Recommending scientific papers through a method based on bibliometric measures

A. Tejeda-Lorente[1] **C. Porcel**[2] **J. Bernabé-Moreno**[1] **E. Herrera-Viedma**[1]

[1]University of Granada, Department of Computer Science and Artificial Intelligence, Granada, Spain.
[2]University of Jaén, Department of Computer Science, Jaén, Spain.

Abstract

We present a quality-based fuzzy linguistic recommender system for researchers. We propose the use of some bibliometrics measures as the way to quantify the quality of both items and users. The system takes into account the measured quality as the main factor for the re-ranking of the top-N recommendations list in order to point researchers to the latest and the best papers in their domain. To prove the accuracy improvement, we conduct a study involving different recommendation approaches. The results obtained proved to be satisfactory within the research departments who took part on the tests.

Keywords: Recommender systems, item quality, fuzzy linguistic modeling, papers recommendation

1. Introduction

Research is a key task in society nowadays. Science is going through difficult times with the actual crisis facing significant budget cuts in modern economies [1]. Therefore, the use of scientific knowledge by setting up a sustainable industry/science cooperative environment positively affects innovation performance [2]. But nowadays, the spectacular growth of sources providing new information introduces noise to our methods of finding relevant information for research. Tools such as Thomson Reuters Web of Science (WOS) [3], Scopus [4], Sciencedirect [5] or Google Scholar [6], support the research in fields where the volume of publications is vast. However, the ever increasing number of resources in a simple area might make the information consumers misunderstand the relevance of certain research resources, at the risk of missing important information. Staying in the loop of all new papers being published in all the new journals could be troublesome. In this sense, it is essential to have tools that allow us to access to updated and personalized information according to our interests.

A *recommender system* seeks to discover information items that are valuable to the user. Recommender systems may be considered personalized services because they have an independent profile for each user taking into account the particularities of each of them. Due to these reasons, these kinds of systems are becoming popular tools for reducing information overload and to improve the con-version rate in e-commerce web sites [7]. Different approaches could be used to generate recommendations. All of them have their advantages and disadvantages, so that a widely used solution is the *hybrid approach*, to minimize the disadvantages while maximizing the benefits [8]. The representation of user preferences as well as other characteristics of some items presents subjectivity and uncertainty making the process a complex task. In order to deal with that, fuzzy linguistic modeling has been proved as an efficient tool to face the problem [8].

Another aspect to consider is the quality of recommended items. In [9] we presented a quality-based recommender system to disseminate information in a University Digital Library (UDL) where we combined the quality of an item with the user preferences. Capitalizing on the improvement we reached with our system in the scope of UDL to tackle the information overload issue, we want to expand the scope to the much larger research domain. The heterogeneity of information sources from elements coming from a much larger set of database is going to be a bigger challenge compared with the work done for the UDL. Hence, a first step to expand the recommendations in research based on quality is to change not the scope but the database. Instead of using a UDL we are going to use the papers indexed by WOS. A second step is to change the old concept of quality applied in [9] and how to evaluate it.

In this paper, we present a new quality-based fuzzy linguistic recommender system which incorporates methods, algorithms and filters for all the steps in the recommendation process, focused on the re-ranking stage. The system allows the user to be up-to-date regarding all articles that might be relevant for him related to the topics the user manifested interest at a particular moment in a real time window. In addition to the articles that are identified to be of interest for the user, a set of high-quality articles with a certain overlap with the researcher topic are also provided to open up new possibilities and therefore enrich the on-going research. The main novelties of the system are listed below:

- The system implements a *hybrid recommendation strategy* based on a switching hybrid approach [10] to share the user's individual experience and social wisdom and to present the

best recent papers in each area for each user.

- The system uses fuzzy linguistic modeling to improve the representation of user preferences and facilitate user-system interactions [11].
- The system implements a two-phase *feedback process*. They can say if a paper is relevant or not, that means evaluate the recommendation, and then when they read the paper they can evaluate their quality giving it different relevance for the collaboration step.
- The resources recommended in our system are now the *most recent research papers* from one of the most important international science database, WOS by Thomson Reuters, where we can find the best journal and conference essentials for the daily work of a researcher. The items recommended fulfill the necessities of the users since the quality and novelty of the items are guaranteed by the system.
- The system incorporates a *re-ranking* process which takes into account the estimated relevance of an item along with the item quality as well as the users' quality, to re-order the list of possible items to be recommended. This is done, taking into account not only these factors but also others to ensure the items are interesting for the researcher, but also with a small innovative percentage allowing to see the different application of one singular topic.
- The quality of research resources and researchers are evaluated using different rankings provides in the international community by experts, i.e., the Journal Citation Report(JCR) provided by Thomson Reuters[1] for the items, and the H-index [12] for the researchers.

The paper is structured as follows. In Section 2, the preliminaries are presented. Next in Section 3, we describe the proposal, including the evaluation of the system. Finally, some concluding remarks and future works are pointed out in Section 4.

2. Preliminaries

2.1. Recommender systems

Recommender systems try to guide the user in a personalized way towards suitable tasks among a wide range of possible options [7]. Personalized recommendations rely on knowing users' characteristics, which might be tastes, preferences about items as well as the ratings of previously explored items. The system has to maintain users' profiles updated in order to provide good recommendations, the way of acquiring this information may vary from implicit information, that is, analyzing users behavior, or explicit information, where users directly provide their preferences.

On the design of a recommender system an aspect to take into account is the way of generating rec-

ommendations. In the literature we can find them mainly pooled in two categories. In the first one [13] authors consider two different approaches: On one side, *Content-based systems* where they generate the recommendations taking into account the characteristics used to represent the items and the ratings that a user has given to them. On the other side, *Collaborative systems* where the system generates recommendations using explicit or implicit preferences from many users, ignoring the items representation. The second one extends the categorization with another three approaches: *Demographic systems*, *Knowledge-based systems* and *Utility-based systems* [7, 10].

Since each approach has certain advantages and disadvantages, depending on the scope settings, the most adopted solution addressed in the literature is the combination of the previous approaches in what is known as an *hybrid recommender system* [10]. The aim of this hybridization is to combine different approaches to reduce the disadvantages of each one and to exploit their benefits.

As we have commented, we focus on the recommendation of scientific papers. In this scope, recommender systems are very useful helping users to access to relevant information. In fact, we can observe the proliferation of these services in last years, sucha as Google Scholar [6], ResearchGate [14] or ScienceDirect [5] and several proposals found in the literature. In [15] the authors develop a machine learning algorithm for recommending scientific articles to users in an online scientific community; the proposed algorithm uses two types of data: the other users' libraries and the content of the articles. They takes into account both older papers that are important to other similar users and newly written papers whose content reflects the user's specific interests. Their approach combines collaborative filtering based on latent factor models and content analysis based on probabilistic topic modeling. By other hand in [16] the authors focus on multiple criteria for recommending scientific papers; they focus on applying a layered evaluation framework to fit the interaction components of a recommender system. More recently, we can find a new proposal [17], in which the authors consider the domain of scientific publications repositories and propose a content-based recommender based upon a graph representation of concepts built up by linked keyphrases; further, they incorporate a keyphrase extraction system able to generate meaningful metadata for the documents, very useful for providing helpful and explainable recommendations.

But all these proposals use different databases and we propose to use one of the most important and standardized tool, i.e. the papers indexed by WOS. Besides, since the novelty of the items is strongly constraint in the goal of this system, collaborative approaches could have difficulties generating recommendations. For this reason, we pro-

---

[1]http://thomsonreuters.com/journal-citation-reports/

pose to use the *mixed hybrid strategy*, i.e. finding the balance between novelty and importance of non recommended items that other users consider important in a time window. Due to the fact that the items can be recommended when they are inserted into the system in a determined time, the use of collaborative filters is not suitable. This technique is used in a lower number in the following round to recommend possible relevant paper that the system could not recommend due to the restriction of the number of recommendations per user.

The recommendation activity is complete with two more phases: *re-ranking* and *feedback*. The former is oriented to increase the accuracy of recommendations; the re-ranking involves a different criteria to order the items in the final list of recommendation i.e. popularity, average rating, likeability. Counter to this objective, in this work the use of re-ranking is oriented not to increase the diversity but the accuracy. In the latter, feedback is a cyclic process whereby the users provide the system with their evaluations about the recommended items and the system uses these evaluations to automatically update user profiles.

2.2. Fuzzy linguistic approach

The fuzzy linguistic approach is a tool based on the concept of linguistic variable proposed by Zadeh [18]. This theory has given very good results to model qualitative information and it has been proven to be useful in many problems. We describe the 2-tuple fuzzy linguistic modeling and multi-granular fuzzy linguistic approach used to represent the linguistic information used in the system.

2.2.1. The 2-Tuple Fuzzy Linguistic Approach

In order to reduce the loss of information of other methods such as classical or ordinal, in [19] is proposed a continuous model of information representation based on 2-tuple fuzzy linguistic modeling. To define it both the 2-tuple representation model and the 2-tuple computational model to represent and aggregate the linguistic information have to be established.

Let $S = \{s_0, ..., s_g\}$ be a linguistic term set with odd cardinality. We assume that the semantics of labels is given by means of triangular membership functions and consider all terms distributed on a scale on which a total order is defined. In this fuzzy linguistic context, if a symbolic method aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, ..., g\}$, we can represent β as a 2-tuple (s_i, α_i), where s_i represents the linguistic label, and α_i is a numerical value expressing the value of the translation between numerical values and 2-tuple: $\Delta(\beta) = (s_i, \alpha)$ y $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ [19].

In order to establish the computational model negation, comparison and aggregation operators are defined. Using functions Δ and Δ^{-1}, any of the existing aggregation operators can be easily be extended for dealing with linguistic 2-tuples without loss of information [19]. For instance arithmetic mean, weighted average operator or linguistic weighted average operator could be used.

2.2.2. Multi-Granular Linguistic Information Approach

A problem modeling the information arises when different experts have different uncertainty degrees on the phenomenon (for instance an expert has less uncertainty than a novice), then several linguistic term sets with a different granularity of uncertainty are necessary. The use of different label sets to assess information is also necessary when an expert has to evaluate different concepts (for instance importance, relevance, and so on). In such situations, we need tools to manage multi-granular linguistic information [20, 21].

In [20] a multi-granular 2-tuple fuzzy linguistic modeling based on the concept of linguistic hierarchy is proposed. A *Linguistic Hierarchy*, LH, is a set of levels $l(t, n(t))$, where each level t is a linguistic term set with different granularity $n(t)$. In [20] a family of transformation functions between labels from different levels was introduced. To establish the computational model we select a level that we use to make the information uniform and thereby we can use the defined operator in the 2-tuple model. This result guarantees that the transformations between levels of a linguistic hierarchy are carried out without loss of information.

2.3. On evaluation of quality of research resources and authors

The proliferation of different scientific repositories provides access to research resources gathering the most important database in different areas, e.g. Scopus[4] and WOS [3]. All of this repositories provide different search tools to find the information sought. To measure the quality of a scientific publication is a key task for our purpose, for that reason we assume some bibliometric indicators which are commonly extended in different fields for the quality measure of them [22].

The most widely adopted method is to use Garfield's Impact Factor (IF), i.e., the average number of times the published papers are cited up to two years after publication [23].

For authors there are different measures that allow us to analyze them individually. In the literature we can find some metrics also based on citations. Some of the most important indicators are the H-index [12], G-index [24] and H-classics [25].

3. System description

Our proposal [2] is a hybrid recommender system which takes into account users' and items' quality when it comes to the re-ranking stage. The system delivers personalized emails monthly to each user with a selection of the best latest papers for him. In Figure 1 is shown the structure of the system.
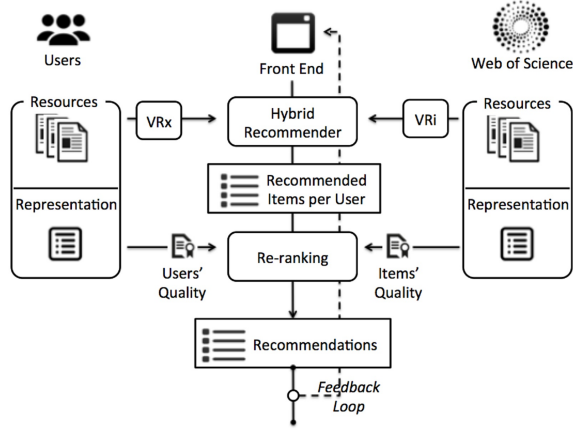

Figure 1: Structure of the system.

3.1. Information representation

In order to represent the different concepts to be assesses by the system, we will use different label sets $(S_1, S_2, ...)$ selected from a LH [20]. The system works with the following concepts:

- **Importance degree** of a discipline with respect to a resource scope, based on the paper's keywords, which is assessed in S_1.
- **Similarity degree** among resources or among users, which is assessed in S_2.
- Predicted **relevance degree** of a resource for a user, which is assessed in S_3.
- **Satisfaction degree** expressed by a user to evaluate the recommendation of one resource, which is assessed in S_4.
- **Quality degree** expressed by a user to evaluate a recommended resource, which is assessed in S_5.

We propose to use a LH composed by the following linguistic terms sets:

- $S^2 = \{a_0 = None = N, a_1 = Total = T\}$.
- $S^3 = \{b_0 = None = N, b_1 = Medium = M, b_2 = Total = T\}$
- $S^5 = \{c_0 = None = N, c_1 = Low = L, c_2 = Medium = M, c_3 = High = H, c_4 = Total = T\}$
- $S^9 = \{d_0 = None = N, d_1 = Very_Low = VL, d_2 = Low = L, d_3 = More_Less_Low = MLL, d_4 = Medium = $

$M, d_5 = More\_Less\_High = MLH, d_6 = High = H, d_7 = Very\_High = VH, d_8 = Total = T\}$

Taking into account the results of several experiments performed to evaluate the user satisfaction, we use the level 4 (9 labels) to represent the similarity degrees and predicted relevance degrees ($S_2 = S^9$ and $S_3 = S^9$), the level 3 (5 labels) are used for the importance degrees ($S_1 = S^5$), the level 2 (3 labels) to represent the quality degree given by a user over a paper ($S_5 = S^3$) and for the satisfaction degree we use the level 1 ($S_4 = S^2$).

3.1.1. Resources representation

The considered resources are the papers imported from WOS monthly. The journals are selected from the subcategories of the JCR of ISI WOS. The main characteristics we use to represent an item are *Title, Authors, Journal, Year* and *time cited*. Additionally, we consider extra characteristics at journal level, like as the Quartile where the journal is indexed in JCR, the impact factor, eigenfactor or characteristics referred to the authors, such as city, email, university, etc.

Once the system has imported all available information about a resource, an internal representation based on the keywords of the paper itself is created. We use the vector model to represent a resource, i.e., a paper i is represented as a vector $VR_i = (VR_{i1}, VR_{i2}, ..., VR_{in})$, where n is the number of keywords and each component $VR_{ij} \in S_1$ is a linguistic assessment of the importance degree of the each keyword. All the keywords are set automatically with the label c_4 of S^5.

3.1.2. User profiles representation

Users' profiles are based on their own papers and papers they are interested in, along with keywords and their importance leading the users' interests. The system automatically tries to identify the resources where the user appears as author or co-author performing a search in WOS *papers* to then import them into the system. The five most important ones according to the quality estimated are then selected. After that, from the five selected papers of a user, the keywords are extracted and flagged as important keywords for this user with the label c_2 of S^5 pending for the user revision.

To represent a user profile i we also use a vector $VU_i = (VU_{i1}, VU_{i2}, ..., VU_{in})$, where n is the number of assigned keywords and each component $VU_{ij} \in S_1$ is a linguistic assessment that represents the importance degree of keyword j for the user i.

3.2. Recommendation scheme

We implement a hybrid recommendation strategy, which combines a content-based recommendation

---

approach and a collaborative one, cooperating both in different percentage. The combination of both occurs after the re-ranking phase. This strategy is based on a matching process developed by a similarity measure. To estimate the similarity between two vectors of users or items, we adopt the standard cosine measure, but defined in a linguistic framework:

$$\sigma_l(V_1, V_2) = \Delta(g \times \frac{\sum_{k=1}^n (\Delta^{-1}(\vartheta_1) \times \Delta^{-1}(\vartheta_2))}{\sqrt{\sum_{k=1}^n (\Delta^{-1}(\vartheta_1))^2} \times \sqrt{\sum_{k=1}^n (\Delta^{-1}(\vartheta_2))^2}})$$

where $\sigma_l(V_1, V_2) \in S_2 \times [-0.5, 0.5]$, g is the granularity of the term set used to express the relevance degree, n is the the number of keywords and $\vartheta_i = (v_{ik}, \alpha_{vik})$ is the 2-tuple linguistic value of keyword k in the vector V_i representing the resource scope or user interest topics, depending of the used filtering strategy.

3.2.1. Content-Based Recommendations

When a new resource i is imported to the system and the time window of recommendations starts, the system computes a first content-based list of candidates to recommendations for a researcher e as follows:

1. Compute $\sigma_l(VR_i, VU_e) \in S_2$.
2. Assuming that $S_2 = S^9$, we consider that a resource i is relevant for a researcher e if $\sigma_l(VR_i, VU_e) > (d_3, 0)$, i.e., if the linguistic similarity degree is higher than the label one step under the mid linguistic label.

3.2.2. Collaborative Recommendations

When the system is already operating and there are at least one evaluation from other users (after the first round), to detect similar users we follow a nearest-neighbor algorithm. Given a researcher e, the recommendations to be added to the previous list of e are obtained in the following steps:

1. Identify the set of users $\aleph_e$ most similar to e that we are estimating in the round t, where each round is the established period between recommendations. For doing this, we calculate the linguistic similarity degree between the vector representing e (VU_e) upon the vectors of all other users VU_y, $y = 1..n$ where n is the number of users, i.e., $\sigma_l(V_e, V_y) \in S_2$. As $S_2 = S^9$, we consider that the user y is near neighbor to e if $\sigma_l(VU_e, VU_y) > (d_4, 0)$.
2. Retrieve the resources positively rated previously by users similar to e, at most, two round before (t_{-1}, t_{-2}). That is, to not to break the novelty property of the recommended items.
3. Due to the singularity of the problem we are tackling, researchers can have different lines of investigation. In order to avoid the problem of recommended items which are not valid for the user e, even if they are relevant from near neighbors, the system calculates the similarity for all the items of the last two round for all the near neighbors and VU_e. Assuming that $S_2 = S^9$, we consider that a resource i is relevant for a researcher e if $\sigma_l(VR_i, VU_e) > (d_3, 0)$. The relevance degree of selected resources is the value $\sigma_l(VR_i, VU_e)$.

3.3. Computing the quality of research resources and users

We compute the quality measure using two elements: papers and authors. The former are measured by using the IF [23] of the journal as well as the quartile of JCR index they are ranked in. Furthermore, the feedback given by users about papers is also taken into account to estimate the quality of an item. The final measure is the aggregation of both measures. However the metric which contributes the most to the quality estimation of an item is the IF. The quality of each item i is estimated individually for each user j:

$$q(i, j) = (\frac{1}{Q_i} + IF_i) * G_j \qquad (1)$$

where we normalize the $q(i, j)$ to the range $[0, 1]$, Q_i is the quartile where the journal i is ranked. IF_i is its impact factor of the journal i and G_j is a constant which measures how appropriate a journal for a user j is, adopting G_j the following values: 1 if $\exists j \in Js_i$, 0.75 if $\exists T(j) \in Ts_i$, 0.5 if $\exists j \in Jns_i$, 0.25 if $\exists T(j) \in Tns_i$ and 0.1 otherwise, where Js_i is the set of journals where the selected papers of the user i are published, $T(j)$ is the function which extracts the topics of a journal j, Ts_i is the set of topics from selected journals of the user i, Jns_i are the journals non selected but part of the profile of the user i and Tns_i is the set of topics from the non selected journal of the user i.

Meanwhile, the latter can be measured by the H-index [12] of authors which quantifies the importance of each author within the research community. The quality of an author despite is independent from the rest of qualities, is weighted with the qualities of similar authors to the user we are calculating. Before of that, they are normalized to the range $[0, 1]$. So, the quality of an author i regarding to the user j is:

$$q_a(i, j) = H(i)/max(Q_a(j)) \qquad (2)$$

where $H(i)$ is the H-index of the author i and $Q_a(j)$ is the set with the H-index values for the authors that are near neighbor to j. The quality of the authors is used to measure how good a resource for a user is, when we are dealing with the collaborative approach.

3.4. Re-ranking

Once the previous lists of resources have been generated with the content-based and collaborative approaches, we proceed with the re-ranking process. The re-ranking phase in both approaches is independent from each other and consists of the reorganization of the previous lists aggregating the linguistic similarities with the scores of quality. The computation of the quality for the resources is also used to support the classification of each resource in different groups based on the selected papers by the users in their profiles. We split the previous list into four smaller lists:

1. *Same Selected Journals List.* Papers from the same journals of the group of the selected one for the user.
2. *Same Selected Topics List.* Papers from the journals with the same topic of the group of the selected one for the user.
3. *Same Non Selected Journals List.* Papers from the rest of the journals for the papers of the group of the non selected one for the user.
4. *Same Non Selected Topics List.* Papers from the rest of the journals with the same topic of the group of the non selected one for the user.

At this point a last step which affects the final relevance degree is applied. The *black list* of the combination of keywords of each user where the feedback provided was negative is compared with the keywords of each list. If the combination is found on one paper, the relevance degree is penalized to half value.

Now, the final list is obtained applying a *diversity filter* in which the percentage which each list provide to the final list was establish as follow: Same Selected Journals List - 40%, Same Selected Topics List - 40%, Same Non Selected Journals List - 20%, Same Non Selected Topics List - [10%]. This last 10% is only used in case there are not enough relevant resources between the other lists. In the final aggregation, the content-based list provides 80% and the collaborative list provides 20%. These percentages are set in this way after serveral experiments but not to dwell too much, we don't include in the system evaluation section. The idea is to achieve an appropriate diversity avoiding excessive specialization in a particular area.

3.5. Feedback phase

In this phase the recommender system calculates and updates the ratings of the recommended resources for the users. The system communicates the recommendations by email and registers the feedback of each user. These ratings provided by the user are taken to improve the quality of the next recommendations for himself and the collaborative approach needs the users' ratings to generate recommendations.

3.6. System evaluation

Due to singularity of the system, we can not compare our proposal with other approaches using a standard data set, because no standard data set has this information about the preferences of users and the adaptation of the existent data sets is not possible. We are not proposing a new recommendation strategy but an approach which combines hybrid recommendation with quality measures as bibliometrics that were widely used in many areas and whose efficiency is proved [26, 27]. So, we performed online experiments where a group of researchers receive recommendations and report their opinions about the recommendation supplied by the system.

3.6.1. Data sets

We made the tests using the collaboration of 100 researchers of the research groups: *Soft Computing and Intelligent Information Systems*[3] from University of Granada, *Delft Multimedia Information Retrieval Lab*[4] from TU Delft, *Intelligent System and Data Mining*[5] and *Intelligent System Based on the Fuzzy Decision Analysis*[6] from University of Jaen and *Department of Civil Engineering*[7] from university of Granada.

By other hand the set of resources was composed by 585000 papers belonging to the year 2013 of 1731 journals related with the areas where the selected users are focused. These resources were downloaded from WOK using the tool described in [28]. Then they were progressively added monthly from WOS obtaining the recent papers. The system filtered these resources when the recommendation round was launched and recommended them to the suitable users. To find out whether the recommendations were appropriate we based our study on the feedback given by the users, removing unusual behaviors such inactive users or users that did not update their profiles. We registered the ratings provided by users about the recommended resources to compare them with the predictions generated by our system.

3.6.2. Experiments results

In our experiments we are not relying on experts who can say if the recommendations are relevant or not. Furthermore, we are implementing a temporal window to avoid overwhelming users with a excessive number of recommendations. Under these conditions, we have no means to identify which items are relevant but are not recommended due to the limitation in the number of recommendations, so

---

[3]http://sci2s.ugr.es/
[4]http://dmirlab.tudelft.nl/content/members-dmir-lab
[5]http://simidat.ujaen.es/
[6]http://sinbad2.ujaen.es/
[7]http://www.icivil.es/web2.0/

that, we cannot build a contingency table. Therefore, in our study we compute the Mean Absolute Error (MAE), which measures the mean absolute deviation between a estimated value and the real value assigned by the user. So, as a rating metric for the system, we estimate the average of the MAE for all the users resulting into a final MAE of **0.30312**. Moreover, we have computed the coverage of the proposed system [29] obatining a value of 100%, i.e., in a fully automated process, we are able to cover the whole number of users and items.

To compare, we have tested the system with other different configurations:

- No re-ranking (NRR) that is, the ten first items of the previous ordered list of items or users similar to the active user, either content-based and collaborative.
- Similarity measured based in abstract (ABS) instead of keywords, using tf-idf algorithm [30] to estimate the similarity between papers and users, either content-based and collaborative, without the use of re-ranking.
- Quality filter for the content-based based on the H-index (QH) of the authors of an article instead of the IF of the journal.
- Same quality filter in journals but without the application of the diversity filter (NDF).
- Actual Configuration (AC), that is, the structure explained above, re-ranking, quality filter based on the IF of the journal and on the H-index for an author and the use of the diversity filter.

Table 1 shows the MAE values for each configuration we used to prove that the chosen one is the optimal. The result obtained after the non-application of the re-ranking (NRR), is lightly worst than the actual configuration (AC), so we can observe certain improvement obtained by the application of the quality as index for the re-ranking process. We also noted that, the use of the analysis of the abstract based on the tf-idf algorithm produced poorer results than expected (ABS); probably the fact that abstract are all written in a similar manner might explain this poor performance. By other hand, the quality criteria of the H-index of authors (QH) of an article in the content-based approach proves that this criteria distorts the results; the reason of this behavior is that these good authors can be oriented to specific areas similar to user interest but with different approaches. Finally, the use of diversity filter (NDF) resulted to be an essential task to mitigate the own distortion introduced by the quality criteria, as the quality of some items could not be relevant if there is no relation with the actual profile.

Configuration	MAE
NRR	0.39
ABS	0.72
QH	0.44
NDF	0.49
AC	0.30

Table 1: MAE for all the different configuration.

4. Conclusions and future work

The personalization of information related to research papers deserves a special attention due to the ever growing information overload problem in the research world. Recommender systems have been successfully applied in both academic and commercial environments obtaining relevant information.

In this paper we have proposed a hybrid fuzzy linguistic recommender system about new research papers and based on the quality of the items and users. On the one hand, a content-based approach to obtain the latest papers inserted into the system, non rated by any user, and on the other hand, a collaborative approach to avoid disregarding potentially important papers for a user, rated in the last two rounds by experts on the field. The system measures both items' and users' quality for the re-ranking process. Beside the quality of items and the quality of the users, the system takes into account the users' profiles for the top-N recommendations.

We have performed online studies with the proposed system and the experimental results yielded satisfactory results.

As future work, we consider enhancing the paper keywords with the automatic extraction of terms from the article abstracts for the users' vectors to estimate similarity. One step in this direction is to use Wordnet [31] in the process for the nouns extraction. Moreover, we want to expand the use of this proposal as a tool for users not only to be up to date on the newest trends but also for researchers who wants to enter in a new area, or even to have the best personalized papers for a wider time window.

Acknowledgments

References

[1] A. Moro-Martín. Spanish changes are scientific suicide. *Nature*, 482(7385):277–277, 2012.

[2] I. Macho-Stadler, D. Perez-Castrillo, and R. Veugelers. Licensing of university inventions: The role of a technology transfer office. *International Journal of Industrial Organization*, 25:483–510, 2007.

[3] Web of science, information.

[4] Scopus.

[5] Sciencedirect.

[6] Google scholar.

[7] R. Burke. Hybrid web recommender systems. *P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): The Adaptive Web, LNCS*, 4321:377–408, 2007.

[8] C. Porcel, A. Tejeda-Lorente, M.A. Martínez, and E. Herrera-Viedma. A hybrid recommender system for the selective dissemination of research resources in a technology transfer office. *Information Sciences*, 184(1):1 – 19, 2012.

[9] A. Tejeda-Lorente, C. Porcel, E. Peis, R. Sanz, and E. Herrera-Viedma. A quality based recommender system to disseminate information in a university digital library. *Information Sciences*, 261(0):52 – 69, 2014.

[10] R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12:331–370, 2002.

[11] S. Alonso, I. J Pérez, F.J. Cabrerizo, and E. Herrera-Viedma. A linguistic consensus model for web 2.0 communities. *Applied Soft Computing*, 13(1):149–157, 2013.

[12] S. Alonso, F.J. Cabrerizo, E. Herrera-Viedma, and F. Herrera. H-index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, 3:273–289, 2009.

[13] U. Hanani, B. Shapira, and P. Shoval. Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11:203–259, 2001.

[14] Researchgate.

[15] C. Wang and D.M. Blei. Collaborative topic modeling for recommending scientific articles. In *KDD '11 Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 448–456, 2011.

[16] N. Manouselis and K. Verbert. Layered evaluation of multi-criteria collaborative filtering for scientific paper recommendation. *Procedia Computer Science*, 18:1189–1197, 2013.

[17] D. Nart and C. Tasso. A personalized concept-driven recommender system for scientific libraries. *Procedia Computer Science*, 38:84–91, 2014.

[18] L.A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning. Part I, Information Sciences 8 (1975) 199-249, Part II, Information Sciences 8 (1975) 301-357, Part III, Information Sciences 9 (1975) 43-80, 1975.

[19] F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 8(6):746–752, 2000.

[20] F. Herrera and L. Martínez. A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multi-expert decision-making. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 31(2):227–234, 2001.

[21] S. Massanet, J.V. Riera, J. Torrens, and E. Herrera-Viedma. A new linguistic computational model based on discrete fuzzy numbers for computing with words. *Information Sciences*, 258:277–290, 2014.

[22] V. Durieux and P.A. Gevenois. Bibliometric indicators: Quality measurements of scientific publication 1. *Radiology*, 255(2):342–351, 2010.

[23] E. Garfield. Citation indexes for science. a new dimension in documentation through association of ideas. *International journal of epidemiology*, 35(5):1123–1127, 2006.

[24] L. Egghe. Theory and practise of the g-index. *Scientometrics*, 69(1):131–152, 2006.

[25] M.A. Martínez, M. Herrera, J. López-Gijón, and E. Herrera-Viedma. H-classics: characterizing the concept of citation classics through h-index. *Scientometrics*, 98:1971–1983, 2014.

[26] RN Kostoff, R Tshiteya, KM Pfeil, JA Humenik, and G Karypis. Power source roadmaps using bibliometrics and database tomography. *Energy*, 30(5):709–730, 2005.

[27] Christopher J Williams, Michael O'Rourke, Sanford D Eigenbrode, Ian O'Loughlin, and Stephen J Crowley. Using bibliometrics to support the facilitation of cross-disciplinary communication. *Journal of the American Society for Information Science and Technology*, 64(9):1768–1779, 2013.

[28] S. Alonso, F.J. Cabrerizo, E. Herrera-Viedma, and F. Herrera. Wos query partitioner: A tool to retrieve very large numbers of items from the web of science using different source-based partitioning approaches. *Journal of the American Society for Information Science and Technology*, 61(8):1582–1597, 2010.

[29] Mouzhi Ge, Carla Delgado-Battenfeld, and Dietmar Jannach. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 257–260. ACM, 2010.

[30] S. Robertson. Understanding inverse document frequency: on theoretical arguments for idf. *Journal of documentation*, 60(5):503–520, 2004.

[31] C. Fellbaum. *WordNet*. Wiley Online Library, 1999.