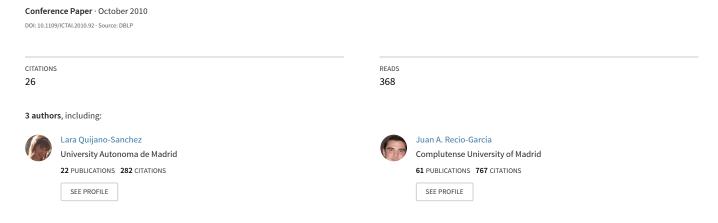
Personality and Social Trust in Group Recommendations



Some of the authors of this publication are also working on these related projects:

Project

Make it Personal: A Social Explanation System Applied to Group Recommendations View project

Personality and Social Trust in Group Recommendations

Lara Quijano-Sánchez, Juan A. Recio-García, Belén Díaz-Agudo Universidad Complutense de Madrid, Spain, email: {lara.quijano, jareciog,}@fdi.ucm.es, belend@sip.ucm.es

Abstract—In this paper we describe some new ideas to improve recommendations to groups of people. Our approach maximizes the global satisfaction for the group taking into account people personality and the social relationships among people in the group. We present some results with two cases of study based on the movie recommendation domain with heterogeneous groups. The first case study uses synthetically generated groups of people to test how the group composition affects the accuracy of the recommendation. Our second case study uses real users and groups where the topology of the groups is based on a social network. This second case of study with real users confirms the wide conclusions of the preliminary experiment with synthetic data, which allows us to conclude that it is possible to realize trustworthy experiments with synthetic data.

I. INTRODUCTION

Recommender systems represent a wide range of applications with a raising impact in the current web [1], [2]. Although most of the most popular recommender systems are focused on recommending items for individual users, the need of systems capable of performing recommendations for groups of people is getting more interest as there are many activities that are carried out in groups, like going to the cinema with friends, watching TV at home or listening music in the car. Our recent work [3] involves the improvement of current group recommendation techniques by introducing a novel factor: the personality of every individual. Intuitively, when a group of friends chooses a movie there are some members that are only happy if they impose their opinion, whereas other individuals do not care letting other people decide. Therefore, we have used a personality test to obtain the different roles that people play when interacting in a decision making process and studied how individual personalities influence the results and the satisfaction for the whole group.

Besides personality, this paper introduces the novelty of taking into account the social structure of the group to influence the recommendation process. Our approach reflects in a realistic way the relationships between groups of users connected in a social network. These relationships are measured through social factors, like the distance in the social network or the number of common friends, that are extracted from the social networks and that are used to compute the trust values between the members of the group.

In this paper we describe our theories about making recommendations for groups of people with different personalities and connected through social network structures. Our method proposes making recommendations to groups using existing techniques of collaborative filtering [4], taking into account the group personality composition and the social connections between the individuals of the group. We have tested our method in the movie recommendation domain using two test datasets. The first case study uses synthetically generated data to create simulated groups of people to test how the group composition affects the accuracy of the recommendation. Our second case study uses real users and groups where the topology of the groups is based on a social network. This second case of study with real users confirms the wide conclusions of the preliminary experiment with synthetic data, which allows us to conclude that it is possible to realize trustworthy experiments with synthetic data.

The paper runs as follows: Section II describes existing techniques to obtain recommendations for groups and gives some details of our personality aware recommendation process and our proposal for including social network topologies in the decision process. Section III explains our case of study: Movie Recommendation. Section IV shows the results of our experiment. Section V concludes the paper and presents the main lines of future work.

II. RECOMMENDATION TO GROUPS

Recommender systems have traditionally recommended items to individual users, but there has recently been a body of work about recommenders that extend their recommendations to groups of users [2]. When moving from individuals to groups many new issues arises. For example, acquiring the preferences of the group, helping the group to decide the more convenient option, or explaining the recommendation to the group. Depending on the size and homogeneity of the group the recommender system has to choose the option that satisfies the biggest number of people taking into account the individual user preferences. As stated in [2] the main approaches to generate a preference aggregation based on the individual user preferences are (a) merging the recommendations made for individuals, (b) aggregation of ratings for individuals and (c) constructing a group preference model.

The most common employed approaches in group recommenders are (b) and (c). They are used in many different fields like selecting the background music of a fitness center [5] recommendation of video clip sequences [6], movies [7] among others. These strategies usually try to maximize the "average satisfaction" of the group. The work described in [7] recommends movies for groups based on the inferred



ratings by MovieLens and using the "least misery" strategy to generate the preference aggregation. This strategy supposes that a small group of people will be as happy as its least happy member. The work in [8] criticizes the aggregation strategies like the one employed in PolyLens because they claim that these strategies combine the ratings always in the same way without considering how the members in the group interact with each other.

A. Personality Aware Recommendation to Groups

Most of the previous works in group recommendation consider the preferences of every member of the group with the same degree of importance and try to satisfy the preferences of every group member. However, groups of people can have very different characteristics like size and can be made of people with similar or antagonistic personal preferences. It is a fact that when we face a situation in which the concerns of people appear to be incompatible a conflict situation arises. Our approach determines that the general satisfaction of the group is not always the aggregation of the satisfaction of its members as different people have different expectations and behaviour in conflict situations that should be taken into account. In [3] we have presented a method for recommendation to groups where we distinguish between different types of individual personalities in a group. Our research characterizes people using the Thomas-Kilmann Conflict Mode Instrument (TKI) [9] that describes a person behavior in conflict situations along two basic dimensions: assertiveness and cooperativeness. These two dimensions of behavior can be used to define five personality modes of dealing with conflicts: competing, collaborating, avoiding, accommodating and compromising. We had our users fill up the TKI test to determine their personality, once each user u_i completes it, we calculate a value that represents how selfish or cooperative they are, we call this value the (CMW_i) (Conflict Mode Weight,), it represents the predominant behaviour in that particular user according to his TKI evaluation. We note that people with strong personality have a high CMW value, while a CMW value represents an easy going personality. The CMW function returns values in a range of [0,1], being 0 the reflection of a very cooperative person and 1 the reflection of a very selfish one. We studied how the group personality composition influences the recommendation accuracy for the group, and how it is improved for certain types of groups compared with different simple group recommendation algorithms.

Our approach creates a group recommendation by mixing individual recommendations. The individual recommender is a collaborative system based on the ratings of other users. Every user rates movies that (s)he has seen before, and the recommendation process consists comparing her/his ratings with the rating of other users to obtain her/his most similar users (the ones that rated the same movies in a similar way). Finally, the ratings of the most similar users are used to infer unknown ratings in movies and to create a new recommendation.

Our proposal uses the type of personality to weight the

influence of his ratings during the recommendation process. If we consider $ir_{i,m}$ the rating of a given user u_i to a certain product m we can see in the following algorithm to compute personality based recommendations (pbr(u,m)), this is, how the personality can be taken into account with the aggregation function of average satisfaction.

$$pbr(u_i, m) = \frac{1}{|G|} \sum_{i \in G \land i \neq i} (ir_{i,m} + pd(u_i, u_j))$$
 (1)

where
$$pd(u_i, u_j) = (CMW_j - CMW_i) \cdot \alpha$$

represents the personality difference, α value has been experimentally selected and it is employed to modify the impact of the personality differences on the modified rating, |G| represents the number of components of the group and CMW_j is the conflict mode weight of the user j.

What we observe is the difference between the personality pairs in the individuals of the group. It is based in a modification on the average satisfaction aggregation method. This strategy reflects that assertive characters will have more influence in the average satisfaction than the cooperative characters. This factor is computed as the distance of CMW values in the personality difference function $pd(u_i, u_j)$. We can find the algorithms applied for other merging functions like least misery in [3].

In this paper we use this personality based recommendation method in a first experiment using simulated users and groups with different features to see how the personality composition of the member of the groups affects the recommendation results. Then, in a second experiment we compare these results with real users and groups of people to test the accuracy of our simulated dataset. Before presenting the case study we describe how our approach consider the group structure into the recommendation process.

B. Social Trust in Recommendation to Groups

Current research has pointed out that people tend to rely more on recommendations from people they trust (friends) than on recommendations based on anonymous ratings [10]. This factor is even more important when we are performing a group recommendation where users have to decide an item for the whole group. This kind of recommendations usually follows an argumentation process, where each user defends his preferences and rebuts other's opinions. Here, the trust between users is the major factor when users must change their mind to reach a common decision.

A promising approach is to collect the trust knowledge from existing social networks. Social networks have been one of the most important topics in the last few years, with nets like Facebook, Twitter and MySpace, among others. The use of social networks and trust when building a recommender system is not new [11], [12]. One contribution of our approach is the use of generic social network topologies to reflect the real interactions of the users in the group. Our working hypothesis is that instead using the topology "all connected",

which is the default topology used in typical group recommendation approaches, with the social network topology we give a more realistic structure and organization of the group, which is closer to how the argumentations would take place in a real group when they argue on which movie to watch. Our main goal is to improve the recommendations by taking into account both, the network topology and the group personality composition. Moreover, our goal is to identify which are the most important factors of the social networks that must be taken into account when computing the trust between users. Examples of these factors are: the number of shared messages, common pictures, direct friends, etc. To perform this task we review several existing works [13], [14] and select the most relevant and feasible factors. These factors are detailed in Section III-B and their impact is reported in Section IV.

III. CASE OF STUDY: MOVIE RECOMMENDATION

We have evaluated our method in the movie recommendation domain. We have initially realized our experiment with synthetic data because we wanted to explore extreme cases that could appear in conflict situations. We wanted to have control of the distribution of the data, which didn't happen if we used real data. This synthetic data let us explore every group composition and personality distribution within the group. It also lets us reproduce the behaviour of large groups that are very difficult to organize in experiments with real users. Next we performed a second experiment using real data in order to verify the results that we obtained first. This experiment with real users confirms the wide conclusions of the preliminary experiment with synthetic data.

We built the recommender systems using the jCOLIBRI framework [15]. jCOLIBRI is currently a reference platform in the Case Based Reasoning (CBR) community and includes an extension to build recommender systems.

A. Experimental set-up with synthetic people

Our experiment runs as follows: (1) we generated randomly many groups of users with different personality profiles and social topology; (2) we developed an individual recommender following the collaborative approach; (3) we created 3 different group recommender systems that use the individual system: a standard recommender that only aggregates preferences; a group recommender using only the personality profile, as explained in section II-A; and a final one that takes into account the personality and social topology, as reflected in equation 2. Finally, (4) we have compared the results obtained with different recommenders and different synthetic group configurations, because we wanted to study if these configurations affected the final recommendation.

These 4 group recommenders use the formula shown in equation 2. Depending on the recommender we set up the value of some parameters to 0. The baseline of our experiment is an standard recommender without personality or social factors (referred as *Base* in the results). Next, the group recommender that only takes into account the personality always uses a trust function that always returns 0 (we refer

to this recommender as *Personality*). Finally, our complete *social group recommendation method* (referred as *Personality & Trust*) uses the complete equation:

$$sgr(u_i, m) = \frac{1}{|G|} \sum_{j \in G \land j \neq i} (ir_{i,m} + pd(u_i, u_j) + trust(u_i, u_j))$$
(2)

In order to simulate the social network inside each group we have randomly generated friendship links between users. We have defined a *trust function* that analyzes these links to compute the trust among users depending on their distance inside the social network. As we present in Section III-B, the second experiment computes the trust value using several factors obtained from the real social network users belong to. However, we have chosen a simple approach in this initial experiment to infer trust along the network (as users are synthetically generated):

$$trust(a,b) = \begin{cases} 1 & \text{if } distance(a,b) == 1\\ 0.5 & \text{if } distance(a,b) == 2\\ 0 & \text{a.o.c.} \end{cases}$$

To generate the group of users we have used a set of 100 people. Every person was assigned to a random value CMW to reflect their personality -in works with real data values, this is the value that was computed using the results of the TKI personality test [9]. This value is employed in function CMW_i that will return the personality weight of every individual. We basically define five different types of personality according to this range: very selfish, selfish, tolerant, cooperative and very cooperative. For example if we consider a selfish person his CMW value must be contained in a range of [0.8,1.0]. When we realized the TKI test to real users, like in [3], there were some of these ranges that were unexplored because people usually don't have such extreme personalities, this is why decided to use the synthetic data. To be able to study the effects of the different types of personalities we generated 20 users for each type of personality. We grouped users in sets of 3, 5, 10, 15, 20 and 40 people. For each group size we selected the components of the group so that the personality distributions presented all the possible combinations: groups of very selfish, selfish, tolerant, cooperative, very cooperative, very selfish & very cooperative, very selfish & tolerant, ... and so on until we reach 13 possible combinations. In the end we had 76 groups (13 different distributions for each size, except for the 40 people group where we only had 11 combinations due to the resemblance of personalities in such big groups).

The next element required by our experiment is the evaluation function to measure the accuracy the recommendation. We have to figure out which movies would each of our users have chosen individually from a movie listing of a cinema, and afterwards determine which of that movies the group as a whole would have finally decided to watch. Our evaluation function is based on the content of movies. As we have explained the recommender systems we are evaluating are rating-based (collaborative). It implies that the content or features description of the recommended item is not taken into account

during the recommendation process. Instead collaborative filtering there is an alternative recommendation technique called content based recommendation [1] that compares the features of new items to the items already selected by the user, and sorts them according to their similarity following the typical retrieval algorithms in CBR. We use this kind of content-based technique in the evaluation function to figure out which movies each user likes.

We selected a list of 50 heterogeneous movies from the MovieLens data set [16] and we rated them (with a range of 0.0 to 5.0) for each user according to a random profile we assigned. These profiles were constructed according to typical preferences in movies of real life people according to their age, sex and preferences. For example, the ratings that a user with a childish profile would give were very high ratings to animation, children or musical movies and very low ratings to drama, horror, documental, etc.

Afterwards, for each user we obtained which movies would be rated with 4.5 or more and define the set of favourites movies for each user. With this information we used a content-based similarity test to organize the listing of the cinema in order of preference. We chose the top 3 and marked them as the individual favourites if. Secondly we needed to obtain the decision of the group. Now that we knew which movies would the individual users argue for, we reproduced a real life situation were everyone discussed their preferences, taking into account the personalities and the friendship between them and then we finally obtained the real favourite movies for the group rgf. We use this information to evaluate the accuracy of our recommender by comparing how many of the first three recommended movies -the proposed group favourites gf-belong to the rgf set of that group.

B. Experimental Setup with real users

Although the first experiment with synthetic data let us explore extreme group configurations, we required a second experiment with real users to validate the obtained results (see Section IV). In order to perform our second experiment in the movie recommendation domain with real users, we create two events in two different social networks, Facebook¹ and Tuenti². In these events we ask some of our friends in the social network for completing three questionnaires³. The first questionnaire serves to obtain the personality profile by asking the 30 questions from the TKI personality test [9]. Second questionnaire gets the individual preferences of the user about cinema. Users evaluate 50 heterogeneous movies from the MovieLens data set [16](rating them with a range of 0.0 to 5.0). This way, we can compute the collaborative filtering algorithm to compute the individual predictions $ir_{i,m}$. Finally, third test asks users to choose their 3 favorite movies from a list of 15 recent movies (of the 2009 year), that represents a movie listing from a cinema. This movie listing was chosen heterogeneously from the MovieLens database. These movies are the ones they would actually like to watch or had enjoyed best. The goal of this test is to measure the accuracy of the individual recommender. The answers to these questionnaires are analyzed to define the user profile of each participant. 58 real users have participated in our experiment. To measure the accuracy of the group recommendation we create groups with our participants and we ask them to simulate that they are going to the cinema together. We provide them the 15 movies that represent our movie listing and we ask them to choose which 3 movies they actually would watch together. We manage to gather 15 groups of 9, 5 or 3 members. The three movies that each group chooses are stored as the real group favourites set rgf. This way, to evaluate the accuracy of our recommender we can compare the set proposed by the recommender -the gf set- with the real preferences rgf. We measure the number of movies in gf that are also in rgf. Once we have the tests, we need the personality and trust factors. To obtain the personality value, we calculate the CMW_i (Conflict Mode Weight) from the results of the first test given to the users. To compute the trust factor we reviewed several existing related works [13], [14] to decide which are the specific factors that must be taken into account. We have selected 10 factors that are combined to get a final trust value. Moreover, we have evaluated (see Section IV) which factors have the highest impact in the recommendation process. The specific trust factors obtained from the social network are:

- f_1 : Distance in the social network.
- f_2 : Number of common friends.
- f₃: Intensity of the relationship: how often they write each other on their walls.
- f₄: Intimacy of the relationship: We classify relationships by finding keywords that represent different intimacy levels.
- f_5 : Duration: how long they know each other.
- f₆: Reciprocal services: number of posted videos/songs/webs, shared games/ applications.
- f₇: Structural variable: common interests described in the users profile like movies, books, or general interests.
- f_8 : Social distance: how many of the following properties are shared: political, educational, religious and demographical information.
- f₉: Status: Value depending on the kind of status: couple, family, best friends..etc
- f₁₀: Pictures: Percentage of of pictures where they appear together.

The final trust value $trust(u_i, u_j)$ is a weighted average of the previously described factors:

$$trust(i,j) = t_{ij} = \sum_{k=1}^{10} \alpha_k f_k(u_i, u_j)$$
 (3)

These weights have been experimentally obtained using a genetic algorithm (GA). Our GA manages a population of vectors of weights (α_i) . These vectors can be combined and mutated. The fitness function to maximize is the group recommendation accuracy.

¹http://www.facebook.com

²http://www.tuenti.com

³Questionnaires are accessible at http://www.lara.warhalla.com/(spanish)

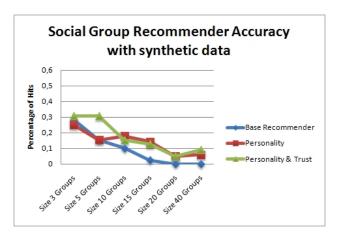


Fig. 1. Results for synthetically generated users.

Again, we build three different recommenders: (1) Base is a standard group recommender using the standard average satisfaction aggregating function; (2) Personality Aware Recommender only uses the personality data; and (3) Personality & Trust recommender takes into account both the social trust and the personality of each member. Next section details the results of our experiment.

IV. EVALUATION RESULTS

As the individual recommender has been only used as a base line for the group recommender we do not discuss its performance. However, let's notice that by improving this system the whole system will improve because the whole group recommender system is based on the individual preferences of each user. Another important factor in the recommender's performance is the configuration of the set of 15 movies that conform the movie listing of the cinema and the similarity of these movies to the set of 50 movies that the users had rated. These data was randomly chosen, so the performance of the recommender may change depending on the values.

The figures that analyze the different strategies of aggregation show three different lines that represent the results of the recommender when considering their friendship relations and their personality—*Personality & Trust* item—, just their personality—*Personality* item—, and just the simple base group recommender system *BaseRecommender*.

Figure 1 summarizes the results for synthetically generated users. This figure shows better results when combining personality and social trust in the group recommendation process. Therefore, we can conclude that recommender systems for groups could be improved when using a social factors.

Regarding the size of the groups we can clearly conclude that our recommender algorithm obtains better results for small groups than for big groups. It reflects the real world because with more people there are more different opinions and it is more difficult to arrive to a consensual solution.

On the other hand, when comparing how the distribution of the personalities in the group affected the percentage of hits, in Figure 3 we can see that the percentage was higher around

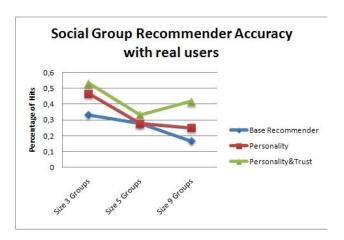


Fig. 2. Results for real users.

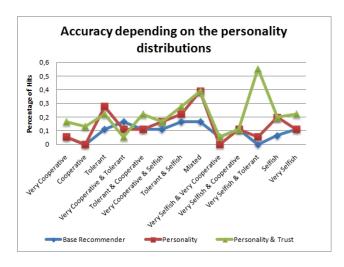


Fig. 3. Results depending on personality distribution.

the center of the graphic, which matches with the groups that have more varied personalities. Therefore heterogeneously in the personality of the group implies a better performance in the recommender. We can also observe that the right side of the graphic, representing groups of people with at least one leader role, has better percentage than the groups situated on the left representing groups of very cooperative people, were no one will impose their opinion.

Regarding the experiment with real users, Figure 2 shows the performance of this experiment (when considering personality and social factors, just personality and the simple base group recommender). Figure 2 summarizes the average results taking into account all the groups. We conclude that our *social group recommendation* method, explained in section II-A, obtains the best results. This figure confirms our theories and shows that we improve the recommendations when taking into account the social trust and the personality of each individual. It also reflects a similar behaviour when comparing these results with the synthetically generated data of Figure 1. In both Figures the *personality based recommendation* works better

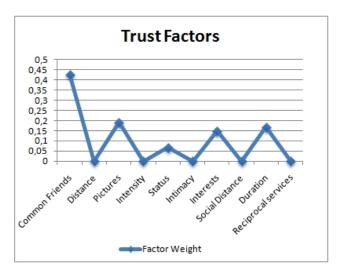


Fig. 4. Importance of the trust factors

than the standard recommender but worse than the *personality plus trust recommendation*. This second experiment only takes into account small and medium groups due to the evident impossibility of reproducing the experiment with large groups that are not real in recommendation processes.

Regarding the importance of each of the factors that conform the trust value, we can see in Figure 4 how they are taken into account in order to maximize the performance of our recommender. These weights (α_i) were obtained using the Genetic Algorithm, explained in section III-B. We can see that the most relevant is the number of friends in common, followed by the pictures, the duration, the common interests and the status. We think this experimental result is very relevant and should be taken into account when implementing real applications. Besides it relates with the used factor of friendship used in the experiment with synthetic data.

Finally, we have also studied if the accuracy of the recommendation was linked to the distribution of the personalities trying to compare it with the previous results with synthetic data. To do so, we calculated the standard deviation of the personality values for each group. After running several statistical studies, we concluded that there is not correlation at all between the distribution of the personalities of the group and the accuracy of our algorithms. This is due to the small deviation in real group of users, this is: normal groups are usually conformed by mixed personalities.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have proposed and experimented with a novel method of making recommendations for groups taking into account the group personality composition and the social structure of the group. We have shown that personality profiles and the social relationships between the users improve the accuracy of the recommendation for a group of people.

We have tested our method in the movie recommendation domain using two test datasets. The first experiment uses synthetically generated data to create simulated groups of people to test how the group composition affects the accuracy of the recommendation. We have also performed another experiment with real data, where we created two events in two different social networks, Facebook and Tuenti. In both experiments we have used groups of different size and personal preferences, where we have proved that by introducing the trust factor and the personality awareness we do improve the results of the recommendations. Our working hypothesis is that the personality and the social organization of the structure of the group will affect and improve the result of the recommendation, mainly because with the social network topology we give a more realistic structure and organization of the group. Our experiments confirm our hypothesis. One main conclusion of this paper is that it is possible to realize trustworthy experiments with our synthetically generated data as the second case study with real users confirms the wide conclusions of the preliminary experiment with synthetic data. As future work we are integrating our algorithms in the jCOLIBRI framework and we are creating a system with memory of the previous recommendations as a necessary step when providing a whole set of recommendations.

REFERENCES

- Bridge, D., Göker, M.H., McGinty, L., Smyth, B.: Case-based recommender systems. Knowledge Engineering Review 20 (2006) 315–320
- [2] Jameson, A., Smyth, B.: Recommendation to groups. The Adaptive Web LNCS 4321 (2007) 596–627
- [3] Recio-García, J.A., Jimenez-Diaz, G., Sánchez-Ruiz, A.A., Díaz-Agudo, B.: Personality aware recommendations to groups. In: Procs. of the 2009 ACM Conference on Recommender Systems, ACM (2009) 325–328
- [4] Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. Communications ACM 35 (1992) 61–70
- [5] McCarthy, J.F., Anagnost, T.D.: Musicfx: An arbiter of group preferences for computer aupported collaborative workouts. In: CSCW. (1998) 363–372
- [6] Masthoff, J., Gatt, A.: In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. User Modeling and User-Adapted Interaction 16 (2006) 281–319
- [7] O'Connor, M., Cosley, D., Konstan, J.A., Riedl, J.: Polylens: A recommender system for groups of user. In: ECSCW. (2001) 199–218
- [8] Chen, Y.L., Cheng, L.C., Chuang, C.N.: A group recommendation system with consideration of interactions among group members. Expert Syst. Appl. 34 (2008) 2082–2090
- [9] Thomas, K., Kilmann, R.: Thomas-Kilmann Conflict Mode Instrument. Tuxedo, N.Y. (1974)
- [10] Sinha, R.R., Swearingen, K.: Comparing recommendations made by online systems and friends. In: DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries. (2001)
- [11] Golbeck, J.A.: Computing and applying trust in web-based social networks. PhD thesis, College Park, MD, USA (2005)
- [12] Avesani, P., Massa, P., Tiella, R.: A trust-enhanced recommender system application: Moleskiing. In: SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, NY, USA, ACM (2005) 1589–1593
- [13] Gilbert, E., Karahalios, K.: Predicting tie strength with social media. In: CHI '09: Procs of the 27th Int. Conf. on Human factors in computing systems, New York, NY, USA, ACM (2009) 211–220
- [14] Golbeck, J.: Combining provenance with trust in social networks for semantic web content filtering. In: Provenance and Annotation of Data, Int. Provenance and Annotation Workshop, IPAW 2006. Volume 4145 of LNCS., Springer (2006) 101–108
- [15] Recio-García, J.A., Díaz-Agudo, B., González-Calero, P.A.: Prototyping Recommender Systems in jCOLIBRI. In: Proceedings of the 2008 ACM conference on Recommender Systems, NY, USA, ACM (2008) 243–250
- [16] Bobadilla, J., Serradilla, F., Hernando, A.: Collaborative filtering adapted to recommender systems of e-learning. Knowl.-Based Syst. 22 (2009) 261–265