

Improving Social Recommendations by applying a Personalized Item Clustering policy

Georgios Alexandridis
School of Electrical and
Computer Engineering
National Technical University
of Athens
Zografou, 157 80, Greece
gealexandri@islab.ntua.gr

Georgios Siolas
School of Electrical and
Computer Engineering
National Technical University
of Athens
Zografou, 157 80, Greece
gsiolas@islab.ntua.gr

Andreas Stafylopatis
School of Electrical and
Computer Engineering
National Technical University
of Athens
Zografou, 157 80, Greece
andreas@cs.ntua.gr

ABSTRACT

In online Recommender Systems, people tend to consume and rate items that are not necessarily similar to one another. This phenomenon is a direct consequence of the fact that human taste is influenced by many factors that cannot be captured by pure Content-based or Collaborative Filtering approaches. For this reason, a desirable property of Recommender Systems would be to identify correlations between seemingly different items that might be of interest to a particular user. This course of action is expected to improve the novelty and the diversity of the recommendations and therefore increase user satisfaction.

In this paper, we address this problem by proposing a socially-aware personalized item clustering recommendation algorithm. We are trying to locate patterns between the items that a user has evaluated by grouping them into different clusters according to the rating behavior of the members of his *Personal Network*, which includes the individuals in his direct social network and those other persons that the user exhibits a similar item evaluation behavior. Once the clustering phase has been completed, we use each cluster's members as seed items in order to construct an item consumption network. Then, by performing a random walk on the aforementioned network, we are able to produce recommendations that are accurate and at the same time novel and diverse. Preliminary results reveal the potential of this idea.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Clustering;
G.3 [Probability and Statistics]: Stochastic processes;
F.1.2 [Modes of Computation]: Probabilistic computation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

RSWeb 13 Hong Kong, China

Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...\$15.00.

Keywords

social recommendation, personal network, spectral clustering, item consumption network, random walk

1. INTRODUCTION

Research over *Recommended Systems* (RS) has grown almost exponentially over the past years. Initial approaches tried to model the user to item interaction evident in every RS in a plethora of ways; collaborative filtering algorithms exploited the said relationship in order to infer similarities and dissimilarities in taste between users that are not otherwise related to one another (*implicit* user to user interaction). On the other hand, content-based approaches tried to estimate item relevance by accumulating all available evaluations for each item and then compare them in some metric space. Lastly, hybrid systems combined both approaches, along with other possible sources of information (i.e. in the form of metadata) in an effort to produce better and more meaningful recommendations.

It is an indisputable fact that the recommendation process has an inherent social dimension as well. Apart from the fact that opinion and taste are formulated by a person's interaction with his environment, it is also a common everyday practice to turn to family, friends and acquaintances when we want to make a decision or a purchase. It could be further elaborated that we also tend to affiliate and establish bonds with people that we share the same interests with.

The emergence of Web 2.0 technologies and the advent of *Online Social Networks* (OSNs) introduced the social relationships into the digital era and inevitably in the recommendation process, brining about the so-called *Social Recommender Systems* (SRS). The user to item interaction of traditional RS has been extended in order to include *explicit* user to user connections. Those ties are extremely useful for the recommendation process as one of the most fundamental characteristics of social networks is *Homophily* [14]. This term means that the members of a social network tend to be more similar to those individuals they are connected with than other persons with whom they don't share direct relationships. Or, more strictly speaking, the acquaintances of any given member of an OSN do not constitute random samples drawn from the whole of the underlying population. This observation has also been experimentally confirmed by the work of Singla et al [13].

Homophily is influenced by many factors such as age, gender, educational level, ethnicity etc. In our work, we plan

to utilize this phenomenon in a novel approach by trying to examine the extent to which an individual's interests are influenced by the people he knows. We start by constructing a personal network for each user; one that includes the persons he has some interactions with, both direct and indirect. This personal network is the basis for the item clustering algorithm that places the items that this user has already consumed in different clusters, according to the consumption behavior of her peers. Once the clustering phase has been completed, the items of each cluster are expected to share some common characteristics. Therefore, using them as seed items and by performing a random walk on a specially constructed item consumption network for this purpose, we will be able to get recommendations that will be accurate, novel and diverse.

2. RELATED WORK

The application of clustering methodologies into the traditional RS domain is not new [1]. Researches have been using them in order to uncover the implicit relationships between users (or items) based on the available evaluations. In most cases, the purpose of clustering is to aide or to substitute user and item neighborhood formation.

The fusion of explicit user ties, in the form of social information, into the clustering process is a relatively novel research field. Most approaches use them as an alternative information source, upon which they apply existing clustering algorithms. For example, in [4] the authors propose a methodology to compute the inferred trust value between any pair of users and then use those values in order to perform a correlation clustering of the user space. They introduce their method into traditional memory-based algorithms (CF and trust-based) and witness a relative improvement in the recommendation accuracy.

In [5], the authors apply the *affinity propagation* clustering algorithm in order to dynamically identify the user clusters. The distance measure they employ is the Jaccard coefficient; the number of common neighbors in the trust network between a pair of users. Their experiments have shown that their system outperforms other traditional clustering techniques (such as *k-means*) in terms of the accuracy of the recommendations as the number of user clusters grows.

In [12], the authors perform a hierarchical clustering of the user's social relationships in order to form user neighborhoods. They use *modularity*, a graph-based criterion to stop the cluster formation. In short, modularity counts the fraction of the in-cluster edges against the number of edges pointing to other clusters. In the next step, those neighborhoods are fed into user-based and item based CF algorithms that produce the recommendations. They have tested their approach into two different datasets and have witnessed improvement both in the accuracy and the usefulness of the recommendations as measured by metrics such as precision and recall.

Our approach differs from those mentioned above in three key aspects; firstly, the clustering level is local (personalized) instead of system-wide, rendering our algorithm more scalable in terms of the dimensionality of the data involved in the computations. Secondly, while most approaches are geared towards grouping similar users, our approach is focused on locating similar items through the identification of common access patterns in the personal network. Finally, the social network information is used in the filtering phase

rather than the clustering algorithm itself.

3. THE PERSONAL NETWORK

So far, a person's taste has been discussed in relation to other people. However, taste itself is also a multidimensional concept; after all most people are not confined to having a single interest only. For example, if someone has an interest in gardening and in playing music, she is expected to evaluate gardening and music books. And since not all people who are interested in gardening are also interested in music, his rating behavior might seem bizarre from the RS point of view; indeed, even in large scale RS, few people would be expected to rate items of both the aforementioned categories.

We may further elaborate on this example by considering that this user's social network would consist of roughly two categories of people; those she's met via her hobbies and other, more general acquaintances. It is therefore quite likely that the target user and some of her gardening peers would have evaluated some gardening items in common. Our intuition is to try to locate such relationships in the RS and it is for this reason that we are performing a *personalized clustering* of the consumed items. The personalized clustering in this sense means that not every item evaluation in the RS is taken into account. Instead, they are filtered by those users that are explicitly (through the social network) or implicitly (through similar consumption behavior) related to the target user.

Our proposed algorithm works in two stages; initially, the ratings of a specific target user above a certain relevance threshold are retrieved. Then, these items are put into one or more different clusters according to some well-defined criteria. Following, the items of each cluster are used as a seed in order to construct an item consumption network. This structure is traversed in a random walk fashion for the recommendation of new items to the user.

3.1 Relationship Types

The explicit user to user ties that are evident in OSN in general and in SRS in particular lead to a variety of connections among them. It is therefore necessary to define the relationship types that would be considered valid for the creation of their personal network. The discipline of social network analysis examines simple and advanced structures present in social networks [15]. However, the nature of our research places the emphasis on *Dyadic Relationships*. They constitute the most common relationship type in OSN and directly involve two persons or actors. Those links may either be symmetric (or bidirectional) as in the case of *Friendship* or non-symmetric (directional). The most common directional link in SRS is *Trust*, where a user expresses his opinion on another user's behavior. These remarks are usually private to the user who has issued them and signify how useful she has found the interaction with the other user. Trust statements are either binary (i.e. trust/distrust) or assume a broader set of values (usually in the $[0, 1]$ interval). The SRS that incorporate trust statements in the recommendation process are also known as *Trust-aware Recommender Systems*.

3.2 Network Formation

As it has been stated before, we are interested in providing a personalized filtering of the available items in the

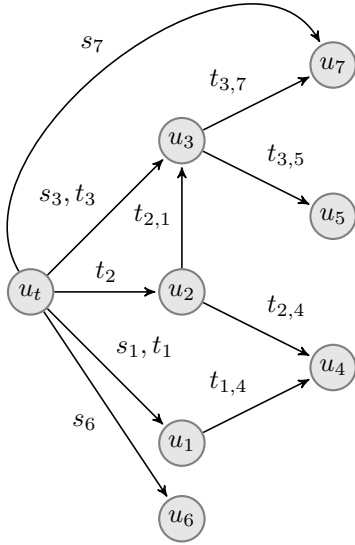


Figure 1: Personal Network

RS for each specific user. The basis of our approach is the *Personal Network* (PN) of each user which is constructed by two different but not necessarily distinct pools of users; those that are part of the target user’s social network and those that are similar (defined by some notion of similarity) to him. Members of the personal network of a specific user may be further categorized in the following groups

- Users in the direct social network of the target user that bear a similarity to her
- Other similar users
- Other users in the social network (Friend-of-a-Friend scheme) that might be similar to him
- Other users in the social network

Most widely-adopted indices of similarity in RS are the Pearson correlation coefficient, the cosine similarity and the Manhattan similarity. The first two measure similarity quite well when there is a large overlap between the items rated by different users, while the latter is more suitable in those cases where the overlap in ratings between different users is small.

Generally, the social and the similarity graphs could be accessed in a number of ways and indeed there is a wealth of methodologies in the SRS literature. As an initial approach, in order to estimate user proximity, we opted for weighted path counting criteria that are dependent on the structure of the PN. For example, in Figure 1, u_3 is considered to be the most “close” user to the target user u_t since he belongs both to his direct social network (edge weight t_3), the direct social network of u_t ’s friend, u_2 (edge weight $t_{2,1}$) and finally to u_t ’s similarity network (edge weight s_3). Following, users u_4 and u_7 may be reached by two simple paths of lengths one and two that originate from u_t . In order to estimate who is most proximate to him, we compute a total value of each path by accumulating the respective weights.

4. PERSONALIZED CLUSTERING

4.1 The Item-to-Item Adjacency Matrix

After determining the importance of each member in a user’s PN, we proceed into creating the item matrix A . Let I_t be the items that have already been evaluated by target user u_t above a given relevance threshold r_{rel} (that might be dependent on u_t). Then A is the $n \times n$ adjacency matrix (where $n = |I_t|$) whose $a_{i,j}$ and $a_{j,i}$ elements denote the frequency items $i, j \in I_t$ have been accessed together (above the relevance threshold) by members of u_t ’s PN. Since not all users in the PN are equally proximate to the target user, this number is scaled by each user’s relevant importance (subsection 3.2).

By definition, matrix A is symmetric and may be viewed as the adjacency matrix of an undirected graph whose nodes are the items already evaluated by u_t and whose edges represent evaluation patterns on the same itemset by users in u_t ’s PN. Naturally, one would expect some items to be accessed together more often than others and this phenomenon is reflected on the graph by the formation of item clusters. We wish to distinguish those clusters and for this reason we apply a spectral clustering algorithm [11] on the matrix A that is outlined in the following subsection.

4.2 The Clustering Algorithm

For the given symmetric item adjacency matrix A :

1. Compute diagonal Degree matrix D whose elements are $d_{i,i} = \sum_j a_{i,j}$
2. Compute the normalized symmetric Laplacian matrix $L = I - D^{-1/2}AD^{1/2}$
3. Compute the n eigenvalues of L $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and the corresponding eigenvectors $v_1 \dots v_n$. Since L is symmetric, all its eigenvalues are real, but not necessarily different; that is some eigenvalues might appear more than once, or more formally, have a multiplicity larger than 1.
4. According to the Perron-Frobenius theorem, the smallest eigenvalue of a non-negative symmetric matrix is always 0 and its multiplicity k specifies the connected components of the graph induced by the similarity matrix A . We set k as the number of the desired item clusters.
5. Construct the $n \times k$ matrix U whose column vectors are the k largest eigenvectors of L , $v_1 \dots v_k$.
6. Cluster the n row vectors of U into k clusters ($C_1 \dots C_k$) using an appropriate clustering methodology (e.g. k-means clustering). Then place each item of its corresponding cluster by using a distance criterion.

It should be noted that the eigen decomposition of step 3 is not a time consuming task, since in the overwhelming majority of the cases, matrix A (and consequently matrix L) are of low dimensionality (i.e. $n < 100$).

5. THE ITEM CONSUMPTION NETWORK

For each cluster created in the previous step of our algorithm, we construct an *Item Consumption Network* (ICN)

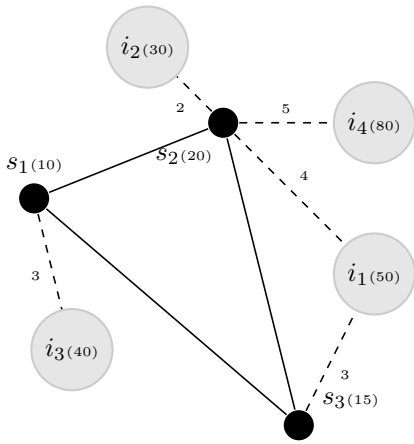


Figure 2: Item Consumption Network

[7]. An example of such a network is illustrated in Figure 2. The black dotted nodes s_1 to s_3 are all members of the same cluster and for this reason they are represented as a *clique* on the graph (interconnected with continuous edges). The gray nodes i_1 to i_4 represent other items that have been evaluated by members of the target user’s PN but not by the target user herself. An edge from a node of the former category to a node in the latter represents the fact that these two items have been accessed together by members of u_t ’s PN the number of times indicated by the weight of the respective edge. In the aforementioned figure, items s_2 and i_1 have been evaluated together by 4 peers in u_t ’s PN. Finally the number in parentheses at each node shows the number of evaluations this particular item has received from the members of the PN.

In order to produce recommendations, the ICN is modeled as a graph with the purpose of performing a modified random walk on it. Indeed, the ICN graph has the properties of being connected and non-bipartite (since it contains a clique) and for this reason it may be viewed as a *symmetric time-reversible finite Markov chain* [8]. The symmetric property stems from the fact that the transition probability $p_{i,j}$ (of moving from item i to j) equals $p_{j,i}$ (this could also be easily deduced from the undirected edges of the graph of Figure 2). The term modified is a consequence of the fact that $p_{i,j}$ is not set to be inversely proportional to the degree of the current node, as in uniform random walks. Instead, it is the result of the computation of the edge weight and the number of the evaluations of both the current and the following node in the walk.

A fundamental property of the random walks on finite Markov chains is that they reach their steady state distribution π regardless of which is the starting node each time [8]. If M is the $m \times m$ symmetric transition matrix of the chain (where m is the cardinality of the itemset in the ICN) such that $M = (p_{i,j}), i, j \in V, (i, j) \in E$ then the following equation holds true

$$M^T \pi = \pi \quad (1)$$

which may be interpreted as the fact that the matrix M has its largest eigenvalue equal to 1 and that the corresponding left eigenvector is the steady state distribution π . Consequently, if we perform an eigen decomposition on matrix M , we are able to retrieve the steady state vector π whose

entries correspond to the probability of each item node being visited by the random walk.

In some cases, however, the dimensionality of M could be large and therefore eigen decomposition may become a time consuming task. For this reason, we follow the iterative approach outlined below in order to approximate π [8]:

1. Create the $m \times 1$ vector r and set its entries to 0 apart from those which correspond to the seed items that assume the value $\frac{1}{|S|}$, ($|S|$ is the cardinality of the seed itemset)
2. Compute $w = M \times r$.
3. While $dist(w, r) > \epsilon$. Here $dist(w, r)$ denotes a function that computes vector distance in a metric space (e.g. euclidean distance) and ϵ is the vector similarity threshold (e.g. $\epsilon = 10^{-4}$)
 - (a) Set $r = w$
 - (b) Compute $w = M \times r$
4. End While
5. Return as recommendations the N non-seed nodes with the largest probability value in w .

6. EXPERIMENTS

We performed a series of initial experiments of our algorithm and of other reference content-based, collaborative and trust based systems on the Epinions dataset crawled from the Epinions website by Massa and Bhattacharje [10]. Table 1 outlines the characteristics of this dataset. It should be noted that trust statements in this case are unweighted (binary). A first remark is that the dataset is extremely sparse both in terms of the ratings’ density and of the connectedness of the trust network as measured by the global clustering coefficient. Besides sparsity, the dataset contains a large proportion of users and items with too few ratings. These characteristics greatly affect the quality of the produced recommendations.

Another peculiarity of this specific dataset is that the ratings of the items are not uniformly distributed in the rating scale (1 to 5) but are skewed towards the upper scale (4 and 5) by a ratio of 1 to 3. This is attributed to the behavioral phenomenon of users not rating items they’ve consumed in general, but of predominantly rating items they have both consumed and liked. Therefore, any RS that would blindly recommend an item with a preference rating between 4 and 5 would exhibit satisfactory performance. This observation is both an indication that RS efficiency should not solely rely on optimizing the accuracy of the predicted preference value and a justification of examining other aspects of RS performance, such as the novelty and the diversity of the recommendations (that will be discussed below).

Since one of our stated goals has been the production of accurate recommendations, we have chosen to evaluate RS performance on the *Statistical Accuracy* metrics. Their purpose is to measure how close the recommendation $\widehat{r}_{u,i}$ is to the actual rating $r_{u,i}$. The most widely used metric in this sense is the *Root Mean Square Error* (RMSE), which is defined over a Test Set T as:

$$RMSE = \sqrt{\frac{1}{|T|} \sum_{n=1}^{|T|} (\widehat{r}_{u,i} - r_{u,i})^2} \quad (2)$$

Table 1: The Epinions Dataset [10]

Epinions	
Users	49290
Items	139738
Ratings	664824
Ratings' Density	0,01%
Trust Statements	487182
Global Clustering Coefficient (Trust)	0,0002

where $|T|$ is the cardinality of the test set.

Another important evaluation metric of RS is the *Ratings' Coverage*; that is the percentage of ratings for which the system manages to produce recommendations. It should be pointed out that a RS which exhibits satisfactory results in the statistical accuracy metrics is still considered to perform poorly if it manages to produce recommendations only for a handful of users or items. More formally, the *Rating's Coverage* is defined as

$$Coverage = 100 \frac{|T_R|}{|T|} \quad (3)$$

where $|T_R|$ is the cardinality of the set of the items for which the RS produced recommendations (generally, $T_R \subseteq T$).

Our objective has also been to measure the quality of the recommendations in terms of how novel the new items are, compared to what the target user has already evaluated. Clearly, a RS that makes palpable recommendations is considered to be of low performance even if it is accurate. In our experiments, we used the definition of *Distance-based Item Novelty* [3], which models the novelty of each item (in a list of recommended items) with respect to all the other already evaluated items on an euclidean space. More specifically, it is defined as the average distance between the item i at hand and the other already consumed items (that form the set S):

$$novelty(i|S) = \sum_{j \in S} p(j|S) d(i, j) \quad (4)$$

where $p(j|S)$ denotes the probability of item j to have been already evaluated and $d(i, j)$ is the distance measure. Most commonly, distance is related to similarity via the relation $d(i, j) = 1 - sim(i, j)$. The similarity measure can be any of the similarity measures mentioned before (Pearson, cosine, etc). In our experiments, we used the Manhattan distance measure which is defined as the absolute distance between items i and j (Equation 5)

$$d(i, j) = |r_{u,i} - \widehat{r_{u,i}}| \quad (5)$$

The last evaluation measure we have examined is diversity. Diversity measures how different are the recommended items in a list from one another. Ideally, a RS that is able to capture the multitude of human taste should be equally able to propose items that would fulfill most of the interests of its users. We used the definition of *Intra-List Diversity* [3]

$$diversity(R|u) = \frac{2}{|R|(|R| + 1)} \sum_{k < n} d(i_k, i_n) \quad (6)$$

where R is the list of n recommended items ($n = |R|$) and $i_k \in R$.

6.1 Other Systems

In order to thoroughly examine the performance of our proposed algorithm, we have implemented a number of RS (both traditional and social) and have evaluated them on the same Epinions dataset.

6.1.1 Baseline Systems

Those systems are given for reference purposes, in order estimate the relative performance improvement of the other approaches. *ItemMean* recommends each item according to the mean value of the evaluations received for that item so far, treating each rater equally (that is, without considering if there exist any relationships between the raters). On the other hand, *UserMean* recommends each item according to the mean value of the evaluations given by the target user to other items so far, treating each item equally (and not considering any correlations in-between the items).

6.1.2 Collaborative Filtering and Item-Based Recommendation

At the heart of those traditional RS is the prediction formula proposed in [2]:

$$\widehat{r_{u_t, i_t}} = \bar{r}_{u_t} + \frac{\sum_{i=1}^{|N|} w_{u_t, u_c} (\bar{u}_t - r_{u_t, i_t})}{\sum_{i=1}^{|N|} w_{u_t, u_c}} \quad (7)$$

where u_t is the target user and $u_c \in U_N$ all of his neighbors with whose the similarity value w_{u_t, u_c} is computable. Equation 7 takes an equivalent form for the item based recommender, where \bar{r}_{u_t} is substituted by \bar{r}_{i_t} and w_{u_t, u_c} by w_{i_t, i_c} and $i_c \in I_N$ is now the set of items similar to i_t (whose similarity value w_{i_t, i_c} is computable).

6.1.3 Trust-based Systems

Trust-based systems are roughly put into two large categories according to the way trust values are processed [6]. The first approach is to accumulate all available trust statements in the RS in order to estimate the system-wide influence of each user. The RS that follow this principle are also known as *reputation systems* and rely on *global trust metrics* in order to calculate the reputation of each and every user. The second approach is to examine trust in a user-centric level; the emphasis is placed on each individual user and the RS departs from her in order to explore her trust network.

Since our algorithm processes the trust network on a local rather than a global level, we sought comparisons with two other SRS based on *local trust metrics*. The first SRS is based on the *gradual trust metric MoleTrust* [9] proposed by Massa and Avesani. The trust graph is firstly transformed into an acyclic form (a tree) by removing all loops in it and then the trust statements are accumulated in a depth-first fashion, starting from each user, up to each and very other user (in the trust network). The *propagation horizon* determines the length of the exploration; the most common forms being MoleTrust-1, where only the users that target user trusts are considered, and MoleTrust-2, where the exploration also includes those trusted by those the target user trusts. More formally, if T_{u_t} is the set that includes all users in u_t 's trust network that have rated item i_t (which has not been evaluated by the target user yet), then the recommendation value $\widehat{r_{u_t, i_t}}$ is approximated using the following

Table 2: Results on the Epinions Dataset (for a list of 5 recommended items)

Performance Metrics	RMSE	Coverage	Novelty	Diversity
A. Baseline				
A.1 ItemMean	1.09	86.43%	11.89%	24.23%
A.2 UserMean	1.20	98.58%	9.70%	19.42%
B. Collaborative Filtering				
B.1 Manhattan Similarity (All Neighbors)	1.07	79.57%	20.11%	56.23%
C. Item-Based Recommendation				
C.1 Manhattan Similarity (All Similar Items)	1.20	39.29%	16.86%	45.26%
D. Trust-based Approaches				
D.1 MoleTrust-1	1.23	25.58%	29.16%	43.62%
D.2 MoleTrust-2	1.16	56.52%	32.31%	54.02%
D.3 MoleTrust-3	1.12	70.89%	42.13%	56.65%
D.4 TidalTrust	1.08	74.67%	45.38%	59.17%
E. Our Recommender				
E.1 Personalized Item Clustering	1.05	58.17%	53.11%	63.04%

formula (*trust-based collaborative filtering*):

$$\widehat{r}_{u_t, i_t} = \overline{r}_{u_t} + \frac{\sum_{u \in T_{u_t}} t_{u_t, u} (r_{u, i_t} - \overline{r}_u)}{\sum_{u \in T_{u_t}} t_{u_t, u}} \quad (8)$$

where \overline{r}_{u_t} is the mean of the ratings u_t has provided so far.

The second RS is based on another popular gradual trust metric, *TidalTrust*, proposed by Golbeck [6]. *TidalTrust* is different from *MoleTrust* in the sense that no propagation horizon is required for the accumulation of trust; instead the shortest path from the target user to each other user in the trust network is computed. All trust paths above a predefined threshold form the *Web of Trust (WOT)* for that particular user. If there exist more than one trust paths between two users, then the one with the biggest trust value is chosen. If WOT_{u_t} is the set that includes those users in u_t 's web of trust network that have rated item i_t , then the recommendation value \widehat{r}_{u_t, i_t} is approximated using the formula (*trust-based weighted mean*):

$$\widehat{r}_{u_t, i_t} = \frac{\sum_{u \in T_{u_t}} t_{u_t, u} r_{u, i_t}}{\sum_{u \in T_{u_t}} t_{u_t, u}} \quad (9)$$

6.2 Results

Table 2 summarizes some first results on the set of the performance metrics presented earlier in this section. All results were obtained on the Epinions dataset by using the *leave-one-out* validation methodology, for a list of 5 recommended items. Naturally, a lower RMSE score means more accurate predictions while higher scores for coverage, novelty and diversity signify that the RS is able to produce recommendations for more of its users, that are more novel and more diverse respectively (the last three metrics are given in the percentage scale). A RS is then considered to outperform another if it achieves a better outcome on all (or most of) the performance metrics.

A first observation is that our algorithm exhibits a steady performance lead in terms of the accuracy, the novelty and the diversity of the recommendations. We attribute these encouraging results, especially the accuracy of the recommendations, to the way the personal network of each user is processed and his influential neighbors are located (sub-

section 3.2). Moreover, the increased novelty and diversity of the recommended items are due to the personalized item clustering strategy that we have followed. Indeed, in many cases, this technique manages to capture the diverse interests of each individual user.

Some of the trust-based system display a very satisfactory behavior in terms of the ratings' coverage. This stems from the aggressive manner in which they process the trust network, by accumulating all users up to a certain depth. It is for this reason, of not being selective, that they fail to achieve better results in terms of the accuracy of the recommendations, even though they are able to cover more users. However, trust statements in the dataset used in the experiments are binary; had they assumed values in a broader range, then the Equations 8 and 9 would have revealed their full potential.

Lastly, the baseline systems may exhibit very satisfactory results in terms of the accuracy and coverage metrics (particularly *ItemMean*) but those are rather the consequences of the peculiarities of the dataset at hand. As it has already been discussed before, a RS that would recommend any item with a value in the range [4, 5] would exhibit satisfactory results. It is obvious, however, that this approach is not plausible for any practical RS.

7. CONCLUSIONS AND FUTURE WORK

In this work, we proposed a novel social recommendation methodology based on personalized item clustering. Even though the initial results are satisfactory, we feel there is room for further enhancements on our algorithm. More specifically, it would be desirable to improve the procedure of the formation of the personal network. A line of research we are currently examining is to model trust and similarity as the marginal probabilities of an unknown joint probability distribution that we would like to approximate. Then user proximity would be estimated by directly sampling from the aforementioned distribution.

The spectral clustering algorithm could also be further developed by introducing some criteria that would examine the size and the quality of the produced clusters. This may be achieved by considering different clustering policies, such as the fuzzy k-means clustering, along with different distance functions (e.g. Chebysev or Mahalanobis distance) for the placement of the items within the clusters.

Finally, the properties of the random walk on the item consumption network (i.e. mixing rate) may be fine-tuned if we consider different transition probabilities, ones that would incorporate more information about the characteristics of each item node. A possible line of research in this area is to use more fine-grained approaches to model the relationship between the items.

8. REFERENCES

- [1] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutierrez. Recommender systems survey. *Knowledge-Based Systems*, 46(0):109 – 132, 2013.
- [2] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithm for collaborative filtering. In *Proceedings of the 14 th Conference on Uncertainty in Artificial Intelligence*, pages 43–52, 1998.
- [3] P. Castells, S. Vargas, and J. Wang. Novelty and diversity metrics for recommender systems: Choice, discovery and relevance. In *International Workshop on Diversity in Document Retrieval (DDR 2011) at the 33rd European Conference on Information Retrieval (ECIR 2011)*, Apr. 2011.
- [4] T. DuBois, J. Golbeck, J. Kleint, and A. Srinivasan. Improving recommendation accuracy by clustering social networks with trust. In *Recommender Systems & the Social Web*, volume 532, pages 1–8, 2009.
- [5] X. Z. Georgios Pitsilis and W. Wang. *5th IFIP WG 11.11 International Conference, IFIPTM 2011, Copenhagen, Denmark, June 29 - July 1, 2011. Proceedings*, chapter Clustering Recommenders in Collaborative Filtering Using Explicit Trust Information, pages 82–97. 358. Springer Berlin Heidelberg, 2011.
- [6] J. A. Golbeck. *Computing and applying trust in web-based social networks*. PhD thesis, College Park, MD, USA, 2005. AAI3178583.
- [7] Q. Liu, B. Xiang, E. Chen, Y. Ge, H. Xiong, T. Bao, and Y. Zheng. Influential seed items recommendation. In *Proceedings of the sixth ACM conference on Recommender systems, RecSys '12*, pages 245–248, New York, NY, USA, 2012. ACM.
- [8] L. Lovasz. Random walks on graphs: A survey, 1993.
- [9] P. Massa and P. Avesani. Trust metrics on controversial users: balancing between tyranny of the majority and echo chambers. *International Journal on Semantic Web and Information Systems*, 2007.
- [10] P. Massa and B. Bhattacharjee. Using trust in recommender systems: An experimental analysis. In *In Proceedings of iTrust2004 International Conference*, pages 221–235, 2004.
- [11] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, pages 849–856. MIT Press, 2001.
- [12] M. C. Pham, Y. Cao, R. Klamma, and M. Jarke. A clustering approach for collaborative filtering recommendation using social network analysis. *Journal of Universal Computer Science*, 17(4):583–604, feb 2011.
- [13] P. Singla and M. Richardson. Yes, there is a correlation: - from social networks to personal behavior on the web. In *Proceedings of the 17th international conference on World Wide Web, WWW '08*, pages 655–664, New York, NY, USA, 2008. ACM.
- [14] J. Sun and J. Tang. A survey of models and algorithms for social influence analysis. In C. C. Aggarwal, editor, *Social Network Data Analytics*, pages 177–214. Springer US, 2011.
- [15] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1st edition, November 1994.