A System based on Multiple Criteria Analysis for Scientific Paper Recommendation

Nikolaos F. Matsatsinis, Kleanthi Lakiotaki and Pavlos Delias

Technical University of Crete
Decision Support Systems Laboratory
University Campus, 73100, Chania, Greece
{nikos, klio, pavlos}@ergasya.tuc.gr

Abstract

Systems able to suggest items that a user may be interested in are usually named as Recommender Systems. The new emergent field of Recommender Systems has undoubtedly gained much interest in the research community. Although Recommender Systems work well in suggesting books, movies and items of general interest, many users express today a feeling that the existing systems don't actually identify them as individual personalities. This dissatisfaction turned the research society towards the development of new approaches on Recommender Systems, more user-centric. A methodology originated from Decision Theory is exploited herein, aiming to address to the lack of personalization in Recommender Systems by integrating the user in the recommendation process.

Keywords: Recommender Systems, multiple criteria analysis, UTASTAR algorithm

1. Introduction

Traditionally, people used to recommend products or services to one another by the process of "word-of-mouth". Nowadays the amount of information increases enormously day by day and is provided to people predominantly via the World Wide Web. The development of automated Recommender Systems (RS) was thus a natural expectancy. To find information of quality from multiple heterogeneous sources is increasingly difficult. This problem finds its origin in the fast growth of the number of information available. The goal of a Recommender System is to reduce information overload by selecting a subset of items from a universal set based on user preferences. Specifically, the recommendation problem can be formulated as follows [Adomavicius, G. et al. 2005]: Let C be the set of all users and let S be the set of all possible items that can be recommended. Both spaces can be very large extending to millions or more of users or items. Let u be a utility function that measures usefulness of item s to user c, i.e., $u : C \times S \rightarrow R$, where R is a totally ordered set. Then, for each user $c \in C$, we want to choose such item $s \in S$ that maximizes the user's utility. More formally:

$$\forall c \in C, s_c^{'} = \arg\max u(c,s), \quad s \in S \tag{1}$$

A major issue in Recommender Systems is personalization, which considers the process of collecting user-information during interaction with the user. This information is subsequently used to deliver appropriate content and services, according to user's needs. User satisfaction is the most important goal of personalization. It is motivated by the condition that a user has needs, and addressing them successfully is likely to lead to a satisfactory relationship and re-use of the services offered.

As stated in [McNee, S. M. et al. 2006], Recommender Systems need a deeper understanding of users and their information seeking tasks. They propose the Human-Recommender Interaction (HRI) framework to incorporate knowledge for the user into the Recommender System. The article in the Wall Street Journal [Zaslow, J. 2002] underlies the pretence for stirring the Recommender Systems researchers towards a more user oriented perspective.

Recommender Systems are usually classified in three main categories depending on how the recommendation process takes place. These are: 1) Content based methods 2) Collaborative filtering methods and 3) Hybrid methods. Although traditional techniques in RS work well so far in predicting users' preferences in goods, they clearly need some improvement to be able to provide better and more benefits to user. New approaches need to be proposed to overcome the existing systems limitations. An overview of these limitations can be found in [Adomavicius, G. et al. 2005] while they are also mentioned in [Shardanand, U. et al. 1995] and in [Balabanovic, M. et al. 1997].

For the content based recommendation techniques limitations are mostly related to the content dependency. A major shortcoming for RS is the cold-start problem. This flows from the fact that since new items are added regularly to Recommender Systems and collaborative systems rely solely on users' preferences to make recommendations, we may easily infer that until the new item is rated by a substantial number of users, the Recommender system would not be able to recommend it.

To overcome the above mentioned limitation, many hybrid techniques which combine in various ways both methods of collaborative filtering and content based filtering have been proposed by [Burke, R. 2002].

A totally different approach to address this issue is proposed herein. We advocate the use of Multiple-criteria Decision Aiding (MCDA) methodologies to deal with the user-recommender interaction theme. Although MCDA methodologies have been extensively studied in Operations Research community [Figueira, J. et al. 2005], not much has been done yet in the field of RS [Adomavicius, G. et al. 2005].

Decision aiding can be defined as follows [Roy, B. 1996]: Decision aiding is the activity of the person who, through the use of explicit but not necessarily completely formalized models, helps obtain elements of responses to the questions posed by a stakeholder in a decision process. These elements work towards clarifying the decision and usually towards recommending, or simply favouring, a behaviour that will increase the consistency between the evolution of the process and this stakeholder's objectives and value system. In this definition, the word "recommending" is used to draw attention to the fact that both analyst and decision maker (DM) are aware that the DM is completely free to behave as he or she sees fit after the recommendation is made.

The methodology discussed in this paper belongs to the broader family of MCDA techniques and since it is a clearly interactive user-system method, it addresses the cold start problem of RS's, simply by the fact that the user introduces his/her preferences to the system by ranking a familiar set of items at the beginning of the process. This ranking is then used to asses users' value system and provide future recommendations according to user's preferences.

As a demonstration of the proposed methodology we deal with the problem of recommending academic papers to the research community. The problem of recommending scientific articles is multicriteria by nature and an efficient autonomous system for paper recommendation would inevitably become an integral tool for a researcher.

This paper is organized as follows: First, in section 1 Recommender Systems are briefly introduced, formulated and major shortfalls of these systems are mentioned. A short literature review on Recommender Systems is provided in section 2 and existing tools for research work retrieval are presented in section 3. In section 4, the proposed methodology is analytically discussed and an illustrative example is implemented in section 5. Finally, conclusions and future aspects are discussed in section 6.

2. Literature review

While research in Recommender Systems grew out of information retrieval and filtering, the topic has steadily advanced into a challenging research area of its own in the mid-1990. Since then, various approaches have emerged mainly from the field of information retrieval, artificial intelligence and management science. There has been much research work conducted both in industry and academia on developing new approaches to Recommender Systems.

The term "Recommender System" was coined by Resnick and Varian [Resnick, P. et al. 1994; Resnick, P. et al. 1997] as a generic replacement for "collaborative filtering", a phrase proposed earlier by [Goldberg, D. et al. 1992] for the first Recommender System, named Tapestry, a filter-based electronic mail system designed for the intranet at the Xerox Parc Palo Alto Research Centre.

In general, Recommender Systems manage information overload by helping users choose among an overwhelming number of alternatives.

As stated in [Adomavicius, G. et al. 2005] these systems broadly fall into three main categories based on the techniques they use to narrow the range of possible choices:

Content-based filtering systems utilize the users' past likings to recommend new items. Content-based Recommender Systems are most popular with systems that work on textual content, for example web pages or news pages [Pazzani, M. et al. 1996]. As stated by [Mooney, R. J. et al. 2000] this approach has the advantage of being able to recommend previously unrated items to users with unique interests and to provide explanations for its recommendations.

Collaborative filtering (CF) systems try to predict the usefulness of items for a particular user based on the items previously rated by other users. It is the method of making automatic predictions (filtering) about the interests of a user by collecting taste information from many users (collaborating). The underlying assumption of CF approach is that those who agreed in the past tend to agree again in the future. Collaborative filtering (CF) technique is the most widely used in RS's [Herlocker, J. L. et al. 2004] and a large number of applications has emerged from that field of research. For example, the Ringo system [Shardanand, U. et al. 1995] applied collaborative filtering to recommend music to individuals. GroupLens [Resnick, P. et al. 1994] and PHOAKS [Terveen, L. et al. 1997] are two examples of collaborative based systems to recommend messages from USENET news.

There are three main categories of CF algorithms: 1) memory-based algorithms, 2) model-based algorithms and 3) hybrid algorithms [Breese, J. et al. 1998].

Hybrid systems combine both content based and collaborative filtering methods in order to overcome certain limitations of each method [Burke, R. 2002; Melville, P. et al. 2002].

A taxonomy of Recommender Systems on the internet can be found in the review [Montaner, M. et al. 2003].

3. Searching for scientific literature

The most tedious and time consuming part of a researcher's responsibilities is searching for related work. It is possible during this procedure to miss important developments in a specific area. The procedure of finding related work usually starts by locating a small number of initial papers and navigating the citation web near those papers. The number of scientific articles catalogued in the internationally recognized peer-reviewed set of Scientific & Engineering journals covered by the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) grew from approximately 466,000 in 1988 to nearly 700,000 in 2003, an increase of 50% [Board, N. S. 2006]. It is impossible for an individual scientist to discover all the available related work even in the bounds of a specific field. Moreover, with the trend of science towards the grouping of interdisciplinary fields, Recommender System adapted to researchers' profile and preferences become even more valued.

Many tools have been developed to enhance the effectiveness of research paper retrieval. The major sources that provide citation data is the Web of Science (http://portal.isiknowledge.com/), Google scholar (http://scholar.google.com/intl/en/scholar/about.html) and Citeseer [Giles, C. L. et al. 1998]. CiteSeer for example is a free access scientific literature digital library and search engine that focuses primarily on the literature in computer and information science. As of November 10, 2006 it contains 756,208 documents for which it provides citation analysis. Google Scholar, launched in November 2004 provides a way to broadly search for scholarly literature. It sorts articles by weighing the full text of each article, the author, the publication in which the article appears, and how often the piece has been cited in other scholarly literature. Web of Science is a popular bibliographic source containing journal citation reports and other resources. Information about the impact factor of 8,500 journals is provided through ISI Web of Knowledge. Additionally, other search tools and bibliographic sources are available like ACM Digital Library (http://portal.acm.org/dl.cfm), IEEE Xplore (http://ieeexplore.ieee.org) and SpringerLink (http://www.springerlink.com/home/main.mpx).

In the work of Judit Bar-Ilan [Bar-Ilan, J. 2006] a citation study for a specific author was accomplished using several citation databases implementing the advantages and shortcomings of each tool. It was concluded that the different collection and indexing policies of the different databases led to considerably different results.

Different tools providing information to users respond successfully in an explicit query. A researcher is found of information retrieval tools. As the list of search tools and available databases grows larger every day a researcher becomes overloaded with available information. The goal of a Recommender System in this field would be to eliminate this excess of available papers and be able to propose a list of most important papers to the researcher that would match his/her interests.

Classic recommendation techniques like collaborative filtering have been used and evaluated on the recommending citations for research papers [McNee, S. M. et al. 2002]. Given a target research paper, traditional recommender algorithms are able to propose citations as suitable additional references. Hybrid algorithms have also been tested for recommending research papers [Torres, R. et al. 2004].

The focus of this paper is to propose a methodology capable of identifying important for a researcher bibliography from a list of possible alternatives and eliminate by this way the vast amount of available scientific material. Our goal is to provide a sorting of most preferred scientific literature according to each user preferences.

4. Applying the multiple criteria paradigm

Multiple Criteria Analysis is a well established field in Decision Science and it comes into a large variety of theories, methodologies, and techniques [Figueira, J. et al. 2005]. In this study, we exploit the approach according to which, Multiple Criteria Analysis is a set of methods or models enabling the aggregation of multiple evaluation criteria to support the decision(s) related to a set of actions A.

We follow the general modelling methodology of decision-making problems as proposed by Roy back in 1985 [Roy, B. 1985]. This seminal methodology is expressed in the four following levels:

- Level 1: Object of the decision, including the definition of the set of potential actions (alternatives) A and the determination of a problem statement on A.
- Level 2: Modelling of a consistent family of criteria assuming that these criteria are non-decreasing value functions, exhaustive and non-redundant.
- Level 3: Development of a global preference model, to aggregate the marginal preferences on the criteria.
- Level 4: Decision-aid, based on the results of level 3 and the problem statement of level 1.

4.1 Definition of the decision object (Level 1)

Roy [Roy, B. 1985] distinguishes four reference problem statements (choosing, sorting, ranking and describing the actions), each of which does not necessarily preclude the others. In our case, the ultimate problem is to rank research papers in order to recommend them to the end users. Such an output (a visible order of the papers) will allow the user to have a sharp view on the results; thus to counterbalance any possible inconsistencies in the order proposed by the automated method.

4.2 Modelling the criteria (Level 2)

In most cases, the first step of a decision maker (DM) consists of building n criteria with $n>1$. They constitute what we call the family F of criteria. In order to ensure that F is able to play its role in the decision aiding process correctly, the criteria considered all together need to satisfy some logical requirements which are: monotony; exhaustiveness; cohesiveness; and non redundancy. Under these requirements we ensure consistency of the family.

Each criterion g_i is a non-decreasing real valued function defined on A, where A is the set of potential actions (or objectives, alternatives, decisions) as follows:

$$g_i : A \to \left[g_{i*}, g_i^* \right] \subset \Re / \alpha \to g(\alpha) \in \Re \tag{2}$$

where $\left[g_{i*}, g_i^* \right]$ is the criterion evaluation scale, g_{i*} and g_i^* are the worst and the best level of the i-th criterion respectively, g_i is the evaluation or performance of action α on the i-th criterion and $g(\alpha)$ is the vector of performances of action α on the criteria vector.

From the above definitions, the following preferential situations can be determined:

$$\begin{cases} g_i(\alpha) > g_i(b) \Leftrightarrow \alpha \succ b \ (\alpha \text{ is preferred to b}) \\ g_i(\alpha) = g_i(b) \Leftrightarrow \alpha \ \square \ b \ (\alpha \text{ is indifferent to b}) \end{cases} \tag{3}$$

It is important to observe that none of the above requirements implies that the criteria of F must be independent. The concept of independence is very complex, and if dependence is desirable, it is necessary to specify what type of independence is needed. Multicriteria analysis has led to important distinctions between structural independence, preferential independence, and utility independence.

In the context of recommending papers to a researcher the following criteria are considered. They constitute a consistent set and represent, to the best of our knowledge, the most characteristic principles for academic paper evaluation.

g_1 – Publication year: We advocate that the most recent the publication is, the most interesting would be for the user. We propose the metric for this criterion to be the result of the subtraction Present Year – Publication's Year. A 3-fold scale is attributed to this criterion considering the fact that the recency of the publication must counterbalance the two year period needed for a publication to be cited [Garfield, E. 2003].

g_2 – Keywords Relevance: This criterion refers to the actual content of the publication. We ask the user to input its preferable keywords that he/she would look for in the publications' corpus. The method uses the TF algorithm [Salton, G. et al. 1986] to calculate the frequency in which those terms appear in the document. Finally the criterion metric is calculated based on the following equation:

$$g_2 = average(\sum_n tf_i) * (1 + \log(N)) \tag{4}$$

where tf_i is the result of the TF algorithm for the i-th term (keyword) and N is the total count of the terms. The second factor of the above equation is necessary in order to incorporate with the fact that the greater the number of the keywords the user is searching, the smallest their average frequency in the documents will be.

g_3 - Impact Factor of publication's journal: The impact factor of the publication's journal is chosen as an individual criterion for the analysis. Since it's conception by Eugene Garfield (http://www.garfield.library.upenn.edu/), the impact factor as a bibliometric indicator is widely utilized by researchers and institutions [Garfield, E. 2003]. All the necessary information was mined from the ISI's Web of Knowledge database. As of March 2006, the total number of Web of Science records is approximately 37 million. Here the journal's impact factor is normalized with the average impact factor of the journal's category.

Data was retrieved from all of the three major citation databases, each used according to problems' convenience. Web of science was used to extract the journal's impact factor, Citeseer to find the authors' citation identity (see g_5 below) and Google Scholar to locate the citation index of each publication (g_4).

g_4 - Citation Index of paper: As stated above, Google Scholar was used to determine the citation index of each paper. The actual number of citations for each paper is used as an input parameter for the assessment of the corresponding criterion.

g_5 – Author(s): This metric is calculated as the average of the citations of the author and the co-authors. The number of citations of each author was retrieved from the Citeseer database.

g_6 – Acknowledgments: Acknowledgment, usually stated at the end of publications is an expression of appreciation for assistance or contribution in the completion of the study described in that paper. According to [Giles, C. L. et al. 2004], acknowledgments may be made for a number of reasons but often imply significant intellectual debt. Just as citation indexing proved to be an important tool for evaluating research contributions, acknowledgments can be considered a metric parallel to citations in the academic review process.

Not all publications enclose acknowledgments. Thus, all the examined papers are categorized into a 3-fold quantitative scale. The word "top" denotes that the acknowledged entity is included in the 15 most acknowledged entities according to the list of [Giles, C. L. et al. 2004]. These entities are categorized as follows: funding agencies, companies, educational institutions, and individuals. The word "yes" implies that although acknowledgments exist, they are not included in the top 15 list mentioned above. In the case that no acknowledgements appear in the bulk of the paper we assign the word "no".

g_7 – Affiliation: The authors' affiliation is considered as a criterion for the ranking procedure as well. The new edition of the Academic Ranking of World Universities of the Shanghai Jiao Tong University, published during August 2006 (http://ed.sjtu.edu.cn/ranking.htm), was used. The 500 highest ranked universities was divided into 4 categories.

A 3-fold scale was used for the first 6 criteria and a 4-fold for the last criterion, concerning the affiliation. The scale fold is decided by the DM and the discritization of the interval $\left[g_{i*}, g_i^*\right]$ is calculated using linear interpolation according to the following equation in case of quantitative criteria:

$$g_i^j = g_{i*} + \frac{j-1}{a_i - 1}(g_i^* - g_{i*}) \qquad \forall \; j = 1, 2, ...a_i \tag{5}$$

4.3 Global Preference model (Level 3)

In order to construct the global preference model the disaggregation-aggregation approach [Siskos, Y. et al. 2005] is exploited. This approach aims at analyzing the behaviour and the cognitive style of the user. Special iterative interactive procedures are used, where the components of the problem and the user's global judgment policy are analyzed and then they are aggregated into a value system.

The UTASTAR methodology [Siskos, Y. et al. 1985] is applied here, an improved version of the original UTA model, a regression based approach that has been developed as an alternative to Multiattribute Utility Theory (MAUT) [Keeney, R. et al. 1993]. UTA (UTilitès Additives) methods not only adopt the aggregation-disaggregation principles, but may also be considered as the main initiatives and the most representative examples of preference disaggregation theory.

The UTASTAR regression aims to estimate additive utilities:

$$U(\underline{g}(\alpha)) = \sum_{i=1}^{m} u_i(g_i) \tag{6}$$

subject to the following constrains:

$$\begin{aligned} & u_i(g_{i*}) = 0 \quad \forall i \\ & \sum_{i=1}^{m} u_i(g_i^*) = u_1(g_1^*) + u_2(g_2^*) + ... + u_m(g_m^*) = 1 \end{aligned} \tag{7}$$

The goal of the UTASTAR method is to guide the decision maker to a process of gradual learning his/her preferences.

Initially, the user is asked to verify his/her global preferences on the alternatives (publications) set taking into account the performances of the reference alternatives on all the criteria. The set of the alternatives is comprised by a set of familiar publications (training set), each having a specific vector of performance on the criteria. These indicative publications can be easily judged by the user to perform global comparisons and to match a ranking among them.

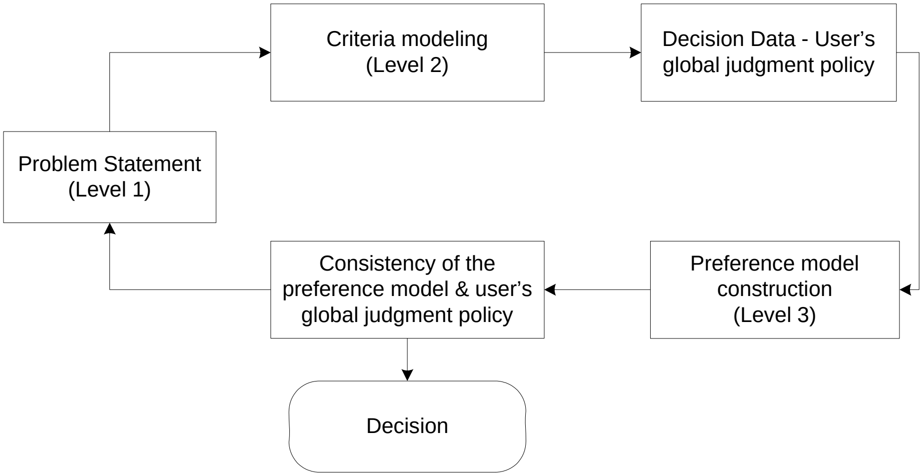

Figure 1. ***The aggregation-disaggregation approach [Figueira, J. et al. 2005]***

Thus, having such a weak-order preference structure on a set of publications (the weak- order is achieved through equations (3)) the methodology invokes the UTASTAR methodology [Siskos, Y. et al. 1985] to adjust an additive utility function, namely a global preference model. The UTASTAR methodology provides self-contained consistency and efficiency measurement techniques such as the Kendall's τ between the initial weak order and the one produced by the estimated model. This method has been chosen particularly because of its ability to enable each individual decision maker to analyze his/her behaviour and cognitive style according to the general framework of preference disaggregation approach [Siskos, Y. 1980; Jacquet-Lagrèze, E. et al. 1982; Jacquet-Lagrèze, E. et al. 2001]. The goal of this approach is to aid the decision maker (DM) to improve his/her knowledge about the decision status and also about his/her preference/value system in order to reach a consistent decision. The acceptance of such a preference model is accomplished through a repetitive interaction between the model and the DM. This means that the DM is able to change either the initial ranking on the training set of alternatives or the number and the context of the criteria or add/remove alternatives. This iterative procedure is continued until the DM's preferences totally converge with the proposed model. The assumption of the UTASTAR method is that the model of DM's preferences is additive, which of course is not true in all decision problems. However, the assumption of a linear preference system simplifies the problem and makes the assessment of the DM's preference system easier.

The disaggregation – aggregation methodological framework is depicted in Figure 1.

4.4 Provide a final decision Support (Level 4)

The ultimate outputs of the UTASTAR method are the marginal value functions for every criterion and its weight of significance against the others. Exploiting these results, the user can assess a value (U_i) for each new publication he would like to evaluate. By repeating this step as many times as needed to evaluate all the candidate publications, the user will end up with an ordered list of publications. Then he could rationally select the top-ones and focus on them.

5. Results - An illustrative example

As an implementation of the methodology described in this paper an example of recommending scientific papers to the academic community is examined. A set of familiar for the DM articles was used initially to trigger the UTASTAR methodology. The criteria for the evaluation of the research papers are mentioned in section 4. The DM provides a weak order ranking of preference over this demo set of articles. Through this order and after the invocation of UTASTAR mehtodology the marginal utilities of each criterion are calculated.

The marginal utility functions for the criteria “keywords” and “affiliation” are presented below as indicative results.

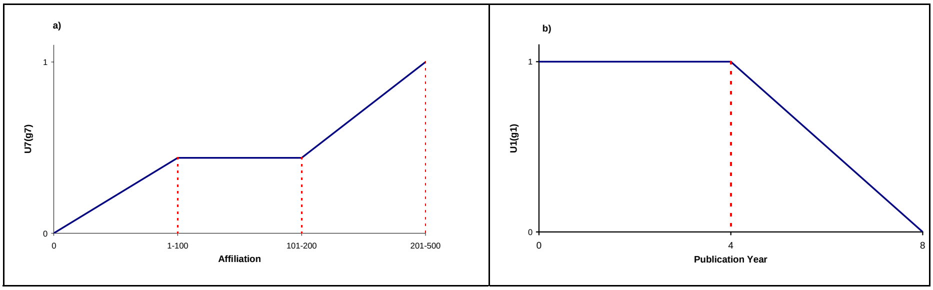

***Figure 2.** Marginal Utilities of criteria a) “keywords” and b) “affiliation”*

Since the weights of each criterion are already known from previous steps, we may thus calculate the utility function for each alternative on a new set of 9 unread this time papers.

The only necessary information for this part is the evaluation of the alternatives on each criterion. For instance, for each paper we need just to know the impact factor of the journal (g_3) or the number of citations of the authors (g_5) etc. Let us use the code names PAPER1 [Yukun, C. et al. 2007]; PAPER2 [Tang, T. et al. 2004]; PAPER3[Montaner, M. et al. 2003]; PAPER4 [Adomavicius, G. et al. 2005]; PAPER5 [Yang, F. et al. 2005]; PAPER6 [Yager, R. R. 2003]; PAPER7 [Schafer, J. B. et al. 2001]; PAPER8 [Herlocker, J. L. et al. 2004]; PAPER9 [Ha, S. H. 2006] for the unrated set of alternatives. The utility for each alternative is calculated as the sum of its marginal utilities over the criteria set. Table 1 shows the sorting of the alternatives in descending utility order and the weights of significance for each criterion.

Table 1. *Criteria weights results and final ranking of the alternatives*

a) Criteria weights

Criterion	*Weight of significance*
g1	*0,235*
g2	*0,060*
g3	*0,018*
g4	*0,309*
g5	*0,045*
g6	*0,090*
g7	*0,243*

b) Final ranking

Alternative	*Final score*
U(PAPER2)	*0,784446*
U(PAPER7)	*0,731583*
U(PAPER9)	*0,666251*
U(PAPER4)	*0,566788*
U(PAPER8)	*0,562758*
U(PAPER5)	*0,522031*
U(PAPER1)	*0,443493*
U(PAPER6)	*0,440748*
U(PAPER3)	*0,359103*

Knowing the weights of significance and thus the importance of each criterion the DM is able to evaluate any unknown set of scientific papers.

In the specific example the user is provided with a descending preference order of nine unknown scientific papers. Seven criteria were used to evaluate these alternatives.

6. Conclusions

By applying techniques and methods from the field of Multi-Criteria Decision Aiding, a new approach to overcome the lack of user-system interaction in existing Recommender Systems is proposed. A methodology for recommending scientific publications to the research community was demonstrated as a pioneering example for the incorporation of MCDA techniques to Recommender Systems. More specifically the UTASTAR algorithm was applied to rank scientific publications according to the researcher's unique preferences. The specific algorithm is extensively studied and applied in Operations Research. It takes as an input the performances of the alternatives (publications to be ranked) in all the predefined criteria, as well as an initial ranking of these publications by the DM and provides as an output the weights of significance for each criterion and the global utilities of the alternatives. Once the algorithm is trained according to DM's preferences it can be used to rank any unread assembly of publications, given the performances upon the criteria. It is crucial to mention at this point that the methodology proposed herein is user subjective since the weak preference order of the alternatives is derived by analysing the initial ranking provided by the user. A different initial ranking of the publications given by the user or a different formulation of the criteria may lead to totally different results. This is an important feature of the current methodology since it provides the user the option to evaluate the outcome and if she/he is not satisfied to change the above mentioned parameters (initial order or criteria formulation) until an adequate result is obtained. With the proposed system a researcher may choose among the vast amount of literature available, the best publications concerning his/her preferences and focus explicitly on them. The discussed system offer the user/decision maker the opportunity to express his/her dynamic intentions/preferences and in that sense it is utterly user-oriented. It considers the decision maker as an individual distinctive personality and relies explicitly on him/her to attain all the necessary for the methodology information. It is mentioned herein that different DM's may provide different model formulation and thus attain respective results. The proposed system does not need the contribution of other people (collaborative filtering) or any advanced content analysis tools to perform. Our prospect is to be able to provide automated feedback for the decision maker that will optimize the current methodology by incorporating specified learning procedures in the recommendation process.

Acknowledgements

This work was financially supported by The Hellenic Ministry of Development, General Secretariat for Research and Technology and the European Commission under the project PENED 2003-Operational Program Competitiveness.

References

Adomavicius, G. and Tuzhilin, A. (2005), *Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions,* IEEE Transactions on Knowledge and Data Engineering vol. 17 (6), pp. 734-749.

Balabanovic, M. and Shoham, Y. (1997), *Fab: content-based, collaborative recommendation,* Commununications of the ACM vol. 40 (3), pp. 66-72.

Bar-Ilan, J. (2006), *An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes,* Information Processing and Management vol. 42 (6), pp. 1553–1566.

National Science Board (2006). *Academic Research and Development*, Science and Engineering Indicators 2006, Arlington, VA, National Science Foundation. vol.1, pp. 5.6-5.7.

Breese, J., Heckerman, D. and Kadie, C. (1998), *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*, in Proc. Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, Madison, Wisconsin.

Burke, R. (2002), *Hybrid Recommender Systems: Survey and Experiments* User Modelling and User-Adapted Interaction vol. 12 (4), pp. 331-370.

Figueira, J., Greco, S. and Ehrgott, M., Eds. (2005). *Multiple criteria decision analysis: State of the art surveys*. Boston, Springer.

Garfield, E. (2003), *The meaning of the Impact Factor,* RIPCS/IJCHP vol. 3 (2), pp. 363-369.

Giles, C. L., Bollacker, K. D. and Lawrence, S. (1998), *CiteSeer: An Automatic Citation Indexing System*, in Proc. Digital Libraries 98 - Third ACM Conference on Digital Libraries, New York, ACM Press.

Giles, C. L. and Councill, I. G. (2004), *Who gets acknowledged: Measuring scientific contributions through automatic acknowledgment indexing,* PNAS vol. 101 (51), pp. 17599–17604.

Goldberg, D., Nichols, D., Oki, B. M. and Terry, D. (1992), *Using Collaborative Filtering to Weave an Information Tapestry,* Comm. ACM vol. 35 (12), pp. 61-70.

Ha, S. H. (2006), *Digital Content Recommender on the Internet,* IEEE Intelligent Systems vol. 21 (2), pp. 70-77.

Herlocker, J. L., Konstan, J. A., Terveen, L. G. and Riedl, J. T. (2004), *Evaluating collaborative filtering recommender systems,* ACM Trans. Inf. Syst. vol. 22 (1), pp. 5-53.

Jacquet-Lagrèze, E. and Siskos, Y. (1982), *Assessing a set of additive utility functions for multicriteria decision making: The UTA method,* European Journal of Operational Research vol. 10 (2), pp. 151–164.

Jacquet-Lagrèze, E. and Siskos, Y. (2001), *Preference disaggregation: 20 years of MCDA experience,* European Journal of Operational Research vol. 130 (2), pp. 233–245.

Keeney, R.L. and Raiffa, H. (1993). *Decision with Multiple Objectives: Preference and Value Tradeoffs*. NY, Cambridge University Press.

McNee, S. M., Albert, I., Cosley, D., Prateep Gopalkrishnan, Lam, S. K., Rashid, A. M., Konstan, J. A. and Riedl, J. (2002), *On the Recommending of Citations for Research Papers*, in Proc. CSCW'02, New Orleans, Louisiana, USA.

McNee, S. M., Riedl, J. and Konstan, J. A. (2006), *Making Recommendations Better: An Analytic Model for Human-Recommender Interaction*, in Proc. ACM Conference on Human Factors in Computing Systems (CHI 2006), Montreal, Canada.

Melville, P., Raymond J. Mooney and Nagarajan, R. (2002), *Content-Boosted Collaborative Filtering for Improved Recommendations* in Proc. AAAI-2002, Edmonton, Canada.

Montaner, M., Lopez, B. and Rosa, J. L. D. L. (2003), *A Taxonomy of Recommender Agents on the Internet,* Artificial Intelligence Review vol. 19, pp. 285–330.

Mooney, R. J. and Roy, L. (2000), *Content-based book recommending using learning for text categorization* in Proc. The fifth ACM conference on Digital libraries, San Antonio, Texas, United States, ACM Press.

Pazzani, M., J, M. and Billsus, D. (1996), *Syskill & Webert: Identifying interesting web sites*, in Proc. The National Conference on Artificial Intelligence, Portland, OR.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J. (1994), *GroupLens: An open architecture for collaborative filtering of Netnews*, in Proc. ACM conference on Computer supported cooperative work, Chapel Hill, North Carolina, United States ACM Press.

Resnick, P. and Varian, H. R. (1997), *Recommender systems,* Commun. ACM vol. 40 (3), pp. 56-58.

Roy, B. (1985), *Méthodologie Multicritère d'Aide à la Décision,* Economica.

Salton, G. McGill, M. J. (1986). *Introduction to Modern Information Retrieval*. NY, USA, McGraw-Hill, Inc.

Schafer, J. B., Konstan, J. A. and Riedl, J. (2001), *E-Commerce Recommendation Applications,* Data Min. Knowl. Discov. vol. 5 (1-2), pp. 115-153

Shardanand, U. and Maes, P. (1995), *Social information filtering: Algorithms for automating 'word of mouth'*, in Proc. Conference on Human Factors in Computing Systems, Denver, Colorado, USA.

Siskos, Y. (1980), *Comment modéliser les préférences au moyen de fonctions d'utilité additives,* RAIRO Recherche Opérationelle vol. 14, pp. 53–82.

Siskos, Y., Grigoroudis, E. and Matsatsinis, N. (2005), *UTA Methods*. Multiple criteria decision analysis: State of the art surveys. Figueira, J.et al. Boston, Springer, pp. 297-344.

Siskos, Y. and Yannacopoulos, D. (1985), *UTASTAR, An ordinal regression method for building additive value functions,* Investigaçao Operacional vol. 5 (1), pp. 39–53.

Tang, T. and McCalla, G. (2004), *Beyond Learners' Interest: Personalized Paper Recommendation Based on Their Pedagogical Features for an e-Learning System,* Lecture Notes in Computer Science pp. 301-310.

Terveen, L., Hill, W., Amento, B., McDonald, D. and Creter, J. (1997), *PHOAKS: A System for Sharing Recommendations,* Communications of the ACM vol. 40 (3), pp. 59-62.

Torres, R., McNee, S. M., Abel, M., Konstan, J. A. and Riedl, J. (2004), *Enhancing digital libraries with TechLens*, in Proc. Joint ACM/IEEE Conference on Digital Libraries, Tuscon, AZ, USA.

Yager, R. R. (2003), *Fuzzy logic methods in recommender systems,* Fuzzy Sets and Systems vol. 136, pp. 133–149.

Yang, F., Han, P., Shen, R. and Hu, Z. (2005), *A Novel Resource Recommendation System Based on Connecting to Similar E-Learners,* Lecture notes in computer science vol. 3583, pp. 122-130.

Yukun, C. and Yunfeng, L. (2007), *An intelligent fuzzy-based recommendation system for consumer electronic products,* Expert Systems with Applications vol. 33 (1), pp. 230-240.

Zaslow, J. (2002), *If TiVo Thinks You Are Gay, Here's How To Set It Straight* The Wall Street Journal (sect. A).