

Fuzziness, OWA and Linguistic Quantifiers for Web Selection Processes

Ronald R. Yager
Machine Intelligence Institute
Iona College, New Rochelle, USA
yager@panix.com
Visiting Distinguished Scientist
King Saud University
Riyadh, Saudi Arabia

Marek Z. Reformat
Electrical & Computer Eng.
University of Alberta
Edmonton, AB, Canada
Marek.Reformat@ualberta.ca

Giray Gumrah
Electrical & Computer Eng.
University of Alberta
Edmonton, AB, Canada
gumrah@ualberta.ca

Abstract—The Internet becomes an enormous source of information containing billions of documents. Users deal with an overwhelming number of alternatives, and continuously make decisions anytime they want to obtain meaningful information.

The paper describes an approach for a simple yet effective selection of the most suitable information that fits user's needs. The novelty of the approach is twofold: the concept of lexicographical-like preferences used for a multi-criteria decision-making with elements of fuzziness and OWA operator; a simple fuzzy number based mechanism for estimating user's degrees of acceptance of criterion satisfaction. The lexicographic preferences allow for mimicking user's attitude that some criteria should be satisfied before other criteria are considered. The acceptance of criterion satisfaction levels are defined with a single threshold that represents a boundary value between acceptable and unacceptable values of attributes of alternatives.

The paper includes results of a simple case study preformed on a prototype of a web selection system built using the proposed approach.

Keywords—component; multi-criteria decision-making, fuzziness, lexicographic preferences, degrees of acceptance, satisfaction levels, OWA operator

I. INTRODUCTION

The Internet is a multi-user, multi-source information repository with billions of documents and millions of users. In order to transform the Web into a user-friendly environment where everyone can easily and precisely find suitable things there is a need for tools and systems that support users in coping with constant inflow of information and selection of the most suitable things they are looking for, e.g., books, hotels, or movies. In many cases, users' search is limited to a few or several alternatives – more extensive search leads to more time spent on browsing the web.

Up to date, a variety of systems – called recommender systems – supporting users in their selection activities have been proposed. The most popular ones apply the concept of collaboration filtering. This selection mechanism provides suggestions to the user based on selections made by users that are similar in their selecting/purchasing patterns with the user under consideration. These types of systems work pretty well for simple items. However, in more complex situations – more sophisticated requirements and more complicated items –

these systems are not very effective. In order to address that issue, knowledge-based systems has been proposed. They use knowledge about a user. The suggestions they provide are derived using different decision making mechanisms based on users' direct input [1]. They build users' models and apply aggregation techniques, examples of some of them can be found in [2] and [3]. In many cases, these are complex and computationally intense approaches.

In this paper, we present a new approach of making suitable selection among all alternatives obtained from the web. The process takes into consideration multiple requirements that could be associated with different levels of importance. Additionally, the process of evaluating alternatives requires minimal input from the user representing user's estimation of usability of a given alternative.

The novelty of the proposed method is based on the application of a very simple yet effective method of ranking user's criteria. It uses the idea of a lexicographical (hierarchical) ordering. This ordering combined with a basic fuzzy operation is used for calculating values of weights required for an OWA-based aggregation of degrees of acceptance of criteria satisfaction levels. Our intention is to mimic the human method of making selections when some criteria dominate other criteria. In such case, levels of satisfaction of higher ranked criteria have influence on the ability of lower ranked criteria to contribute to the final decision. This influence is controlled by a t-norm operation.

As mentioned above, the presented approach for supporting selection processes relies on just a few pieces of information about the user. The user needs to identify her criteria of interest, provide ranking of those criteria, and give a basic information – in a form of single values – required for construction of fuzzy numbers representing user's levels of acceptance of criteria satisfaction levels.

This novel approach has been used to build a prototype system called *Personal Evaluation Tool (PET)*.

II. SELECTION AND DECISION-MAKING PROCESS

A. Systems Supporting Human Activities on the Web

It is quite common to use the web for finding information, buying different items, making travel arrangements, and looking for different sorts of entertainment. A process of

finding suitable items translates into a process of identifying the best-fitted alternative among a set of responses the web provides as the result of user's request. In a nutshell, the purpose of systems performing such duties – called recommender systems – [4] is quite simple: they suppose to find an alternative – a product, a service, or a piece of information – that fits user's interests. Such systems can be applied in a variety of activities, from online shopping to service selection.

There are multiple ways how the most suitable item or service response can be recognized. It can be done based on 1) finding similarities between users based on ratings of different alternatives/items provided by them using so called collaborative filtering systems [5], or 2) identifying similarities between alternatives and products that users have already bought, used, or liked – content-based recommender systems [6]. Some of the most popular similarity measures used for recommendation purposes are cosine similarity, naive Bayesian (NB) classifier, and Pearson correlation. Research activities in the area of recommender systems have led to many different variations of both types of those systems.

A need for more tuned and personalized recommendations have resulted in a new set of selection techniques focusing on matching user's needs and preferences to a set of available alternatives [1]. Systems that consider user's information – such as their preferences and satisfaction levels, belong to the category of knowledge-based recommender systems [7]. Those systems can learn a profile of user's interests implicitly based on attributes of alternatives rated by a user in the past [6], or explicitly gaining this information from a user [8].

A successful selection of the most suitable alternative depends on modeling of user's preferences. The preferences represent user's expectations and requirements set against the most desirable alternative, i.e., the alternative that satisfies user's demands to the highest degree. In the case of constructing a user's model (preferences and needs) a number of techniques have been investigated: a causality-driven user's model representing more domain-relevant aspects of the selection process [3], incomplete preference relations [9], question-based dialogue [10], dialogue-wise techniques based on a finite state model [11], and cross recommender mediation [12] to name a few.

B. Multi-criteria Decision Making in Web Applications

The problem of identifying the most suitable alternative in the presence of user's needs is also addressed by algorithms and techniques originated in the domain of multi-criteria decision-making. The topic of finding a best alternative has already drawn a lot of attention for many years [13]. In a nutshell, a selection process of multi-criteria decision-making consists of evaluating each alternative based on multiple measures, comparing obtained evaluations, and selecting an alternative with the best (usually the highest) evaluation value.

Elicitation of preferences is not a trivial process and may cause problems in obtaining a true reflection of user's needs. In some decision-making approaches, a user provides preferences inspecting directly all possible alternatives. Other techniques, such as verbal decision analysis and its variations,

have been proposed. Once user's preferences are known, a multi-criteria decision-making system has to induce a final selection. Many approaches identifying a final selection, such as the ones based on the concepts of Utility Theory, Analytic Hierarchy, and outranking approach, assume that a user provides preferences without any imprecision.

Another representation of user's needs is a set of criteria. Criteria represent principles by which a user would judge alternatives. A user specifies criteria as “constraints” imposed on the values of alternative attributes (an alternative is treated as a set of attributes. The values assigned to attributes constitute an “instance” of alternative). Based on comparison of a single criterion with an attribute value, a criterion evaluation value is generated. That value reflects how well a given criterion is satisfied by a given alternative. This process is performed for each criterion-attribute pair. A single alternative “generates” a number of criterion evaluation values (one for each criterion). These values have to be combined to obtain a final evaluation value of alternative. This step is called an aggregation of evaluation results. A number of possible aggregation operators can be applied here. Some of most popular ones are [2]:

- ordered weighted aggregation (OWA) operator [14];
- quasi-arithmetic means (different variations such as: quasi-linear means, weighted root-power mean, weighted harmonic mean, and weighted arithmetic mean);
- weighted median, and Sugeno interval.

Most of the above methods require weights. However, a process of identifying values of weights for aggregation calculations is not easy. Techniques dedicated to elicitation of weights involve asking a decision maker to answer simple questions about relative importance of criteria, analyzing these responses, and estimating ‘true’ weights. The weights used for OWA, and quasi-arithmetic means are calculated using criteria and regular non-decreasing quantifiers. For weighted median and Sugeno interval, weight values are obtained via a “try and error” approach process of satisfying user's criteria.

C. Lexicographical Rules

As mentioned above, once a given alternative is evaluated against each criterion, the results are aggregated and a single value representing goodness of this alternative is obtained. These processes are built based on diverse multi-criteria decision-making approaches suitable for aggregating multiple values [15]. The very captivating issue here is to perform such aggregation in a human-like way [16].

Mimicking a human decision-making process has been a subject of research activities for many years. One of the most popular techniques that fits a description of a human-like method and is applied for aggregation purposes are lexicographical rules [17, 18].

The lexicographical-based comparison of two alternatives is straightforward: it relays on a simple idea of comparing alternatives using one criterion at a time. The criteria are sorted based on their priorities, and a comparison sequence starts with a criterion of the highest priority. The choice between two alternatives A1 and A2 is made if the first criterion can identify a better alternative, if not, the second

criterion is used, and so on. If none criterion can do it, a random selection is applied. The method has been already applied in a number of different applications required simple, yet human-like decision making processes [19, 20].

The “popularity” of lexicographical decision-making rules has been also confirmed by studies in different application areas: for brand quality allocation, for estimating values of rural landscape attributes, or for management decision-making in petroleum refinery industry for distribution of oil to the various depots in a form of lexicographic goal programming (LGP) model.

D. OWA and Linguistic Quantifiers

Aggregation of different pieces of information is a common aspect of any system that has to infer a single outcome from multiple facts. A very interesting class of aggregation operators is called Ordered Weighted Averaging (OWA) [21]. In the simplest possible statement, this operator is a weighted sum over ordered pieces of information. In a formal representation, the operator, defined on the unit interval I and having dimension n , is a mapping $F_W : I^n \rightarrow I$ such that:

$$F_W(a_1, \dots, a_n) = \sum_{j=1}^n (w_j \cdot b_j) \quad (1)$$

where b_j is the j th largest of all arguments $a_1, a_2, \dots, a_n$, and w_j is a weight such that w_j is in $[0,1]$ and $\sum_{j=1}^n w_j = 1$.

In the mid-nineties, Yager showed that a linguistic quantifier could be used to obtain weights associated with an OWA aggregation. He has introduced parameterized families of the Regular Increasing Monotone (RIM) quantifiers [22]. These quantifiers are able to guide aggregation procedures by verbally expressed concepts in a description independent dimension. A RIM quantifier is a fuzzy subset Q over $I=[0,1]$ in which for any proportion $r \in I$, $Q(r)$ indicates the degree to which r satisfies the concept indicated by the quantifier Q [22]. A fuzzy subset Q represents a RIM quantifier if: $Q(0)=0$, $Q(1)=1$, and if $r1 > r2$ then $Q(r1) > Q(r2)$.

Assuming a RIM quantifier, we can associate with Q an OWA weights such that for $j=1$ to n ,

$$w_j = Q\left(\frac{j}{n}\right) - Q\left(\frac{j-1}{n}\right) \quad (2)$$

where n is a number of pieces of information to be aggregated. Using this expression the values of weights can be obtained directly from the expression representing the quantifier.

For example, let us take a look at the parameterized family $Q(r) = r^p$, where $p \in [0, \infty)$. Here if $p=0$, $w_1=1$ and $w_j=0$ for $j \neq 1$, we obtain the *existential (MAX)* quantifier, which makes the OWA operator F closer to an “or” operator; when $p \rightarrow \infty$, $w_n=1$ and $w_j=0$ for $j \neq n$, we have the quantifier *ALL (MIN)*, which makes the OWA operator F closer to an “and” operator; and when $p=1$ we have $Q(r)=r$, $w_j=1/n$ and we deal with the quantifier *SOME*.

The OWA aggregation has been used in multiple different applications including web-based information retrieval [23].

III. FUZZY-BASED SELECTION PROCESS: HUMAN-CENTRIC ASPECTS

There is a genuine need for tools and systems that are capable of supporting individual users in their unique way of exploring and utilizing information and services available on the Web. These systems should be able to mimic human selection processes in finding a single alternative that “fits” users’ needs when multiple alternatives are available.

The most important challenges encountered by those systems are the abilities to perform multi-criteria decision-making that replicates user’s selection behavior, and to elicit user’s criterion satisfaction and acceptance levels in an easy and simple way.

In real life, selections performed by people can be quite complex. They can involve decisive factors influencing evaluation of each alternative. Alternatives can be evaluated using multiple criteria, and different strategies can be used to identify the most suitable alternatives. A truly human-friendly system has to deal with multi-criteria evaluation of alternatives, and be able to handle criteria with differing importance levels. We will look at features of such a human-friendly system from the following aspects: importance of criteria, acceptance of criterion satisfaction levels, and aggregation procedure.

A. Criteria Ranking

It is clear that utilization of multiple criteria does not mean that all criteria are equally important. Some of them can be critical, while some of them can represent requirements that are not essential – kind of “good to have” requirements. As discussed in [14] user’s behavior can be summarized in the following way: if the higher priority criteria are not satisfied it is irrelevant if lower priority criteria are satisfied or not. That attitude can be also observed in many other situations related to Internet activities. Thus a user needs a certain degree of satisfaction to their higher priority criteria before allowing some compensation by lower priority criteria.

Such approach for evaluating alternatives can be implemented when a user provides a ranking of his/her criteria. This ranking defines an ordering in which evaluations are aggregated, and how evaluation values of different criteria contribute to the evaluation value of a given alternative [14].

B. Acceptance (Process) Levels

A ranking list of available alternatives should reflect the user’s degree of acceptance of these alternatives. An acceptance level for each alternative is estimated based on how well this alternative satisfies user’s criteria, i.e., what is the acceptance level associated with each criterion, and how these levels can be aggregated. In order to evaluate an alternative based on a criterion, the user has to identify how different values of alternative’s attribute related to this criterion satisfy his/her needs. In other words, the user has to provide some information allowing us to estimate up to what degree he/she is able to accept a particular value of an attribute. This information is applied to construct fuzzy numbers that are used for this estimation process.

An important objective of the proposed approach is to keep the process of elicitation of information required for determination of criterion acceptance levels as simple as possible. Since it is much easier for humans to provide values of directly observable attributes as opposed to introspective satisfaction levels of criteria, we have tried as much as possible to make use of values of attributes to determine the alternative's level of acceptance. The user provides only single threshold values for attributes associated with criteria. Once those thresholds are known, the system is able to generate acceptance levels directly from the values of attributes associated with given criteria.

In the approach proposed, the user is informed about minimum and maximum values of an attribute, and has to specify a threshold value within this range. It is explained to the user that the threshold is the attribute value at which he/she feels the full acceptance begins to deteriorate. More specifically, for a criterion whose acceptance level increases with bigger values of the attribute value its threshold can be described as “the smallest value of the attribute” which gives full acceptance (Fig. 1(a)). For those, such as cost, in which acceptance decreases with bigger values of the attribute value the threshold can be explained as “the biggest value of the attribute” which gives the full acceptance (Fig. 1(b)). In Fig. 1 we illustrate constructed fuzzy numbers used to transform attribute values into criterion acceptance levels. This figure represents two plots of criterion acceptance degrees – we will call them acceptance fuzzy numbers. As it can be seen, the acceptance levels below the threshold (Fig. 1(a)) are equal to *max_level* (usually 1.0), while the values above the threshold are monotonically decreasing. The decrease represents user's changes in criterion acceptance degree when the values of attribute become larger than the threshold. The maximum attribute value is associated with a small, non-zero value – *min_level*. Equivalently the same concept is applied when the desired attribute values should be larger – closer to the maximum (Fig. 1(b)). More sophisticated fuzzy numbers can be used if we are willing elicit more introspective information from a user.

C. Lexicographical-based Aggregation Method: Concept

The process of evaluating or scoring a single alternative based on a set of criteria depends on combining the following pieces of information:

- degrees of acceptance of individual criteria;
- weights reflecting priorities assigned to criteria.

One of the simplest ways of aggregating multiple pieces of information in order to obtain a single score representing the acceptance level of an alternative is a weighted sum. The formula is:

$$Score = \sum_{i=1}^{no_of_attributes} criterion_weight_i * acceptance(attribute_i) \quad (3)$$

Each component of the sum is a product of an acceptance level that is the result of application of an acceptance fuzzy number to an *i*-th *attribute_i* and a weight associated with the equivalent criterion *i* – *criterion_weight_i*. The weights have to be

determined for each criterion. They should denote user's preferences and his/her objectives regarding contribution of acceptance levels of individual criteria to the overall score.

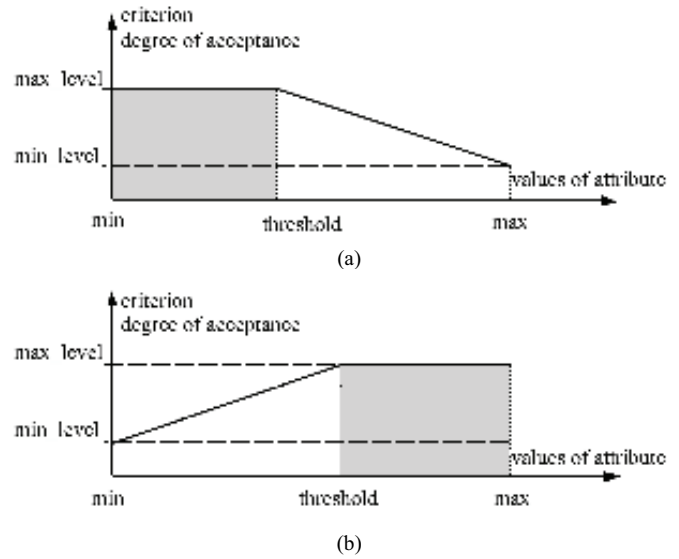


Fig. 1. Graphical representation of fuzzy numbers representing degree of acceptance for two scenarios: decreasing acceptance with increasing attribute value (a), and increasing acceptance with increasing attribute value (b).

The weights can be assigned in a universal way so they can be used in determining the scores for different alternatives. However, this universal assignment of weights independent of consideration of the alternative being evaluated is not always right. It is possible that if the highest criterion is not fully satisfied, high acceptance levels of low priority criteria can bring the *score* to the level that is higher than the *score* of another alternative satisfying the highest priority criterion but not satisfying low priority criteria. In such a case, a selection mechanism would indicate an alternative with low satisfaction of highest priority criteria as a winner.

Here we shall use a procedure, proposed in [14], for determining the criterion weights that avoids this problem. The weights are calculated via a unique combination of the criteria ranking and criterion acceptance levels. This leads to a schema that is very similar to lexicographic preferences. The criteria weights for a given alternative are calculated based on to the following formula where the term *acceptance(attribute_k)* represents the level of acceptance of the *k*-th criteria for the alternative being scored

$$criterion_weight_i = \prod_{k=1}^{i-1} acceptance(attribute_k) \quad (4)$$

for *i*=2, ..., *n*, where *n* is a number of criteria, while **T** is a t-norm. It can be observed that the weight of *i* criterion depends on the acceptance levels obtained for all *i-1* criteria that have a higher priority than *i* attribute (The indexing of criteria is equivalent to a hierarchy of criteria). For *i*=1 (the first criterion of the highest priority), its *criterion_weight* is equal to 1.0. Such approach prevents cases where lower priority

criteria can contribute to the overall evaluation value so much that this alternative is more “attractive” than an alternative with higher acceptance of high priority criteria but lower acceptance of low priority criteria.

The degree of contribution of the higher priority criteria on the weight of lower priority criteria is “controlled” by the type of applied t-norm (4). For the product t-norm, the weight of the attribute i is a product of acceptance levels of all criteria that have higher priority levels. It means that each criterion of lower priority has lower or the same weight comparing with the weights of criteria of higher priority. For the min t-norm, the weight value is equal to the smallest acceptance level among all $i-1$ criteria. It means that the weights of all criteria with priorities lower than the criterion with the lowest acceptance have the same weight and potentially the same contribution to the overall score.

Please note that the criteria weights in this approach will be different for each alternative as they have different acceptance levels for the criteria.

IV. LEXICOGRAPHIC-BASED AGGREGATION IN MULTI-CRITERIA DECISION-MAKING

The proposed method for determining an alternative's aggregated score can be used in situation in which there are ties between criteria with respect to their level of importance. This induces a grouping of criteria. In this case, there are several groups of criteria, and each group has different priority, while all criteria from the same group have the same priority.

Let us assume that there are g different groups (levels) of criteria. The criteria $C_{11}, C_{12}, \dots, C_{1b}, \dots, C_{1n_1}$ belong to the group of the most important criteria, $C_{21}, C_{22}, \dots, C_{2b}, \dots, C_{2n_2}$ belong to the second most important group, and $C_{g1}, C_{g2}, \dots, C_{gb}, \dots, C_{gn_g}$ to the least important group g . In total, there are $n = n_1 + n_2 + \dots + n_g$ criteria. In this case, the attributes of an alternative x_k are indexed in the following way: $x_{k,(11)}, x_{k,(12)}, \dots, x_{k,(1i)}, \dots, x_{m,(1n_1)}, \dots, x_{k,(21)}, x_{k,(22)}, \dots, x_{k,(2i)}, \dots, x_{m,(2n_2)}, \dots, x_{k,(g1)}, x_{k,(g2)}, \dots, x_{k,(gi)}, \dots, x_{m,(gn_g)}$ (the indexes in the brackets represents indexes of attributes, and they are equal to the indexes of equivalent criteria).

Let $C_{ij}(x_{k,(ij)})$ represents a value of criteria C_{ij} for an attribute j in the group priority i of an alternative x_k . In the case of priority groups the aggregation process obeys the following formula:

$$Score(x_k) = \sum_{i=1}^g \sum_{j=1}^{n_i} w_{ij} C_{ij}(x_{k,(ij)}) \quad (5)$$

The values of weights w_{ij} are obtained using the following procedure. Let S_i represents an acceptance value obtained for an alternative x_k based on the criteria from a priority group i . S_i can be calculated using the OWA operator. In such case:

$$S_i(x_k) = OWA\{C_{i1}(x_{k,(i1)}), C_{i2}(x_{k,(i2)}), \dots, C_{in_i}(x_{k,(in_i)})\} \quad (6)$$

Different linguistic quantifiers can be combined with OWA (Section II.D). The quantifiers can represent different ways of

user's needs imposed on the group criteria. One of them is *ALL* – all criteria in the group should be satisfied, and *SOME* – some criteria have to be satisfied in order to treat the criteria group as satisfied. Those quantifiers represent different “personalities” of users – *ALL* represents users more demanding and strict in their requirements, and *SOME* represents users who are willing to “forget” that all criteria in a group should be satisfied.

The values of S_i are calculated for each group and are directly used in calculations of pre-weights (7). The pre-weights for all criteria that belong to the same group are the same. In this case we talk about group pre-weights instead of individual criteria pre-weights.

$$\begin{aligned} u_{11} &= u_{12} = u_{13} = \dots = u_{1n_1} = 1 \\ u_{21} &= u_{22} = u_{23} = \dots = u_{2n_2} = S_1(x_k) \\ u_{31} &= u_{32} = u_{33} = \dots = u_{3n_3} = \prod_{i=1}^2 S_i(x_k) \\ &\dots \\ u_{g1} &= u_{g2} = u_{g3} = \dots = u_{gn_g} = \prod_{p=1}^g S_p(x_k) \end{aligned} \quad (7)$$

To ensure that an alternative with a full acceptance of all criteria obtains an overall score of 1.0, the values of the weights are obtained by dividing the pre-weights by the total number of criteria. Here, the total number of criteria is:

$$T = \sum_{i=1}^g n_i$$

and then

$$w_{ij} = \frac{u_{ij}}{T}. \quad (8)$$

Assuming that

$$\begin{aligned} w_{11} &= w_{12} = w_{13} = \dots = w_{1n_1} = w_1 \\ w_{21} &= w_{22} = w_{23} = \dots = w_{2n_2} = w_2 \\ w_{31} &= w_{32} = w_{33} = \dots = w_{3n_3} = w_3 \\ &\dots \\ w_{g1} &= w_{g2} = w_{g3} = \dots = w_{gn_g} = w_g \end{aligned}$$

a new form of the equation (5) is:

$$Score(x_k) = \sum_{i=1}^g w_i \sum_{j=1}^{n_i} C_{ij}(x_{k,(ij)}) \quad (9)$$

V. A SIMPLE WEB-BASED PET ARCHITECTURE

The proposed lexicographical-based approach supporting user's selections process is a pivotal element of a *Personal Evaluation Tool (PET)*. The *PET* is a simple web-based system that allows its users to input information about priorities of criteria and values of parameters needed for defining satisfaction level functions. As the result, the *PET* provides users with ranked – based on data entered by users – alternatives. The architecture of the *PET* is presented in Fig. 2. The *PET* consists of three execution components:

Web Interaction Unit: It is a simple unit that initiates a service processing activity based on a query provided by the user. Based on this query – received from the Browser-based User Interaction Unit – the unit performs a simple keyword-based web search. The result of that search – a set of responses matching the query keywords – is provided to the Selection Unit. Each response represents an alternative that is compared against users preferences and levels of acceptance.

Selection Unit: It is a “brain” of the *PET*. This is the place where each alternative is evaluated using acceptance fuzzy numbers. This process is a full implementation of the approach described in Section IV.

Browser-based User Interaction Unit: In order to utilize the presented priority-based approach, a web-based user interface has been developed. It allows a user to rank criteria, and enter her information about different criteria satisfaction levels.

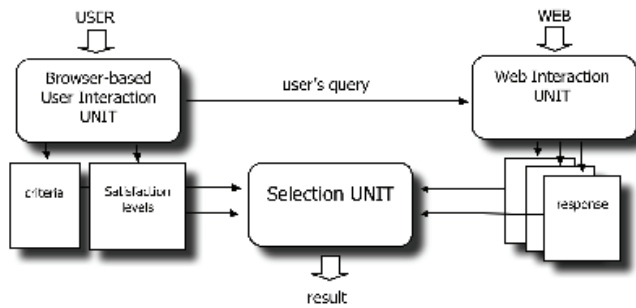


Fig. 2. Architecture of the Web-based *Personal Evaluation Tool (PET)*

The web-based interface has been developed for Hotel Reservation. The following criteria are identified:

- *room price*;
- *location* (here, we assume the distance from a hotel to the downtown);
- *access* (central reservation or not);
- *security* (personal safe, peephole, door chain);
- *services* (airport shuttle, room service, high speed internet, fax, photocopying, laundry/dry cleaning, lounge, bar);
- *tangibles* (family room, minibar, Jacuzzi, parking);
- *facilities* (printer, swimming pool).

The user has to provide information regarding ranking and satisfaction levels for all criteria. Once, the user selects the Hotel Reservation service, he/she ranks all criteria. We allow the user to identify more than one criterion at the same level of importance. Only relevance of positions is taken into consideration.

For acceptance levels, a user navigates to the page that allows for entering pieces of necessary information. For the criteria: *room price* and *location* – the user is provided with sliders that can be moved to positions representing the border (threshold) between a full acceptance and decreasing acceptance level. For example, for the *room price* – the user is satisfied with the price up to \$140 (the level of acceptance for any price below the threshold of \$140 is 1.0), any *room price* above that value means a decreased acceptance level. For location acceptance, a distance up to 2 km is okay for the user,

any larger distance means lower level of acceptance – the value of acceptance reaches 0.1 for the distance of 5 km.

There are also few “radio button” entries that allow the user to select choices related to *access*, *security*, *services*, *tangibles* and *facilities*. Several sub-criteria are defined for each of them. The options provided to the user are “should have” and “don’t care”. In this case the values of acceptance are binary – it equals *one* if an alternative hotel has that feature, or *zero* if it does not.

VI. CASE STUDY

A. Overview

The proposed lexicographical-based approach for effective selection of the most suitable alternative in the context of a web search process is validated, as well as compared with another method. The validation process is performed via comparison of the results obtained from the prototype *PET* with a realistic human selection process. The comparison of *PET*’s results with another method is focused on finding similarities and differences between *PET*’s selections and selections provided by a simple lexicographic-based approach “*Take-The-Best*” (*TTB*).

Take-the-Best

Our motivation for selecting another lexicographical-based method, called *Take-the-Best*, is based on a simple fact that it works surprisingly well when compared to statistical benchmarks [24]. The method uses a very strict and simple lexicographical rule. Once alternatives are evaluated against each criterion they are being pair-wise compared. The choice between two alternatives A and B is made if the first criterion can identify a better alternative, if not, the second criterion is used, and so on. If none criterion can do it, a random selection is applied. This means that once we have knowledge about priorities of criteria we sort them and perform comparison using sorted criteria on one at a time basis.

Normalized Distance-based Performance Measure (NDPM)

NDPM was proposed by Yao [24]. He developed *NDPM* theoretically, using an approach from decision and measurement theory. *NDPM* can be used to compare two different weakly ordered rankings.

$$NDPM = \frac{2P^c + P^o}{2P^t} \quad (10)$$

P^c is the number of contradictory preference relations between the system ranking and the user ranking. A contradictory preference relation happens when the system says that item A will be preferred to item B, and the user ranking says the opposite. P^o is the number of compatible preference relations, where the user rates item A higher than item B, but the system ranking has item A and item B at equal preference levels. P^t is the total number of “preferred” relationships in the user’s ranking (i.e. pairs of items rated by the user for which one is rated higher than the other). This metric is comparable among different datasets (it is normalized), because the numerator represents the distance, and the denominator represents the worst possible distance.

Let us assume that we have a user's ranking of items: *Ia, Ib, Ic, Id, If*. For five items the value of P' is ten – there are ten different pairs of items. The value of *NDPM* is 0.0 when the ranking is compared with another identical ranking. If both rankings differ in one pair the value of *NDPM* is equal to 0.1, two different pairs lead to 0.2. If there are ten differences – it means that the other ranking is totally different: *If, Id, Ic, Ib, Ia* – the value of *NDPM* is 1.0.

B. Hotel Reservation Service

The case study illustrates a Hotel Reservation service. The experiment involves eleven users. The selection process for the most suitable hotel takes into accounts six different criteria: *room price, location* (we assumed that the point of interest for each user is a downtown, therefore the location represents a distance from a given hotel to the downtown), *access, security services, tangibles, and facilities*. Such criteria as *security, additional services, tangibles, and leisure facilities*, contain multiple sub-criteria.

The case study is performed with five different hotels. The detailed information about each hotel is presented in Table 1. Each of users has identified his/her priorities, and provided information needed for construction of acceptance fuzzy numbers. Following that, each user has ranked all five hotels based on all features (Table 1).

TABLE 1. HOTEL RESERVATION SERVICE ALTERNATIVES

		H1	H2	H3	H4	H5
Price (single room)		\$249	\$199	\$129	\$99	\$159
Location (distance from downtown)		0.5km	2.3km	1.9km	3.5km	1.0km
Access	Central reservation	x	x	-	-	-
Security	Personal safe in room	x	-	-	-	x
	Peephole, door chain	-	-	x	x	x
Additional Services	Airport shuttle	x	x	-	-	-
	Room service	x	x	x	-	x
	High speed internet	x	x	x	-	-
	Fax, photocopying	x	x	-	-	x
	Laundry/dry cleaning	x	-	-	-	x
	Lounge/bar	x	x	x	x	x
Tangibles	Family room	-	-	-	-	x
	Minibar	x	x	x	-	x
	Jacuzzi	x	x	-	-	-
	Parking	-	x	-	x	-
Leisure Facilities	Children's leisure	-	-	-	-	x
	Swimming pool	x	x	-	-	-
	Golf course	x	-	-	-	-

The selection process based on the proposed lexicographical-based approach has a number of variations. Such elements as a t-norm or OWA linguistic quantifier should be determined. In the case of a t-norm our choice is the min t-norm. The reason for that is the close resemblance of the proposed approach with the min t-norm to the lexicographical

one. Additional experiments about suitability of different types of t-norm have been left for later. For the selection of a linguistic quantifier required for determining the value of S_i (6) we investigate two quantifiers *ALL* and *SOME*. The results of this work are shown below.

PET vs. Users (NDPM)

Firstly, the evaluation of *PET* results is performed using *NDPM*. As it has been indicated earlier, the ranking of five items leads to ten different preferences pairs. In this case the value of *NDPM* of 0.1 means that the rankings differ in one preference pair. The *NDPM* values calculated for a pair of rankings – one ranking given by the user, and the other offered by *PET* based on priorities and criteria entered by this user – for all eleven users and two linguistic quantifiers *ALL* and *SOME* are shown in Table 2. On average the value of *NDPM* is 0.3 meaning that on average the *PET*'s rankings have three pairs of preferences different from the rankings provided by the users. The smallest difference between the *PET* and the user rankings is observed for the **user_#5** for *ALL* and the **user_#1** and **user_#3** for *SOME*. In the case of **user_#5** the contradictory pair has been observed for positions 3 and 4 – the user's rank of hotels is: **H5, H3, H4, H1, H2** (from the most to the least suitable), while the *PET*'s ranking is: **H5, H3, H1, H4, H2**. The biggest difference is observed for the **user_#7** (for both *ALL* and *SOME*) – in this case, there are six contradictory preferences. The user's ranking is: **H5, H3, H2, H1, H4**, while *PET*'s is: **H1, H2, H5, H3, H4**. A short discussion with the user has led to the conclusion that the user is not consistent – the preferences and criteria (s)he has entered to the *PET* system are different then those (s)he has used to make the selection. As we can observe, there is a small difference – on average – between results for quantifiers *ALL* and *SOME*.

TABLE 2. NDPM VALUES FOR RECOMMENDATIONS OBTAINED FROM PET WHEN COMPARED WITH RANKINGS PROVIDED BY THE USERS

	value of NDPM	
	OWA: ALL	OWA: SOME
user #1	0.25	0.10
user #2	0.25	0.30
user #3	0.15	0.10
user #4	0.20	0.20
user #5	0.10	0.20
user #6	0.30	0.30
user #7	0.60	0.60
user #8	0.30	0.30
user #9	0.55	0.50
user #10	0.30	0.50
user #11	0.25	0.20
average	0.30	0.30

PET vs. TTB (1st recommendation)

The *TTB* method (the simplest possible lexicographical-based selection technique) can only recommend a single hotel. Therefore, the analysis of results is done by comparison of the *PET*'s 1st recommendation with the suggestion provided by *TTB*. The linguistic quantifier used here is *ALL* – this quantifier “is closer” to the original lexicographical approach. The results for *PET* and *TTB* are presented in Table 3. We can see that *TTB* is correct in four cases only. Looking at the first

position of the ranking provided by the *PET*, the *PET* is correct six times.

The *PET* prototype provides ranking of all hotel alternatives. Therefore, we also compare the top-2 and top-3 “suggestions” given by *PET* with the user’s first choice. For the top-2, the *PET* finds the user’s top choice for ten users (Table 3). For the top-3 recommended hotels, the *PET* system is able to identify the user’s first choice for all eleven individuals.

TABLE 3. RESULTS FOR HOTEL RESERVATION SERVICE OBTAINED FROM *PET* AND *TTB* WHEN COMPARED TO THE USER’S FIRST CHOICE

	# of correct	# of incorrect
TTB – first recommendation	4	7
PET – first recommendation	6	5
PET – first two recommendations	10	1
PET – first three recommendations	11	0

VII. CONCLUSIONS

Increased popularity of the Internet triggered interest in development of methods and techniques supporting users’ search for relevant information and services. One of the most challenging requirements is the ability to mimic human way of making selections and recommendations.

The paper proposes a novel technique for selecting the most suitable alternative based on multiple criteria. The technique combines information about ranking of criteria with criterion levels of satisfaction and acceptance. A new algorithm that calculates the evaluation value of each alternative has been proposed and described. The algorithm uses the principle of weighted aggregation – it dynamically calculates weights following the idea of lexicographic preferences combined with fuzzy operations – t-norms – and the OWA operator. The weights are recalculated for each alternative, and their values depend on location of a given criterion in criteria hierarchy and how well the satisfaction of criteria are acceptable for the user. The criteria acceptance levels are estimated using a newly proposed and simple process of constructing fuzzy numbers based on a single parameter representing a boundary value between acceptable and unacceptable values of criterion.

A simple system – called *PET* – has been developed based on the proposed technique. The case study has been preformed to assess usability of the proposed method. It indicated a good correlation of the rankings deduced by the *PET* prototype with the rankings provided by the users. An important conclusion of the performed comparison is a need for tracking changes in users’ criteria – mismatches are the consequence of users’ inconsistency in their preferences. More experiments are planed for investigations focused on different types of t-norms and linguistic quantifiers.

REFERENCES

- [1] R. Burke, “Hybrid Recommender Systems: Survey and Experiments”, User Modeling and User-Adapted Interact, vol. 12, pp. 331-370, 2002.
- [2] R. Smolikova, and M.P. Wachowiak, M.P., “Aggregation operators for selection problems”, Fuzzy Sets and Systems, vol. 131, pp. 23-34, 2002.
- [3] B. Towle, and C. Quinn, “Knowledge Based Recommender Systems Using Explicit User Models”, Knowledge-Based Electronic Markets, the AAAI Workshop, Menlo Park, CA: AAAI Press, pp. 74-77, 2000.
- [4] P. Resnick, N. Iacovou, N. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: an open architecture for collaborative filtering of netnews”, Computer Supported Collaborative Work (CSCW), Chapel Hill, NC, USA, pp. 175-186, 1994.
- [5] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J. T. Riedl, “Evaluating Collaborative Filtering Recommender Systems”, ACM Transactions on Information Systems, vol. 22(1), pp. 5-53, 2004.
- [6] M.J. Pazzani, “A Framework for Collaborative, Content-Based and Demographic Filtering”, Artificial Intelligence Review, vol. 13(5/6), pp. 393-408, 1999.
- [7] R. Burke, “Knowledge-based Recommender Systems”, in A. Kent (ed.), Encyclopedia of Library and Information Systems, vol. 69, Supp. 32, New York: Marcel Dekker, 2000.
- [8] M. Zanker, and M. Jessenitschnig, “Case-studies on exploiting explicit customer requirements in recommender systems”, User Modeling and User-Adapted Interaction, vol. 19, pp. 133-166, 2009.
- [9] L. Martinez, L.G. Perez, M.J. Barranco, and M. Espinilla, “A Knowledge Based Recommender System Based on Consistent Preference Relations”, Studies in Computational Intelligence (SCI) 117, pp. 93-111, 2008.
- [10] C. Thompson, M. Goker, and P. Langley, “A Personalized System for Conversational Recommendations”, Journal of Artificial Intelligence Research, vol. 21, pp. 393-428, 2004.
- [11] A. Felfernig, G. Friedrich, D. Jannach, and M. Zanker, “An Integrated Environment for the Development of Knowledge-Based Recommender Applications”, Int. Journal of Electronic Commerce, vol. 11(2), pp. 11-34, 2006.
- [12] F. Herrera, E. Herrera-Viedma, and F. Chiclana, “Multiperson decision-making based on multiplicative preference relations”, European Journal of Operational Research, vol. 129, pp. 372-385, 2001.
- [13] E. Triantaphyllou, Multi-Criteria Decision Making Methods: A comparative Study, Springer, 2000.
- [14] R.R. Yager, “Prioritized aggregation operators”, International Journal of Approximate Reasoning, vol. 48(1), pp. 263-274, 2008.
- [15] G. Adomavicius, and Y.O. Kwon, “New Recommendation Techniques for Multicriteria Rating Systems”, IEEE Intelligent Systems, vol. 22(3), pp. 48-55, 2007.
- [16] G. Gigerenzer, and D.G. Goldstein, “Reasoning the Fast and Frugal Way: Models of bounded Rationality”, Psychological Review, vol. 103(4), pp. 650-669, 1996.
- [17] K. W. Church, “Approximate Lexicography and Web Search”, International Journal of Lexicography, vol. 21(3), pp 325-336, 2008.
- [18] R. H. Hogarth, and N. Karelaia, “Take-the-Best” and Other Simple Strategies: Why and When They Work “Well” with Binary Cues, Theory and Decision, vol. 61, pp. 205-249, 2006.
- [19] M. Ehrgott, A characterization of lexicographic max-ordering solutions, in A. Göpfert, J. Seeländer, and C. Tammer, editors, Methods of Multicriteria Decision Theory, Proceedings of the 6th Workshop of the DGOR-Working Group Multicriteria Optimization and Decision Theory, pp. 193-202, 1997.
- [20] K.-W. Jee, D. L. McShan, and B. A. Fraass, “Lexicographic ordering: intuitive multicriteria optimization for IMRT”, Physics in Medicine and Biology, vol. 52, pp. 1845-1861, 2007.
- [21] R.R. Yager, “On ordered weighted averaging aggregation operators in multi-criteria decision making”, IEEE Transactions on Systems, Man and Cybernetics, vol. 18, pp. 83-190, 1988.
- [22] R.R. Yager, “Families of OWA operators”, Fuzzy Sets and Systems, vol. 59, pp. 125-148, 1993.
- [23] H. Amiri, A. AleAhmad, F. Oroumchian, C. Lucas, M. Rahgozar, Using OWA Fuzzy Operator to Merge Retrieval System Results, 2nd Workshop on Computational Approaches to Arabic Script-based Languages, pp. 13-21, 2007.
- [24] Y.Y. Yao, “Measuring Retrieval Effectiveness Based on User Preference of Documents”, Journal of the American Society for Information Science, vol. 46, pp. 133-145, 1995.