Social PrefRec framework: leveraging recommender systems based on social information

Cricia Z. Felício¹,², Klérisson V. R. Paixão³, Guilherme Alves², Sandra de Amo²

¹ Federal Institute of Triângulo Mineiro, Brazil
² Federal University of Uberlândia, Brazil
cricia@iftm.edu.br, klerisson@doutorado.ufu.br, guilhermealves@mestrado.ufu.br, deamo@ufu.br

Abstract. Social recommender systems assume a social network among users and make recommendations based on ratings of users that hold a relationship with a given user. However, explicit user's ratings suffer from loss of information. One way to deal with such problem is mining preferences from user's ratings. Even though, for a new user, a preference recommender system also needs techniques to provide accurate recommendations. In this paper, we present Social PrefRec, a social pairwise preference recommender system based on Preference Mining techniques. We focus on leveraging social information on pairwise preference recommender system, corroborating with the idea that matching new people with existing similar people help on providing accurate recommendations. Remark that our approach makes use of social information only on recommendation phase to select among existent recommendation models the most appropriate for a new user. In addition, this is the first step towards a general framework to incorporate social information in traditional approaches, improving upon the state-of-art in this context. We test this idea against two real data sets from Facebook and Flixster. We contribute to this line of work in three ways: (1) Social PrefRec, a social framework for pairwise preference recommender system; (2) a strategy for recommending items based on social metrics; (3) Two publicly available data set of item ratings with social information. For cold start users, the empirical analysis demonstrates that Social PrefRec reaches nDCG@10 equals to 0.9869.

Categories and Subject Descriptors: H.3.3 [Information Storage and Retrieval]: Clustering-Information filtering; J.4 [Computer Applications]: Social and behavioral sciences
Keywords: Pairwise preferences, Social Recommender System, Social Network

1. INTRODUCTION

Social recommender systems are becoming increasingly important to help users to find relevant content. This is in part because of social media contents now account for the majority of content published on web. Typical social recommender systems assume a social network among users and makes recommendations based on the ratings of the users that have direct or indirect social relations with the target user [Jamali and Ester 2010]. However, explicit user’s ratings suffer from two known drawbacks: (i) The problems of calibration (consistency), which consists in incompatible users ratings on same scale, for example, on 1 to 5 star ratings scale, a rating of 4 for user X might be comparable to a rating of 5 for user Y. (ii) Resolution (granularity), this problem states that any numeric scale for ratings, say 1 to 5 stars, may be insufficient to capture all the users interests without loss of information [Balakrishnan and Chopra 2012] [de Amo and Ramos 2014]. Thus, we advance previous work, PrefRec [de Amo and Oliveira 2014], proposing Social PrefRec a social recommender that applies user preference mining and clustering techniques to incorporate social information on the pairwise preference recommender system.

One of the most significant discussions in recommender system field is the user cold start problem. This

We would like to thank all volunteers who took time to participate in our survey. C. Z. Felício would like to thank Federal Institute of Triângulo Mineiro for granting her study leave. We also thank the Brazilian Research Agencies CAPES, CNPq and FAPEMIG for supporting this work.

Copyright © 2012 Permission to copy without fee all or part of the material printed in KDMiLe is granted provided that the copies are not made or distributed for commercial advantage, and that notice is given that copying is by permission of the Sociedade Brasileira de Computação.

Symposium on Knowledge Discovery, Mining and Learning, KDMiLe 2015.
problem appears when users do not receive any recommendation, because they had not previously rated any item (user cold start or new user problem). Furthermore, the recommendation process worsens when it faces data sparsity problem. This latter problem is characterized by a system with too many items to be rated and few ratings per user, and also when the number of items rated in common among users is small or zero. Researches related to social recommendation argue that social information can easily deal with new user problem and data sparsity, because instead of relying on user’s preferences, which are not available, they use available ratings from users whose hold a relationship with the target user [Ma et al. 2011] [Wang et al. 2014]. In this work, we propose an approach to incorporate social rating network to provide recommendations. To leverage social influence in our model, we exploit several well known social network metrics (Section 3.2).

In addition, model-based social recommender systems in general make use of social information to build recommendation models. Thus, for each new user a new model must be built for each of them. In comparison, our approach harnessing pre-existent models. Instead of building a new model from scratch for each new user, we cluster existent users and generate recommendation models for each group. Through social information we select among existent models the most appropriated for a new user.

The main hypothesis of this paper is that matching people through their similarities can help on providing accurate recommendations in a pairwise preference recommender. It is addressed by investigating two research questions:

**RQ 1:** How accurately social information help on pairwise preference recommendation?

**RQ 2:** How relevant are the recommendations made by a social pairwise preference recommender?

**Main Contributions.** The main contributions of this paper can be summarized as follows: (1) The introduction of Social PrefRec, a social recommender system which incorporates social information in pairwise preference approach. (2) Strategies for recommending items based on social metrics. Social PrefRec achieves significantly highly correctness of ranking, calculated using the normalized Discounted Cumulative Gain (nDCG), in particular for cold start users. (3) Two publicly available real life datasets from facebook.com and flixter.com have been used to validate our proposal. The former we crawled and the existing latter we enriched with movie information from imdb.com.

**Organization of the Paper.** This paper reads as follows. Section 2 presents the background knowledge undertaking in this work. Section 3 describes our proposed framework the Social PrefRec, as well as the applied social metrics and recommender model selection strategies. Section 4 describes our experimental settings and results. Then, Section 5 discusses related work and, finally, Section 6 concludes the paper.

## 2. BACKGROUND

In this section we briefly introduce the main concepts underlying this work. Due to the lack of space, please refer to de Amo and Oliveira [2014] for more details on pairwise preference recommender systems.

A *preference relation* on a finite set of objects $A = \{a_1, a_2, ..., a_n\}$ is a strict partial order over $A$, that is a binary relation $R \subseteq A \times A$ satisfying the irreflexibility and transitivity properties. We denote by $a_1 > a_2$ the fact that $a_1$ is preferred to $a_2$. A *contextual preference model* is modeled as a *Bayesian Preference Network* (BPN) over a relational schema $R(A_1, ..., A_n)$. A BPN is a pair $(G, \theta)$ where $G$ is a directed acyclic graph whose nodes are attributes and edges stand attribute dependency, and $\theta$ is a mapping that associates to each node of $G$ a set of probability’s rules of the form: $A_1 = a_1 \land \ldots \land A_m = a_m \rightarrow X = x_1 \lor X = x_2$ where $A_1, \ldots, A_m, X$ are item attributes. The left side of the rule is called the context and the right side is the preference on the values of the attribute $X$. This rule reads: if the values of the attributes $A_1, \ldots, A_m$ are respectively $a_1, \ldots, a_m$ then I prefer $x_1$ to $x_2$ for the attribute $X$. Remark that the preferences on $X$ depends on the values of the context attributes. A contextual preference model is capable to compare items: given two items $i_1$ and $i_2$, the model is capable to predict which one is the preferred.

A *recommendation model* is constituted by a set $M = \{\theta_1, P_1\}, \ldots, (\theta_k, P_k)\}$, where $k$ is the number of groups in user-item matrix, computed by profile similarities, and for each $i = 1, \ldots, k$, $\theta_i$ is the consensual preference.
vector (preferences’ group vector expressed by average of group items rates) and $P_i$ is the preference model extracted from $\theta_i$. The output is a ranking $<i_1, i_2, \ldots, i_n>$ where an item $i_k$ is preferred or indifferent to an item $i_m$, for $k < m$ and $k, m \in [1, \ldots, n]$.

Considering the relational schema of movies attributes in Table I, we build, from $C_1$ consensus ratings (Table II), the pairwise preference relation (Table III). Thus, we are able to define the BPN depicted in Fig. 1 and then compare a set of items pair. For more details see [de Amo et al. 2013].

### Table I: Movie dataset.

<table>
<thead>
<tr>
<th>Title</th>
<th>Decade</th>
<th>Director</th>
<th>Star</th>
<th>Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giangs of New York</td>
<td>2000</td>
<td>Scorsese</td>
<td>Di Caprio</td>
<td>Drama</td>
</tr>
<tr>
<td>Catch me If You Can</td>
<td>2000</td>
<td>Spielberg</td>
<td>Di Caprio</td>
<td>Drama</td>
</tr>
<tr>
<td>The Terminal</td>
<td>2000</td>
<td>Spielberg</td>
<td>Tom Hanks</td>
<td>Drama</td>
</tr>
<tr>
<td>The Departed</td>
<td>2000</td>
<td>Scorsese</td>
<td>Di Caprio</td>
<td>Thriller</td>
</tr>
<tr>
<td>Shutter Island</td>
<td>2010</td>
<td>Scorsese</td>
<td>Di Caprio</td>
<td>Thriller</td>
</tr>
<tr>
<td>Saving Private Ryan</td>
<td>1990</td>
<td>Spielberg</td>
<td>Tom Hanks</td>
<td>Drama</td>
</tr>
<tr>
<td>Artificial Intelligence</td>
<td>2000</td>
<td>Spielberg</td>
<td>Haley J. Osment</td>
<td>Drama</td>
</tr>
</tbody>
</table>

### Table II: Users ratings over movie dataset.

<table>
<thead>
<tr>
<th>User</th>
<th>$r_1$</th>
<th>$r_2$</th>
<th>$r_3$</th>
<th>$r_4$</th>
<th>$r_5$</th>
<th>$r_6$</th>
<th>$r_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ted</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Zoe</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fred</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mary</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rose</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paul</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>John</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table III: $C_1$ pairwise preference relation

- $(t_1 > t_2)$
- $(t_1 > t_3)$
- $(t_2 > t_3)$
- $(t_3 > t_4)$
- $(t_4 > t_5)$
- $(t_5 > t_6)$
- $(t_6 > t_7)$

**Fig. 1:** Bayesian Preference Network over $C_1$ preferences.

### 3. SOCIAL PREFREC

**Social PrefRec** proposes a new approach to address new user problem through social information. It is a PrefRec framework extension, incorporating social information at recommendation module. There were no modifications on how models are built, but at recommendation phase we propose an alternative based on social information to recommend items for new users.

Recommendation process for a new user using social information, in a simple way, could recommend items well rated by his direct friends. Another option is to leverage the tie strength among friends to provide better recommendations. The challenge here is to determine how much influence or similarities exists among user’s relationship. Tie strength among users can be computed through similarities on profiles (profession, age bracket, location, etc), interaction between users (messaging, photos, etc) and degree of influence.

To support this features, PrefRec was extended considering the **Social PrefRec structure**: Let $U$ be a user set and $I$ be an item set. The user set $U$ is composed by user identifier and others attributes related to users, where $A_u = \{a_1, \ldots, a_i\}$ is an attribute set for users. Item set $I$ is composed by item identifier and others attributes related to items, where $A_i = \{a_1, \ldots, a_i\}$ is an attribute set for items. A friendship set over $U$ is defined as $F = \{(u_i, u_k) \mid u_i, u_k \in U\}$, where $(u_i, u_k) = (u_k, u_i)$. We denote by $F_i$ the set that contains all friends of $u_j$ user. The weight function $w : U \times I \rightarrow \mathbb{R}$ computes an **user preference degree** for an item and a function $l : F \rightarrow \mathbb{R}$ defines **tie strength** between $u_i$ and $u_k$. **Social PrefRec structure**, shown on Fig. 2, consists of: one graph $G = (U, I, F, w, l)$. A social network in $G$ is represented by sub-graph $S_N = \{U, F, l\}$. An illustrative example of **Social PrefRec** is shown on Fig. 3. Nodes represent users and edges are friendships relations. Labels on edges indicate computed tie strength. Dashed groups are computed clusters of users. Each cluster is associated with
a recommendation model. Suppose that Paty is a new user, therefore there is no item previous rated for her. The system already knows some Paty’s friends and had previously clustered them. As soon as Paty shows up the tie strength is computed and a suitable recommendation model is selected.

Fig. 2: Social PrefRec structure

Fig. 3: Social network example

3.1 Social PrefRec Framework

Social PrefRec is an extension of PrefRec, a model-based hybrid recommender system framework using pairwise preferences mining and preferences aggregation techniques [de Amo et al. 2013]. The general Social PrefRec architecture, the interactions among the five modules, as well as their respective input and output is presented at Fig. 4. Note, that modules from 1 to 4 are from PrefRec, however we improved the later system where instead of representing user and consensus preferences in a matrix, now they are represented in a vector. This reduces the algorithm complexity, execution time and allows a better clustering step.

Fig. 4: Social PrefRec Framework

Next, we describe how recommendation module works. Recommendation model is given as input for module 5. This input is constituted by a set $M = \{(\theta_1, P_1), \ldots, (\theta_k, P_k)\}$, where for each $i = 1, \ldots, k$, $\theta_i$ is the consensual preference vector associated to cluster $C_i$ and $P_i$ is the preference model extracted from $\theta_i$.

**Recommendation Module.** The aim of this module is to use the recommendation model $M$ to recommend items for new users. It is executed online, differently from the previous modules which are executed offline. Recommendation process could be executed using one out of the two strategies:

**A) PrefRec.** (1) Given a target user $u$ and a (small) set $R_u$ of ratings provided by $u$ on items in $I$, the first task of Module 5 consists in obtaining the preference vector $\sigma_u$ corresponding to $R_u$; (2) the similarity between $\sigma_u$ and each consensual preference vector $\theta_i$ is calculated. Let $\theta_u$ be the consensual vector most similar to $\sigma_u$; (3) consider the preference model $P_u$ corresponding to $\theta_u$; (4) $P_u$ is used to infer the preference between pairs of items in $I$ which have not been rated by the user $u$ in the past. From this set of pairs of items $(i, j)$ indicating that user $u$ prefers item $i$ to item $j$, a ranking can be built by applying one ranking algorithm adapted from the algorithm Order By Preferences introduced in [Cohen et al. 1999].

**B) Social PrefRec metrics.** (1) Given a target user $u$ and a its social network $SN_u$, the first task of Module 5 consists in applying one of the social strategies (described in Section 3.2) to compute the tie strength between $u$ and it contacts; (2) obtaining the consensual vector $\theta_u$ corresponding to the cluster $C_k$ where $u_k \in SN_u$ are $u$
direct contacts using one of the chosen model methods (Average or threshold, described in Section 3.2) ; (3) and (4) are identical to PrefRec strategy. Note that by using this strategy is possible to recommend to an user without taking in consideration any previous ratings, but considering user’s relations in the cluster set.

3.2 Tie strength calculus and Recommendation model selection

We compute tie strength between users through the following metrics: (1) friendship considers that \( l(u_j, u_k) = l((u_j, u_k) \in F) \), where \( l(\cdot) \) is the characteristic function (1 if argument is true, 0 otherwise); (2) interaction level is calculated as \( \frac{a(u_j, u_k)}{a(u_j)} \), where \( a(u_j, u_k) \) is the number of times where user \( u_k \) appears at \( u_j \)'s timeline and \( a(u_j) \) is the number of all occurrences of users \( u_k \) at \( u_j \)'s timeline; (3) mutual friends considers that \( l(u_j, u_k) = F(u_j, u_k) \), where \( F \) is the Jaccard similarity coefficient; (4) in similarity score, function \( l(u_j, u_k) = \text{sims}(u_j, u_k) \) is the average of \( \text{similarity}(u_j, u_k, A_i) \) binary values for all attributes \( A_i \), where \( \text{similarity}(u_j, u_k, A_i) \) represents user \( u_j \) compatibility with an user \( u_k \) considering the demographic user attributes \( A_i \) (1 if similar, 0 otherwise) like Relationship Status, Age Bracket, Sex, Religion, Location, etc; (5) centrality as tie strength is calculated by average of closeness, betweenness and eigenvector centrality measures.

Social PrefRec recommender uses two metrics for model selection based on tie strength value:

- **Minimum threshold**: Let \( \varepsilon \in [0, 1] \) be a tie strength minimum threshold. The strategy \( C_m \) will select the preference model \( P_i \) (associated with model \( M_i \in M \)) with more users who have a tie strength with the target user \( u_j \) above a minimum threshold according to Eq. 1.

  \[
  C_m(F_j, M, u_j) = \arg \max_{M_i \in M} \left[ |u_k : (u_j, u_k) \in F_j \land l(u_j, u_k) \geq \varepsilon| \right]
  \]

- **Average**: The strategy \( C_a \) will select the preference model \( P_i \) with users who have the highest average tie strength with the target user \( u_j \) according Eq.2.

  \[
  C_a(F_j, M, u_j) = \arg \max_{M_i \in M} \frac{1}{|F_j|} \sum_{(u_j, u_k) \in F_j} l(u_j, u_k)
  \]

4. EXPERIMENTS

4.1 Datasets

**Facebook Dataset.** We surveyed this data set through a developed Facebook web application. With volunteers permission we crawled relationship status, age bracket, gender, born-in, lives-in, religion, study-in, last 25 posts in user’s timeline, posts shared and posts’ likes, as well as movies previous rated on Facebook platform. In addition, we asked each volunteer to rate 169 Oscar nominated movies in 1 to 5 star scale. We obtained data from 720 users and 1,454 movies, resulting in 56,903 ratings.

**Flixster Dataset.** Jamali and Ester [2010] published this dataset. However, movie information was restricted to its title, then we improved it by adding genres, directors, actors, year, languages and countries information retrieved from IMDB.com public data.

In our experiments we considered only the 169 movies surveyed, because there are more common movies rated among users. We split Facebook data into two datasets, FB50 and FB100, to represent the set of users that rated at least 50 and 100 movies, respectively. This was done to evaluate the overall system performance under data sets with different sparsity and level of social information. In Table IV we summarize our datasets. The movies attributes considered were genres, directors, actors, year, languages and countries. In FB50 and FB100, user similarity metric was computed using the attributes: relationship status, age bracket, gender, born-in, lives-in, religion and study-in. The interaction_level was computed considering the last 25 posts in user timeline, posts shared and likes. Flixster social information includes friends relationships, mutual friends, friends centrality and users similarities. Similarity between users is computed only through three attributes: gender, age bracket and location. Interaction information is not available on Flixster dataset.
4.2 Experimental Protocol and Evaluation methods

Each experiment was performed against the datasets split into two parts: training and test sets. Fig. 5 shows a comparative scheme of our protocols. PrefRec and Social PrefRec make use of training data to build clusters (K-Means clustering) of similar users. For each cluster they associate a correspondent recommendation model. Then, to recommend items for a given user \( u \), is necessary to select the most similar model (cluster) that fits \( u \). This process is done during test phase. However, those approaches take different directions. Since PrefRec is not able to deal with social information, it relies only on previous ratings of \( u \) to select its best recommendation model, whereas Social PrefRec needs social information for this choice. To better validate our tests we apply the leave-n-out cross validation protocol: for each iteration one user is taken for test purposes and the remaining users assemble the training set. In this case we have \( n \) iterations, where \( n \) is the number of users.

**Table IV: Movies Datasets**

<table>
<thead>
<tr>
<th>Features</th>
<th>FB100</th>
<th>FB50</th>
<th>Flixster</th>
</tr>
</thead>
<tbody>
<tr>
<td># of users</td>
<td>230</td>
<td>361</td>
<td>357</td>
</tr>
<tr>
<td># of items</td>
<td>169</td>
<td>169</td>
<td>625</td>
</tr>
<tr>
<td># of ratings</td>
<td>35,438</td>
<td>44,925</td>
<td>175,523</td>
</tr>
<tr>
<td>Sparsity</td>
<td>8.77%</td>
<td>26.36%</td>
<td>21.33%</td>
</tr>
<tr>
<td>Friends relationship</td>
<td>1,330</td>
<td>2,926</td>
<td>706</td>
</tr>
<tr>
<td>Avg friends per user</td>
<td>6.4</td>
<td>8.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Avg rates per user</td>
<td>154.16</td>
<td>124.44</td>
<td>491</td>
</tr>
<tr>
<td>Users without friends</td>
<td>9.56%</td>
<td>5.54%</td>
<td>29.97%</td>
</tr>
</tbody>
</table>

**PrefRec protocol.** The PrefRec recommendation model is built offline. For the test phase \( m \) random ratings of current test user \( u_k \) were considered for the choice of the most similar cluster \( C_i \). Then, calculating similarity between \( u_k \) and \( c_i \) is a matter of calculating the Euclidian distance between their respective ratings arrays \( p = (p_1, \ldots, p_n) \) and \( q = (q_1, \ldots, q_n) \). Remark that this similarity distance was used for models clustering (training) and selection models (test) phases. Finally, for validation purpose, the remaining ratings of current test user \( u_k \) were used.

**Social PrefRec protocol.** There is no difference regarding recommendation model building on Social PrefRec from PrefRec. However, during test phase, we do not take in account any rating. It relies on social information to find the most similar cluster, \( C_i \), according to a given social metric and a model selection strategy.

Regarding our evaluation methods we present results from two metrics: (1) nDCG is a standard ranking quality metric to evaluate the ability of the recommender to rank the list of top-k items [Shani and Gunawardana 2011]. (2) We also compute the standard F1 score, based on precision and recall, to evaluate the predictions quality of a pairwise preference items [de Amo and Oliveira 2014].

4.3 Results

**RQ1: Quality of recommendation.** Comparative \( F_1 \) scores can be seen in Fig. 6, for minimum threshold (\( \epsilon = 0.4 \)) and tie strength average selection model strategies. In all datasets with a profile length of 30-ratings scenario for PrefRec versus 0-ratings for Social PrefRec, social metrics achieve better results using Minimum threshold strategy. Rate-15-items baseline is widely used to bootstrap traditional recommender systems [Chang et al. 2015]. Thus, to make a fair comparison we provide 30-ratings for PrefRec, which means that all runs have a good safe margin and should not harm its performance.

A Kruskal-Wallis test was performed to check statistical significance among social metrics performance and PrefRec. Regarding Mutual Friends, Interaction, Similarity there are no significant differences. Also Friendship and Centrality results are not significant different from PrefRec (profile length = 30-ratings) result. Thus, the test shows, with 95% confidence, that with the three first metrics we can better accurately recommend in social 0-rating profile scenario than 30-rating profile in a traditional recommender approach. The others social metrics achieved the same result as the traditional approach, but they do not need any previous rate from a given user.

Symposium on Knowledge Discovery, Mining and Learning, KDMILE 2015.
Social PrefRec framework: leveraging recommender systems based on social information

Mutual Interaction Similarity Friendship Centrality PrefRec

0.6 0.65 0.7

FB100 FB50 Flixster

RQ2: Relevance of recommendation. Fig. 7 shows the nDCG results for rank size 5, 10, 15 and 20, considering Minimum threshold (\(\epsilon = 0.4\)) strategy. Rank quality is better in Flixster dataset, because the number of items is greater than Facebook data, generating a richer preference model. Test of statistical significance shows, with 95% confidence, that Mutual Friends metric is better than others. The performance with Centrality achieves equivalent score as PrefRec. Finally, Similarity, Friendship and Interaction results are not significant different.

5. RELATED WORK

Pairwise Preference Recommendation. Balakrishnan and Chopra [2012] have proposed an adaptive scheme in which users are explicitly asked for their relative preference between a pair of items. Though it may provide an accurate measure of a user’s preference, explicitly asking users for their preference may not be feasible for large numbers of users or items, or desirable as a design strategy in certain cases. Park et al. [2009] proposed a pairwise preference regression model to deal with cold start user problem. We corroborate with their idea. They argue that ranking of pairwise users preferences minimize the distance between real rank of items and then could lead to better recommendation for a new user. On the same direction Sharma and Yan [2013] propose a probabilistic latent semantic indexing model for pairwise learning, which assumes a set of users’ latent preferences between pairs of items. We build on previous work [de Amo and Ramos 2014] by adapting a pairwise preference recommender to leverage a graph of information, social network.

Social Recommender. This research field especially started because social media content and recommender systems can mutually benefit from one another. Many social-enhanced recommendation algorithms are proposed to improve recommendation quality of traditional approaches [Canamares and Castells 2014] [Alexandridis et al. 2013]. Moreover, the works of Ma et. al. [2008] [2011] [2011] are the most related to this one. No matter what techniques are developed, the basic assumption employed in these works is that users’ social

Fig. 6: (a) Minimum threshold (\(\epsilon = 0, 4\)), (b) Tie strength average

Fig. 7: nDCG@5, @10, @15 and @20 for PrefRec versus Social PrefRec Metrics
relations can positively reflect users’ interests similarities. Although we also explore users’ relation in our approach, we do it in different way. Instead of embedding social information in the recommendation models, we built a loosely coupled approach based on clustering techniques to incorporate social relation into our system.

6. CONCLUSION

In this paper, we have devised and evaluated Social PrefRec, an approach whose ultimate goal is to help pairwise preferences recommender systems to deal with cold start problem. We built on the shoulders of others, and expand previous work by: (1) Working in a way to incorporate social information in pairwise preference recommender approach; (2) presenting strategies for recommending items based on several social metrics; and (3) evaluating the resulting approach on two real life data sets from facebook.com and flixter.com, which we made publicly available. This work opens several avenues for future research. First, it is worth exploring the use of others networks (graphs) where we can compute a tie from similarities scores among nodes, such as scientific networks. Furthermore, we ought to empirically compare Social PrefRec performance against benchmark social recommenders. From the application point of view, we believe that Social PrefRec framework could be generalized to others hybrid model-based recommenders, allowing traditional approaches to incorporate contextual social information.

REFERENCES


1http://www.lsi.facom.ufu.br/~cricia/

Symposium on Knowledge Discovery, Mining and Learning. KDMILE 2015.