A Knowledge Based Recommender System with Multigranular Hierarchical Linguistic Contexts

Luis Martínez*, Manuel J. Barranco*, Luis G. Pérez*, Macarena Espinilla* and Emilio J. Castellano*

Dpt. of Computer Science, University of Jaén,
Jaén (23071), Spain
**E-mail: martin, barranco, lgonzaga, mestevez@ujaen.es*
www.ujaen.es

Electronic shops provide an excellent choice to buy without leaving home. Nevertheless, people frequently have problems to find what they look for because of the wide range of items that e-shops offer. Recommender Systems are applications that help people in their searches in e-shops. They deal with information that people provide which is usually related to opinions, tastes and perceptions, and therefore, it is difficult to express them by means of precise numeric scales. Fuzzy linguistic approach provides a better way to express this kind of information. In this contribution we propose a Knowledge Based Recommender System that uses the fuzzy linguistic approach to handle the uncertainty of the human opinions and provides a multigranular context that allows users to utilize the term set that better fits with their degree of knowledge.

Keywords: e-commerce; recommender systems; fuzzy linguistic approach; 2-tuple linguistic representation model; multigranular linguistic information.

1. Introduction

In the last years Internet has turned into a powerful tool to search what we need. E-shops offer a large number of products and it is possible we can find everything we are looking for, visiting a few e-shops in Internet. Nevertheless, the searching is not easy because of the large amount of information that e-shops provides. Recommender systems emerged in order to support users in their purchase decisions in electronic shops.

There are different types of recommender systems[1,6,7] but all of them follow the same steps: (i) The system gathers information from users, experts, etc., related to their preferences about the products, (ii) the system deals with the available information about the target user in order to know her preferences and, finally (iii) the products are ranked according to the user's preferences and the recommendation is made. Depending on the technique used for ranking the products and the type of information gathered by the system, recommender systems are classified into several categories. One of them are the Knowledge based recommender systems. They are characterized because they do not need historical information about users. These systems make recommendations only with the knowledge the user gives about her necessities at the moment of her search and the information they have about the items that could be recommended. In a typical situation, the only information that the user provides is an example of her necessities. Starting with this example, the system builds a user profile and recommends the products that better fit with such profile.

Our proposal uses the Fuzzy Linguistic Approach[8] to model the information provided by the users since this information is related to opinions and preferences and so, it has a qualitative nature. Besides, we must take into account that our model deals with multigranular linguistic information because users may have different degrees of knowledge about the items and their attributes. In [5] we proposed a model for content-based recommender systems that deals with linguistic and multigranular information providing good results.

2. Linguistic hierarchies

In order to model a multigranular context in the Recommender System we shall use the 2-tuple fuzzy linguistic representation model[3] and the linguistic hierarchies[4]

A *linguistic hierarchy* is a set of levels, where each level is a linguistic term set with different granularity from the remaining of levels of the hierarchy. Each level of a linguistic hierarchy is denoted as $\mathbf{S}(\mathbf{t}, \mathbf{n}(\mathbf{t}))$, in short $S^{n(t)}$, where t indicates the level of the hierarchy and *n(t)* the granularity of the linguistic term set of the level t. We define a linguistic hierarchy, LH, as the union of all levels t: $LH = \bigcup_t S^{n(t)}$ being $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$

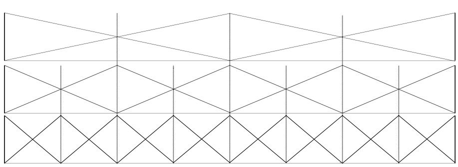

Fig. 1. : Linguistic Hierarchy of 3,5 and 9 labels.

Generally, we can say that the linguistic term set of level $t+1$ $S^{n(t+1)}$, is obtained from its predecessor, $S^{n(t)}$, as: $S(t, n(t)) \rightarrow S(t+1, 2 \cdot n(t) - 1)$

A graphical example of a linguistic hierarchy is showed in Fig. 1. This structure allows to deal with multigranular linguistic information without lost of information because there exists a biyective transformation function between levels.

Definition 1. *Let* $LH = \bigcup_t S(t, n(t))$ *be a linguistic hierarchy whose linguistic term sets are denoted as* $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$*. The transformation function between a linguistic label that belongs to level t and a label in level t'=t+a, with* $a \in Z$*, is defined as:*

$$TF_{t'}^{t} : S(t, n(t)) \longrightarrow S(t', n(t'))$$

$$TF_{t'}^{t}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t') - 1)}{n(t) - 1}) \tag{1}$$

where $(s_i^{n(t)}, \alpha^{n(t)})$ is a 2-tuple linguistic label. In [2] the function Δ is introduced which transforms a tuple into a number and the function Δ^{-1} is its inverse.

3. Knowledge Recommender System with Linguistic Hierarchies

The structure of our knowledge recommendation system is in Fig. 2. This structure shows the necessary elements and phases to obtain a recommendation. Initially the system stores the information about the items in a database $A = \{a_1, \ldots, a_m\}$ in which each item, a_i, is described by a set of features $\{c_1, \ldots, c_l\}$ and its values are kept in a utility vector, $F_{a_j} = \left\{v_1^j, \ldots, v_k^j, \ldots, v_l^j\right\}$ in which v_k^j is a linguistic value of the feature c_k for the item c_j. Each attribute, c_k, is assessed by using the linguistic term set $S^{n(t_k^e)}$ of the linguistic hiearchy LH used by the system.

In order to compute recommendations our system, first of all, builds a user profile, and then, it computes the recommendation using the profile and the item database. Following, we shall describe the system working in detail.

3.1. *Profiling phase*

The aim of this phase is to build a profile for the user from an example and its description. The user profile is obtained after three steps:

3.1.1. *Gathering of user preferred example*

Firstly, the user provides an item whose features will constitute the initial user profile. Let a_u be the item the user has chosen. Then, the *initial user*

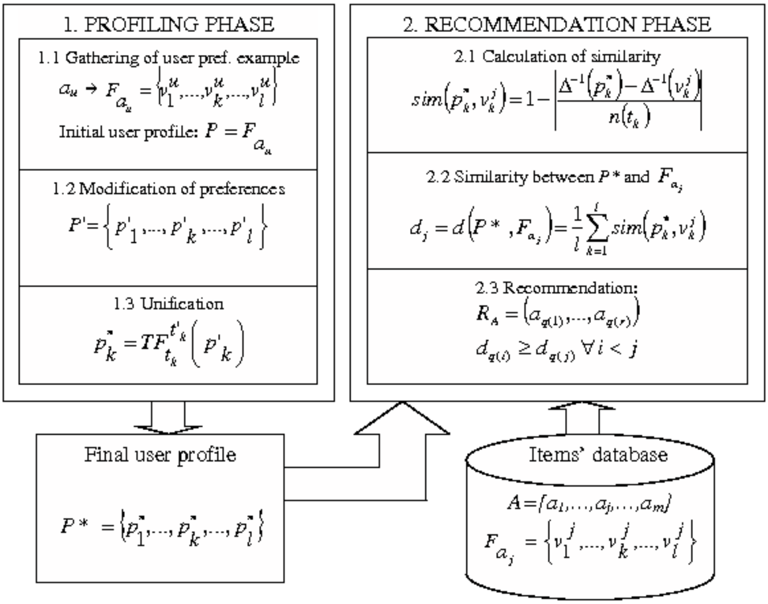

Fig. 2. Recommender System Model

profile $P = \{p_1, \ldots, p_k, \ldots, p_l\}$ is equal to $F_{a_u} = \{v_1^u, \ldots, v_k^u, \ldots, v_l^u\}$, in which v_k^u is a linguistic value that describes the feature c_k for the item a_u.

3.1.2. *Modification of preferences*

The chosen example represents the user's necessities, preferences, tastes, etc. Nevertheless, this example probably does not fit exactly her necessities. Therefore, the system allows her to modify some features of the initial user profile in order to do a better adjustment of her necessities. In this process, the system adapts itself to the user's knowledge providing a linguistic hierarchy $LH = \bigcup_t S(t, n(t))$. So, the user can choose the set that better fits with her degree of knowledge. If the user does not agree with the value for the feature c_k, then, she will choose a linguistic term set $S^{n(t'_k)}$ and will provide a value from this set for assessing this feature. This process can be repeated for other features. Finally, the system will have a *modified user profile* $P' = \{p'_1, \ldots, p'_k, \ldots, p'_l\}$ where $p'_k \in S^{n(t'_k)}$.

3.1.3. *Unification process*

To deal with multigranular linguistic information it is necessary to unify the information of the user profile and the items descriptions in the same linguistic term set in order to match one each other. To do so, all the linguistic information provided using different levels of the LH, $S^{n(t'_k)}$, must be transformed to the linguistic term set $S^{n(t^e_k)}$ by means of the transformation function $TF^{t'_k}_{t^e_k}$ (see eq. (1)).

Then, all the features modified by the user in the before step, are transformed with $TF^{t'_k}_{t^e_k}$. Let c_k be a feature modified by the user. The system will apply the $TF^{t'_k}_{t^e_k}$ function to the element p'_k to obtain the element p^*_k. So, the result of this process is the *final user profile*, $P^* = \{p^*_1, \ldots, p^*_k, \ldots, p^*_l\}$ in which p^*_k is a linguistic value belonging to $S^{n(t^e_k)}$, i.e., the same linguistic term set used for the c_k feature in the items database.

3.2. *Recommendation phase*

Once the information has been unified the system is ready to make recommendations evaluating the similarity between the user profile and items, selecting the most similar ones. This phase consists of two steps:

3.2.1. *Calculation of the similarity between the user profile and the items*

In this step, the system computes the similarity between the user profile and every item $a_j \in A$. To do so, we propose the use of the following function:

$$d_j = d(P^*, F_{a_j}) = \frac{1}{l}\sum_{k=1}^{l} sim(p^*_k, v^j_k) \tag{2}$$

This function computes this similarity as an aggregation of the similarities between the assessments that describe the user profile and those ones that describe the item a_j. The aggregation is accomplished by using an arithmetic mean. In order to measure the similarity between the features of the user profile and the items the following function is used:

$$sim(p^*_k, v^j_k) = 1 - \left| \frac{\Delta^{-1}(p^*_k) - \Delta^{-1}(v^j_k)}{n(t^e_k)} \right| \tag{3}$$

3.2.2. *Recommendation*

Finally, we will select the most similar items to the user profile according to the values d_j obtained in the before step.

Let $A = \{a_1, \ldots, a_m\}$ be the item set, and r the maximum number of items to be recommended. Then, the recommendation is given by the vector, R_A, where the first element is the top one recommended item.

$$\begin{aligned} &R_A = \left(a_{q(1)}, \ldots, a_{q(r)}\right) \text{ where} \\ &q : \{1, 2, \ldots, r\} \rightarrow \{1, 2, \ldots, m\} - \{e\} \\ &d_{q(i)} \geq d_{q(j)} \quad \forall i < j \end{aligned}$$

4. Conclusions

We have proposed a model for a knowledge based recommender system that deals with multigranular linguistic information. This context offers the users a greater flexibility to express their necessities since users can use the best linguistic term set according to their degree of knowledge among the term sets offered by the system. On the other hand, we must realise that because users can use different linguistic term sets of the ones used in the items database, and because each feature can be described using a different one, we are dealing with a multigranular linguistic context. In order to model this context and make easier the process of making recommendations, we have used linguistic hierarchies.

Further researches in the linguistic hierarchy are needed to improve its flexibility since the current representation only allows some granularity degrees for each level of the hierarchy.

5. Acknowledgements

This work is partially supported by the Research Projects TIN-2006-02121, JA031/06 and FEDER funds.

References

1. R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, 2002.
2. F. Herrera, E. Herrera-Viedma, and L. Martínez. A fusion approach for managing multi-granularity linguistic term sets in decision making. *Fuzzy Sets and Systems*, (114):43–58, 2000.
3. F. Herrera and L. Martínez. A 2-tuple fuzzy linguistic computational model. advantages of its linguistic description, accuracy and consistency. *Internation Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9:3348, 2001.
4. F. Herrera and L. Martínez. A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making. *IEEE Transactions on Systems, Man and Cybernetics*, 31(2):227–234, 2001.
5. L. Martínez, L.G. Pérez, and M. Barranco. A multi-granular linguistic content-based recommendation model. *International Journal of Intelligent Systems*, (In press), 2007.
6. Saverio Perugini, Marcos Andre Goncalves, and Edward A. Fox. Recommender system research: A connection-centric surveyrecommender system research: A connection-centric survey. *Journal of Intelligent Information Systems*, 2(23):107143, 2004.
7. P. Resnick and H.R. Varian. Recommender systems. *Association for Computing Machinery. Communications of the ACM.*, 40(3):56, Mar 1997.
8. L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning-i, ii, iii. *Information Sciences*, 8-9:199–249, 301–357, 43–80, 1975.