IT technician.
	Cr_{20}. Self-assurance of the IT technician during the service.
	Cr_{21}. IT technician competencies for execution of the service.
	Cr_{22}. Easy to get in touch with the IT technician.
	Cr_{23}.Easy in solving problems.
D5 - Empathy	Cr_{24}. Service customization in order to meet customer needs.
	Cr_{25}. Education of IT technician during the service.
	Cr_{26}. Knowledge of the IT technician about the client's business.
	Cr_{27}. IT technician care during the service request.
	Cr_{28}. IT technician understanding about the specific customer needs.
	Cr_{29}. Self-assurance of the IT technician to negotiate service issues.
	Cr_{30}. Courtesy of the IT technician.

A combination of the graphic rating scale and the itemized rating scale was used to measure the users' satisfaction degree concerning each criterion. The satisfaction responses vary from 'Very dissatisfied (value 0)' to 'Very satisfied (value 10)'. According to Parasuraman, Grewal and Krishnan (2004), the resulting scale gains the benefits of both types of scales: an itemized rating scale should be easier to respond to and more meaningful from the respondent's perspective and a graphic rating scale allows detection of fine shades of differences in attitudes.

For both MCDA seven categories were considered in this study. Table 3 shows the categories concepts and the boundaries that delimit these categories. Table 3 also presents the rules for assigning a generic alternative X_i into one of the predefined categories regarding the Average Satisfaction Degree.

Table 3 – Categories and Limits

Categories	Limit	Lower Limit	Upper Limit
A (Excellent)	$\overline{SD}(X_i) > 9.0$	(9.0, 9.0, ..., 9.0, 9.0)	---
B^{+} (Very Good)	$8.0 < \overline{SD}(X_i) \leq 9.0$	(8.0, 8.0, ..., 8.0, 8.0)	(9.0, 9.0, ..., 9.0, 9.0)
B^{-} **(**Good)	$7.0 < \overline{SD}(X_i) \leq 8.0$	(7.0, 7.0, ..., 7.0, 7.0)	(8.0, 8.0, ..., 8.0, 8.0)
C (Regular)	$6.0 < \overline{SD}(X_i) \leq 7.0$	(6.0, 6.0, ..., 6.0, 6.0)	(7.0, 7.0, ..., 7.0, 7.0)
D^{+} (Bad)	$5.0 < \overline{SD}(X_i) \leq 6.0$	(5.0, 5.0, ..., 5.0, 5.0)	(6.0, 6.0, ..., 6.0, 6.0)
D^{-} (Very Bad)	$4.0 < \overline{SD}(X_i) \leq 5.0$	(4.0, 4.0, ..., 4.0, 4.0)	(5.0, 5.0, ..., 5.0, 5.0)
E (Terrible)	$\overline{SD}(X_i) < 4.0$	---	(4.0, 4.0, ..., 4.0, 4.0)

According to the ELECTRE TRI, the **indifference** (q_j)**, preference** (p_j) **and veto** (v_j) thresholds are used to take into account the fact that the values of the performances may be arguable because of ill-determined, uncertain or imprecise factors. Considering that such thresholds are related to the variability of the judgments made by the users, the preference threshold and indifference thresholds for each criterion were defined regarding the respective coefficient of variation (CV): $p_j = (CV)_j$ and $q_j = (CV)_j/2$. Table 4 shows these coefficients.
The veto thresholds were not considered. The λ – cutting level was 0.76.

Table 4 – Preference and indifference thresholds.

Criteria	p_i = CV	q_i = CV/2	Criteria	p_j = CV	q_j = CV/2	Criteria	p_j = CV	q_j = CV/2
Cr_1	0.24	0.12	Cr_{11}	0.38	0.19	Cr_{21}	0.32	0.16
Cr_2	0.45	0.23	Cr_{12}	0.33	0.16	Cr_{22}	0.17	0.09
Cr_3	0.53	0.27	Cr_{13}	0.44	0.22	Cr_{23}	0.27	0.13
Cr_4	0.22	0.11	Cr_{14}	0.51	0.25	Cr_{24}	0.25	0.13
Cr_5	0.28	0.14	Cr_{15}	0.44	0.22	Cr_{25}	0.17	0.09
Cr_6	0.50	0.25	Cr_{16}	0.35	0.17	Cr_{26}	0.24	012
Cr_7	0.33	0.17	Cr_{17}	0.32	0.16	Cr_{27}	0.17	0.09
Cr_8	0.30	0.15	Cr_{18}	0.32	0.16	Cr_{28}	0.25	0.12
Cr_9	0.22	0.11	Cr_{19}	0.22	011	Cr_{29}	0.18	0.09
Cr_{10}	0.25	0.12	Cr_{20}	0.29	0.15	Cr_{30}	0.20	0.10

In order to compose the sample of respondents, a report generated by the IT management system informed the users of IT support service on the last 40 days. We consider that during this period the users keep in mind the service they have received. Then the researcher started looking for the users in their workplaces. The original intention is that the users answered the survey as soon as they received the questionnaire. However, most of the time the user was not at the workplace or that moment was not propitious to answer the survey.

Therefore, 38 questionnaires were delivered to the users in the workplaces and they were collected until the next 24 hours. The users returned 31 questionnaires and 29 were valid for analysis. The data were tabulated in a spreadsheet in order to perform the analysis and interpretation of data.

4.1 Service Quality Classification According To The User's Perspective

Considering the ELECTRE TRI results, 19 incomparability relations were registered (See Table 5). Incomparability relations occur when an alternative is assigned to different categories according to the two assignment procedures: pessimistic and optimistic. Divergence exists among the results of the two assignment procedures only when an alternative is incomparable to one or several profiles. In such case the pessimistic assignment rule assigns the alternative to a lower category than the optimistic one. When the evaluation of an alternative are between the two profiles of a category on each criterion, then both procedure assign this alternative to this category (MOUSSEAU, SLOWINSKI & ZIELNIEWICZ, 2000).

According to Costa e Freitas (2005), several different assignment results can indicate the existence of inconsistencies in modeling the classification problem, which must be revised if more reliable results are needed. Importantly, this indicator is not available through traditional classification methods, as the weighted average method. It is suspected that very discrepant performances of an alternative on different criteria may also contribute to the existence of incomparability relations - this finding is supported on the analysis of judgments which resulted in assignment into different and non-subsequent categories (in this study, the evaluations of $A_{5,}$ $A_{14,}$ $A_{15,}$ $A_{21,}$ $A_{25,}$ A_{27} and A_{29} correspond to this finding).

Comparing the results obtained by the Non-weighted Average method and by ELECTRE TRI, it was found that in 9 cases (31.0%) the result was the same between them. However, the majority of results were at least equivalent for the Non-weighted Average method classification with one of the ELECTRE TRI assignment procedures (72.4%) and some results there was no equivalence between the results (27.6%). These results show that the assignment procedures are distinct and it is not relevant to compare the significance of the results. More specifically, the Non-weighted Average method is based on 'traditional model of preferences' and ELECTRE TRI, the 'double threshold model.'

Table 5 - Results of the classification procedures

	Criteria	Service classification on each criterion		User	Service classification according to the user's perspective			
		Non-weighted average			Non-weighted average		ELECTRE TRI	
		$SD(X)_J$	Category		$SD(X)_i$	Category	Pessimistic	Optimistic
D_1 Tangibility	Cr_1	7.39	B^-	A_1	5.86	D-	C	B-
	Cr_2	6.59	C	A_2	9.26	A	A	A
	Cr_3	5.21	D^+	A_3	7.58	B-	B-	B-
	Cr_4	8.34	B^+	A_4	8.24	B+	B-	B-
	Cr_5	7.53	B^-	A_5	7.55	B-	C	A
	Cr_6	5.87	D^+	A_6	9.63	A	A	A
D_2 Reliability	Cr_7	7.22	B^-	A_7	9.00	A	B+	A
	Cr_8	6.82	C	A_8	9.23	A	B+	A
	Cr_9	7.81	B^-	A_9	7.35	B-	B-	B-
	Cr_{10}	7.14	B^-	A_{10}	4.22	D^+	C	B-
	Cr_{11}	6.91	C	A_{11}	9.05	A	B+	A
	Cr_{12}	7.07	B^-	A_{12}	8.13	B+	B+	B+
D_3 Responsivenes s	Cr_{13}	6.47	C	A_{13}	6.50	C	C	B-
	Cr_{14}	6.12	C	A_{14}	7.22	B-	C	B+
	Cr_{15}	6.09	C	A_{15}	7.13	B-	C	B+
	Cr_{16}	6.54	C	A_{16}	7.07	B-	C	B-
	Cr_{17}	7.13	B^-	A_{17}	7.63	B-	B-	B-
	Cr_{18}	7.03	B^-	A_{18}	7.20	B-	C	B-
D_4 - Warranty	Cr_{19}	7.24	B^-	A_{19}	5.61	D^+	D+	C
	Cr_{20}	7.33	B^-	A_{20}	7.30	B-	B-	B-
	C_{21}	6.84	C	A_{21}	6.21	C	C	B+
	Cr_{22}	8.45	B^+	A_{22}	5.78	D^+	C	B-
	Cr_{23}	7.17	B^-	A_{23}	6.19	C	C	C
D_5 - Empathy	Cr_{24}	7.54	B^-	A_{24}	9.13	A	A	A
	Cr_{25}	8.66	B^+	A_{25}	8.96	B+	B-	A
	Cr_{26}	7.50	B^-	A_{26}	7.68	B-	B-	B+
	Cr_{27}	8.27	B^+	A_{27}	2.24	E	E	B-
	Cr_{28}	7.14	B^-	A_{28}	8.70	B+	B+	A
	Cr_{29}	7.51	B^-	A_{29}	2.77	E	D-	B-
	Cr_{30}	8.19	B^+					

A (Excellent) **B^+ (**Very Good) **B^- (Good)** **C (**Regular) **D^+** (Bad) **D^-** (Very Bad) **E** (Terrible)

From the point of view of the decision maker / IT service manager, the amount of users who assigned the quality of services into each category may be relevant for making operational decisions (See Table 6). For example, if it is possible to identify the user and if his/her experience with the service provided is taken into account, the assignment results can provide additional information that will contribute to improving the quality of service.

Table 6 – Sorting results according to the users' point of view

Categories	Non-weighted Average	Pessimistic	Optimistic
A (Excellent)	20.7% (A_2, A_6, A_7, A_8, A_{11}, A_{24})	10.3% (A_2, A_6, A_{24})	31.0% (A_2, A_5, A_6, A_7, A_8, A_{11}, A_{24}, A_{25}, A_{28})
B$^+$ (Very Good)	13.8% (A_4, A_{12}, A_{25}, A_{28})	17.2% (A_7, A_8, A_{11}, A_{12}, A_{28})	17.2% (A_{12}, A_{14}, A_{15}, A_{21}, A_{26})
B $^-$(Good)	34.5% (A_3, A_5, A_9, A_{14}, A_{15}, A_{16}, A_{17}, A_{18}, A_{20}, A_{26})	24.1% (A_3, A_4, A_9, A_{17}, A_{20}, A_{25}, A_{26})	44.8% (A_1, A_3, A_4, A_9, A_{10}, A_{13}, A_{16}, A_{17}, A_{18}, A_{20}, A_{22}, A_{27}, A_{29})
C (Regular)	10.3% (A_{13}, A_{21}, A_{23})	37.9% (A_1, A_5, A_{10}, A_{13}, A_{14}, A_{15}, A_{16}, A_{18}, A_{21}, A_{22}, A_{23})	6.9% (A_{19}, A_{23})
D$^+$ (Bad)	10.3% (A_{10}, A_{19}, A_{22})	3.5% (A_{19})	-----
D $^-$(Very Bad)	3.5% (A_1)	3.5% (A_{29})	-----
E (Terrible)	6.9% (A_{27}, A_{29})	3.5% (A_{27})	----

4.2 Sorting results on each criterion

ELECTRE TRI algorithm does not provide sorting results when each criterion is only considered (mono-criterion sorting/classification problem). In terms of multi-criteria sorting/classification problem, such algorithm does not provide information about the influence of each criterion on the sorting/classification results.

A relatively simple procedure is to use the Non-weighted Average method and aggregate the judgments of all users into the index $\overline{SD}(X)_j$ (Average Satisfaction Degree on each criterion *j*) and then compare this index to the frontiers limits in order to assign the IT support service quality into one of the predefined categories.

Table 7 shows that the quality of IT support services was more critical on the criteria which resulted in the assignment into categories D^-, D^+ and C. These criteria are associated with providing the support service (information about when the service will be done (Cr_{13}), immediate care of the request (Cr_{14}), availability of technicians to serve requests (Cr_{15}), number of employees for a quick service (Cr_6), written records about the entire service (Cr_{11}), and IT technician competencies for execution of the service (Cr_{21})), and also in relation to the appropriateness and timeliness of the software used by IT technician (Cr_2) and the physical facilities of the IT service sector (Cr_3). Technically speaking, these criteria would be prioritized by the IT manager for corrective and preventive actions in order to improve the quality of services provided.

5. Conclusions

Measure and classify the quality of services - in particular, IT support services - is a complex task. The difficulties associated with this problem are related in part to the evaluation process - especially the lack of agreement among users - and, partly, by choosing the method of classification that makes the results easy to understand for the manager and helps him to prioritize the services that need improvement.

The purpose of this article was to present an alternative approach for helping managers of IT service companies to investigate and solve users' non-satisfaction problems. By means of a case study, it was showed a clearer view of how MCDA methods can complement each other and contribute as a tool for making management decisions and for research. The Non-weighted Average method was used for classifying the quality of the service provided by the organization concerning each criterion - analysis not so easily obtained by the ELECTRE TRI method. On the other hand, the ELECTRE TRI allowed to classify the quality of services under a pessimistic and optimistic perspective and allowed the identification of incomparability relations - analysis not available by means of the Non-weighted Average method. A brief report was presented to the IT manager and he stated he was satisfied with the results.

However, there is one aspect that still needs further investigation: the use of techniques or statistical procedures for elimination of outliers when assessing the quality of services. More specifically, it is questioned whether more scattered values are actually considered outliers or if they really represent the perception of the evaluators about the service received.

This questioning is resulting of the fact that scattered judgments may have different origins, such as: values wrongly assigned by the evaluator; values intentionally assigned by the evaluator, and, judgments arising from the perception of a completely atypical service.

It is worth mentioning that a service is never exactly provided alike (different attendants/IT technicians, different customers/IT users, different times, etc.) making more complex the activity of standardization and the measurement of quality by different users. Or rather, if services were delivered simultaneously to a large number of users / customers under the same conditions (attendant/IT technician, workplace, time, etc.), it is possible that different customers establish different values to the performance of the service they received, since humans may have different perceptions concerning the same object of analysis.

In this context, it is suggested the realization of a broader case study involving a larger number of evaluators (users) in order to: (i) verify the reliability of the data collection instrument, (ii) reuse the two MCDA methods used in this study (iii) the use of multivariate statistical techniques, and finally, (iv) investigate the influence of the use of techniques or statistical procedures for elimination of outliers when assessing the quality of services.

References

Albrecht, K., Bradford, L J. (1990). *The Service Advantage*, Richard D. Irwin, Inc.

Arrow, K. J. (1963) Social Choice and Individual Values. London: JohnWiley and Sons.

Bharati, P., Berg, D. (2005) Service quality from the other side: Information systems management at Duquesne Light. International Journal of Information Management, 25. 367–380.

Belacel, N. (2000). Multicriteria assignment method PROAFTN: Methodology and medical applications", Journal of Global Optimization, n.º 125, pp. 175-183.

Bitner, M. J. (1990). "Evaluating service encounters: the effects of physical surroundings and employee responses", Journal of Marketing, n.54, 69-82.

Bolton, R.N., Drew, J.H. (1991a). "A longitudinal analysis of the impact of service changes on customer attitudes", Journal of Marketing, v. 55, p.1-9.

Bolton, R.N., Drew, J.H. (1991b). "A multistage model of customer's assessments of service quality and value", Journal of Consumer Research, v. 17, p.375-384.

Brady, M. K., Cronin, J. J., Brand, R. R. (2002). "Performance-only measurement of service quality: a replication and extension", Journal of Business Research, n.55, 17-31.

Carman, J. H. (1990). "Consumer perceptions of service quality: an assessment of SERVQUAL dimensions", Journal of Retailing, n. 66, 33-55.

Costa, H. G; Mansur, A. F. U; Freitas, A. L. P.; Carvalho, R. A. (2007) ELECTRE TRI applied to customers satisfaction evaluation. *Produção*, v. 17, n.2, 230-245. (*In Portuguese*).

Costa, H. G., Freitas, A. L. P. (2005) Using the ELECTRE TRI method for classifying the customer satisfaction: a case study in a university short-term course. *Revista Portuguesa e Brasileira de Gestão*. v.4, n.4, 66-76. (*In Portuguese*).

Cronin, J.J., Taylor, S.A. (1992). Measuring Service Quality: A Reexamination and Extension, Journal of Marketing, v. 56, p.55-68.

Dabholkar, P. A. (1995). A contingency framework for predicting causality between customer satisfaction and service quality, Advances in consumer research, v. 22, 101-108.

Dias, L.C. Mousseau, V. (2002). IRIS: Um SAD para problemas de classificação baseado em agregação multicritério, 3.ª Conf. da Associação Portuguesa de Sistemas de Informação, Coimbra, Portugal.

Fishburn, P. C. (1970). Utility theory for decision making. New York: Wiley.

Fischer, R.A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, n. 7, 179-188.

Freitas, A. L. P., Costa, H. G. (2012) Development and Testing of a Multi-Criteria Approach to the Assessment of Service Quality: An Empirical Study in Brazil. International Journal of Management, v.29, n.2, Part 2, p. 633-651.

Freitas, A. L. P. (2005) Service Quality in the context of competitivity. *Produção on Line*, v.5, n.1, 1-24. [*In Portuguese*].

Hair, J. F.; Black, W. C.; Babin, B. J.; Anderson, R. E.; Tatham, R. L. (2006) Multivariate Data Analysis, 6th Edition, Pearson Prentice Hall. New Jersey.

IT Governance Institute (2011) Global Status Report on the Governance of Enterprise IT (GEIT)—2011. Available at: http://www.isaca.org/Knowledge-Center/Research/Documents/Global-Status-Report-GEIT-10Jan2011-Research.pdf. [Accessed March 30, 2012]

Kang, H., Bradley, G. (2002). Measuring the performance of IT services: An assessment of SERVQUAL. International Journal of Accounting Information Systems 3, 151–164.

Keeney, R. L.; Raiffa, H. (1976) Decisions with Multiple Objectives: preferences and value tradeoffs New York: John Willey & Sons, 569 p.

Léger, J. Martel, J-M. (2002). "A multicriteria assignment procedure for a nominal sorting problematic", European Journal of Operational Research, n. 138, 349-364.

Marrone, M., Kolbe, L. M. (2011) Uncovering ITIL claims: IT executives' perception on benefits and Business-IT alignment. Inf Syst E-Bus Manage, 9:363–380.

Massaglia, M. Ostanello, A. (1991) N-Tomic: A support system for multicriteria segmentation problems: In: Horhonen, P. (Ed.), International Workshop on Multicriteria Decision Support, Helsinki, 1989. Lecture Notes in Economics and Mathematical Systems, v. 356, Berlin, 167-174.

Mesquida, A. L., Mas, A., Amengual, E., Calvo-Manzano, J. A. (2012) IT Service Management Process Improvement based on ISO/IEC 15504: A systematic review. Information and Software Technology, 54, 239–247.

Moscarola, J. (1977), "Aide à la decision en presence de critères multiples foundée sur une procédure trichotomique: méthodologie et application, Thèse de 3ème cycle Sciences des Organisations, Université de Paris-Dauphine.

Mousseau, V.;, Figueira, J. and Naux, J-Ph. (2001) Using assignment examples to infer weights for ELECTRE TRI method: Some experimental results. European Journal of Operational Research. n. 130. 263 – 275.

Mousseau, V.; Slowinski, R. (1998) Inferring an ELECTRE TRI Model from Assignment Examples. *Journal of Global Optimization*, n. 12, 157 – 174.

Mousseau, V.; Slowinski, R.; Zielniewicz, P. (2000) A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research*. v. 27. 757-777.

Peppard, J. (2003) Managing IT as a Portfolio of Services. *European Management Journal.v.* 21, N. 4, 467–483.

Perny, P. (1998). Multicriteria Filtering Methods Based on Concordance and Non –Discordance Principles. Annals of Operations Research, n.80, 137-165.

Roses, L. K., Hoppen, N., Henrique, J. L. (2009) Management of perceptions of information technology service quality. Journal of Business Research, 62, 876–882.

Roy, B. (1999) Decision-Aiding Today: What Should We Expect? In: Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications. Edited by Tomas Gal, Theodore J. Stewart, Thomas Hanne. Kluwer Academic Publishers, 1st edition, 560p.

Roy, B. (1985) Mèthodologie Multicritère d'Àide à la Dècision, Economica, Paris.

Sureshchandar, G.S., Rajendran, C., Anantharaman, R.N. (2002). The relationship between service quality and customer satisfaction - a factor specific approach. Journal of Services Marketing, Vol. 16. pp. 363 – 379.

Tsaur, S-H., Chang, T-Y., Yen, C-H. (2002). The evaluation of airline service quality by fuzzy MCDM. Tourism Managenement, n. 23, 107-115.

Vincke, P. (1989) *L'aide Multicritère à la Decision*, Editions de l'Université de Bruxelles - Editions Ellipses.

Watson, R.T., Pitt, L.F., Cuningham, C., Nel, D. (1993) User satisfaction and service quality of the IS department. J Inf Technol. 8:257–65.

Yu, W. (1992) ELECTRE TRI, aspects méthodologiques et guide d´utilisation. *Document du Lamsade,* nº 74,Université de Paris: Dauphine.

Zopounidis, C., Doumpos, M. (2002). "Multicriteria classification and sorting methods: a literature review". European Journal of Operational Research, n. 138, 229-246.