Semi-supervised Tag Recommendation - Using Untagged Resources to Mitigate Cold-Start Problems

Christine Preisach, Leandro Balby Marinho, and Lars Schmidt-Thieme

Information Systems and Machine Learning Lab, University of Hildesheim, Germany
{preisach,marinho,schmidt-thieme}@ismll.uni-hildesheim.de

Abstract. Tag recommender systems are often used in social tagging systems, a popular family of Web 2.0 applications, to assist users in the tagging process. But in cold-start situations i.e., when new users or resources enter the system, state-of-the-art tag recommender systems perform poorly and are not always able to generate recommendations. Many user profiles contain untagged resources, which could provide valuable information especially for cold-start scenarios where tagged data is scarce. The existing methods do not explore this additional information source. In this paper we propose to use a purely graph-based semi-supervised relational approach that uses untagged posts for addressing the cold-start problem. We conduct experiments on two real-life datasets and show that our approach outperforms the state-of-the-art in many cases.

1 Introduction

Recently Web 2.0 applications like social tagging systems (or folksonomies) are getting more and more popular. One service often provided by these sites are tag recommender systems that help simplifying the process of tagging for the user. Given that users are free to tag, i.e., the same resource can be tagged differently by different people, it is important to personalize the recommended tags for an individual user. But state-of-the-art methods are not always able to suggest personalized tags for a new user or a new resource. Often these situations are handled by using the content of the new resource or just recommending the most popular tags. But these approaches have several drawbacks, first the recommended tags are not personalized and second, using content is not a generic approach, one needs to use different algorithms for each type of resource, e.g., in Last.fm[1] information needs to be extracted from the audio files, in Flickr[2] images need to be analyzed, in YouTube[3] knowledge from videos has to be extracted and for Delicious[4] and BibSonomy[5] the text of bookmarked web pages or publications belonging to a BibTeX entry needs to be assessed. Thus, this is a costly solution if one needs different approaches for several types of resources.

We propose a content independent, purely graph-based approach, which is based on the observation that user profiles usually contain many untagged posts that could be

[1] http://www.last.fm

[2] http://www.flickr.com

[3] http://www.youtube.com

[4] http://delicious.com

[5] http://www.bibsonomy.org

M.J. Zaki et al. (Eds.): PAKDD 2010, Part I, LNAI 6118, pp. 348–357, 2010.

exploited for improving the recommendations, especially when there are only a few tagged examples available. We investigate two scenarios, first where a new user enters the system and second, where among the untagged posts, there are new resources, i.e., resources that were not tagged by any other user in the system. We will address these problems by means of semi-supervised relational classification, whereby we can benefit from the structural information of untagged posts.

As presented in [1], we first cast the problem of personalized tag recommendations as a relational classification problem, where we use relational semi-supervised classification to profit from the potentially valuable information carried by other untagged resources. In contrast to our approach submitted to the ECML/PKDD Discovery Challenge 2009 (task 2) that achieved the second place, in this paper we focus on the *cold-start* problem (that did not occur in the challenge dataset).

In this paper our contributions are as follows:

1. We formally define the *cold-start* (in terms of new user/resource) problem in social tagging systems.
2. To address this problem we propose and compare different semi-supervised relational methods, which exploit the structural information of untagged posts in the post graph.
3. Finally, we show empirically that our approaches outperform the current state-of-the-art algorithms (FolkRank and PITF, a tensor factorization model), as well as other simpler baselines such as KNN and most popular tags in many cases.

2 Related Work

In [2] the authors compared several personalized tag recommendation algorithms, the best results, were achieved by the FolkRank algorithm [3], an adaptation of PageRank for retrieving information and recommending tags in social tagging systems. More recently Rendle et al. [4] introduced RTF (Ranking with Tensor Factorization), a method for learning a tensor factorization model optimized for the best personalized tag ranking. The model also handles missing values by introducing a new interpretation of the data and learns from pairwise ranking constraints through a gradient descent algorithm. The prediction runtime is independent of the number of observations and only depends on the factorization dimensions but the training time is huge. Another new factorization model for tag recommendation PITF (Pairwise Interaction Tensor Factorization) was introduced in [5,6], it tries to find latent interactions between users, items and tags by factorizing the observed tagging data. Similar to [4] the model is learned by optimizing the Bayesian Personal Ranking method with gradient decent. Although the methods discussed above provide high quality recommendations, they are not robust against new user/resource scenarios. Furthermore, RTF and PITF strictly operate over ternary relations, and thereby are not able to exploit the information of untagged posts. For item recommendation a semi-supervised approach for cold-start problems has been recently published, the authors of [7] have introduced the tied Boltzmann machine that captures pairwise interactions between items. To our best knowledge for cold-start problems in tag recommendation no graph-based, semi-supervised approach has been introduced so far.

Since tagging data forms relations between users, resources and tags, it is natural to exploit these relations by adapting relational methods to the tag recommendation scenario, in [1] we showed that relational classification methods perform very well on the ECML/PKDD Discovery Challenge 2009 dataset (which did not contain new users or resources). Originally relational methods have been applied to areas where entities are linked in an explicit manner, like hypertext documents and scientific publications. Especially iterative semi-supervised relational methods, which use *collective inference* and exploit relational autocorrelation of class labels of related entities, received attention. One of the earliest iterative semi-supervised relational approaches was proposed by Chakrabarti et al. [8], where a probabilistic model for classification of web pages was introduced. In [9,10] the authors presented different semi-supervised iterative models and showed that collective inference increases accuracy.

Here we will focus on the *cold-start problem*, where we expect that simple iterative semi-supervised relational methods outperform supervised approaches, since they allow the usage of unlabeled data, which is particularly important for *cold-start* scenarios.

3 Tag Recommendations

In this section we formalize the general problem of tag recommendations in social tagging systems, and formalize the new user/resource problem.

3.1 Problem Formulation

Social tagging systems data usually comprises a set of users U, a set of resources R, a set of tags T, and a set Y of ternary relations between them, i.e., $Y \subseteq U \times R \times T$. Let

$$X := \{(u, r) \in U \times R \,|\, \exists t \in T : (u, r, t) \in Y\}$$

be the set of all posts in the data. Let $T(x = (u, r)) := \{t \in T \mid (u, r, t) \in Y\}$ be the set of all tags assigned to a given post $x \in X$. We consider train/test splits based on posts, i.e., $X_{\text{train}}, X_{\text{test}} \subset X$ are disjunct and covering all of X: For training, the learner has access to the set X_{train} of training posts and their true tags $T|_{X_{\text{train}}}$. Semi-supervised methods also could exploit the set X_{test} of untagged posts, but of course not their associated true tags. The tag recommendation task is then to predict, for a given $x \in X_{\text{test}}$, a set $\hat{T}(x) \subseteq T$ of tags that are most likely to be used by the respective user for the respective resource.

3.2 New User/Resource

An issue that remains unaddressed by the current literature on personalized tag recommendations refers to the new user/resource problems. A new user u refers to the user who posted for the first time in the system, i.e.,

$$|X_{\text{test}} \cap (\{u\} \times R)| \geq 1 \text{ and } X_{\text{train}} \cap (\{u\} \times R) = \emptyset$$

In other words, all posts of a new user are in the test set. A new resource r, on the other hand, refers to a resource that has never been tagged before by any other user:

$$X_{\text{train}} \cap (U \times \{r\}) = \emptyset$$

Currently, there are no suitable purely graph-based[6] approaches for providing recommendations whenever these situations occur. Unpersonalized content-based approaches are usually used in such scenarios, but since the resource' format can vary across different social tagging systems one would need to develop a specific method for each possible kind of resource.

4 Semi-supervised Relational Methods

In this section we present the types of relations we use and introduce several semi-supervised relational methods for tag recommendation. Here we especially focus on the *new user/resource* scenario.

We propose to represent folksonomy data as a homogeneous, undirected relational graph over the post set, i.e., $G := (X, E)$ in which edges are annotated with a weight $w : X \times X \rightarrow \mathbb{R}$ denoting the strength of the relation. Besides making the input data more compact – we have only a binary relation $\mathcal{R} \subseteq X \times X$ between objects of the same type – this representation will allow us to cast the problem of personalized tag recommendations as a relational classification problem.

Relational classifiers usually consider, relations between objects instead of only taking into account the conventional attribute-value data of objects. A scientific paper, for example, can be related to another paper that has been written by the same author or because they share common citations. It has been shown that relational classifiers usually perform better than purely attribute-based classifiers [8,11,12].

In our case, we assume that posts are related to each other if they share the same user: $\mathcal{R}_{\text{user}} := \{(x, x') \in X \times X \mid \text{user}(x) = \text{user}(x')\}$ or the same resource: $\mathcal{R}_{\text{res}} := \{(x, x') \in X \times X \mid \text{res}(x) = \text{res}(x')\}$ as an alternative we can use both relations together, i.e, posts either share the same user or resource (see Figure 2): $\mathcal{R}^{\text{res}}_{\text{user}} := \mathcal{R}_{\text{user}} \cup \mathcal{R}_{\text{res}}$. For convenience, let $\text{user}(x = (u, r)) := u$ and $\text{res}(x = (u, r)) := r$ denote the user and resource of post x respectively. Iterative relational methods have been shown to work very well because of the following three assumptions [13]: first, the first-order Markov assumption, i.e., in the tag recommendation scenario, only the direct neighborhood is necessary for accurate tag recommendations, second, the assumption of homophily, i.e., similar posts are more likely to be tagged alike and third, the assumption of simple belief propagation, i.e., tags can be propageted to untagged posts.

We are especially interested in the situation where related posts are untagged, thus differently from other approaches (e.g., [2,4]) that use X_{train} and $T|_{X_{\text{train}}}$ allowing us to exploit the structural information of untagged posts using semi-supervised iterative relational methods. One simple relational classification method for tag recommendation is the *WeightedAverage (WA)* which sums up all weights of neighboring posts $x' \in N_x$ that share the same tag $t \in T$ and normalizes this by the sum over all weights in the neighborhood:

$$P(t|x) = \frac{\sum_{x' \in N_x | t \in T(x')} w(x, x')}{\sum_{x' \in N_x} w(x, x')} \tag{1}$$

[6] By graph-based we mean algorithms that do not rely on resources' content but only on the graph induced by the folksonomy data.

with $N_x := \{x' \in X \mid (x, x') \in \mathcal{R}\}$ being the neighborhood. This algorithm (in the following denoted as *WA*) is similar to collaborative filtering, which is based on the k-Nearest Neighbor algorithm, the difference here is that k, does not need to be determined but is given by the number of neighbors.

But what if some of the neighboring posts are untagged, how should we handle this situation? State-of-the-art methods just ignore untagged posts. Semi-supervised iterative methods in contrast do exploit them and thus, increase classification accuracy. One simple way of considering untagged posts is to transform *WA* into an iterative algorithm, i.e., in the first iteration we classify test posts by only using direct neighbors from the training set, in the second iteration, the still unclassified test posts are classified by extracting tag information from neighbors that have been classified in the previous iteration. The procedure stops when all the test instances are classified. This iterative version of *WA* is denoted as *WAOneShot* since all test posts are classified only once, i.e., already classified posts are not re-classified in the following iteration. Note that eq. (1) considers the tags of the neighborhood in a deterministic way i.e., probabilities are not taken into account, so that even if the probability of a estimated tag is very low it is considered in the same way as a high probability tag. To overcome this limitation one can extend eq. (1) to *PWA* (*Probabilistic Weighted Average*):

$$P(t|x) = \frac{\sum_{x' \in N_x} w(x, x') P(t|x')}{\sum_{x' \in N_x} w(x, x')} \tag{2}$$

Now, instead of only summing up edge weights of direct neighbors, we additionally take into account the probability of the tags belonging to those neighbors. Combining eq. (2) with the aforementioned iterative algorithm leads to a probabilistic semi-supervised iterative method which makes use of the uncertainty of tag estimations in previous iterations. This algorithm is denoted as *PWAOneShot*. The algorithms, *WA*, *WAOneShot*, *PWA* and *PWAOneShot* use only the first two properties of relational methods, namely the first-order assumption and the homophily but not the third property. Thus, for both algorithms test posts are not re-classified even if tags of neighbors have changed, i.e., the information cannot be spread in the graph.

To resolve this problem relaxation labeling [8] can be used, i.e., we make use of the third property of simple belief propagation, here the tag probability of all test posts are re-estimated simultaneously in each iteration, i.e., the infomation spreads in the graph which helps to increase accuracy. *PWA* combined with relaxation labeling is denoted as *PWA**. The algorithm for *PWA** is depicted in Figure 1. There we first initialize the untagged posts with the prior probability calculated using the training set, then we compute the probability of each tag t given x iteratively using *PWA*. The procedure stops when the algorithm converges (i.e., the difference of the tag probability between iteration i and $i + 1$ is less than a very small ϵ) or a certain number of iterations is reached. Note that eq. (1) and eq. (2) have been introduced in [14] and applied to relational datasets. The weight w in eq. (1) and (2) is an important factor in the estimation of tag probabilities, since it describes the strength of the relation between x and x'. We used the weight schemes described in [1]. Since we want to recommend more than one tag we need to cast the tag recommendation problem as a multilabel classification problem, i.e., assign one or more tags to a test post. We accomplish the multilabel problem by

learn-PWA*$(X_{\text{train}}, X_{\text{test}}, T, \epsilon)$:
$P(t \mid x)^0 := 1, P(t' \mid x) := 0$ for all $x \in X_{\text{train}}, t \in T(x), t' \notin T(x)$
$P(t \mid x)^0 :=$ prior(t) for all $x \in X_{\text{test}}, t \in T$
for $i := 0, \ldots, I$ do
    for $x \in X_{\text{test}}$ do
        for $t \in T$ do
            $P(t \mid x)^{i+1} := \frac{1}{Z} \sum_{x' \in N_x} w(x, x') P(t \mid x')^i$
        od
    od
od until $\sqrt{\frac{1}{|X_{\text{test}}| \cdot |T|} \sum_{x \in X_{\text{test}}} \sum_{t \in T} (p(t \mid x)^{i+1} - p(t \mid x)^i)^2} < \epsilon$
return $(P(t \mid x))_{x \in X_{\text{test}}, t \in T}$

Fig. 1. Algorithm *PWA**

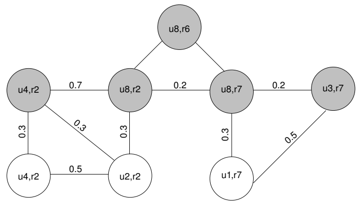

Fig. 2. Relational graph for the new user u_8 and new resource r_6

sorting the calculated probabilities $P(t|x)$ for all $x \in X_{\text{test}}$ and recommend the top n tags with highest probabilities.

In terms of runtime complexity, *PWA** is in $O(I \cdot (|T| \cdot N_x))$ for prediction and $O(1)$ for training. I.e., the runtime is only dependent on the number of iterations, number of tags and the size of the neighborhood.

4.1 Cold-Start Problem

Semi-Supervised relational classification is especially usefull for addressing *cold-start* problems where users have untagged resources in their profiles, since, in contrast to the current state-of-the-art, it is able to exploit the structural information of untagged posts and to propagate this information in the post graph. In general our semi-supervised approach is able to extract information from two sources and two relations, from the tagged posts and untagged posts over $\mathcal{R}_{\text{res}}$ or $\mathcal{R}_{\text{user}}$ which is very beneficial for the coldstart problem.

In figure 2 we illustrate the new user/resource problem. The gray nodes in the given post graph represent the untagged posts, the white nodes belong to the training set. In our example user u_8 is a new user, she has several untagged posts. In order to recommend tags for post (u_8, r_7) for example, we can make use of both information sources, training and test set. First, through $\mathcal{R}_{\text{res}}$ with (u_1, r_7) from the training set and (u_3, r_7) from the test set (this post is an untagged post of user u_3). Second, through $\mathcal{R}_{\text{user}}$ with her own untagged posts (u_8, r_2) and (u_8, r_6). Thus, the system can profit from both, the training set over the resource relation and from the untagged posts belonging to the users own profile. For the new resource problem in contrast we cannot use $\mathcal{R}_{\text{res}}$, but only $\mathcal{R}_{\text{user}}$. In figure 2 r_6 is a new resource, so if we want to recommend tags for (u_8, r_6) one can exploit (u_8, r_2) and (u_8, r_7). Although (u_8, r_2) and (u_8, r_7) are initially untagged, our methods still benefit from them, because they are connected to other posts and this information spreads over the graph. Since our graph can be composed by two kinds of relations at the same time, when one relation is missing (e.g., new user or resource), there is always another relation as backup. The only scenario where this does not hold is when all the resources uploaded by a new user are new.

5 Experiments

The main issues we want to address here is the new users/resources scenario. We conduct two main experiments to show that our semi-supervised methods are able to cope

Table 1. Characteristics of 5-core BibSonomy and 10-core Last.fm

Dataset	$\lvert U \rvert$	$\lvert R \rvert$	$\lvert T \rvert$	$\lvert Y \rvert$	$\lvert X \rvert$	$\lvert E_{\text{user}} \rvert$	$\lvert E_{\text{res}} \rvert$	$\lvert E_{\text{user}}^{\text{res}} \rvert$
BibSonomy	116	361	412	10,148	2,522	64,669	9,108	73,777
Last.fm	2,917	1,853	2,045	219,702	75,565	1,088,023	4,149,862	5,237,885

with these issues. We compare several semi-supervised relational models (*WAOneShot*, *PWAOneShot*, *PWA**) with state-of-the-art methods like WA, FolkRank[7], PITF (tensor factorization model), and the most popular tags on two real-world datasets.

5.1 Datasets

In order to evaluate our approach we use two real-life datasets, BibSonomy and Last.fm[8]. BibSonomy is a social tagging system that allows users to manage and annotate bookmarks and publication references simultaneously. Last.fm on the other hand, is a social online radio station where people can upload, share and tag music/artists/albums they like. Since these systems represent different domains and are evtl. used by different people, we assume that our findings can also be carried over to other social tagging systems. We follow the conventional approach of using the dense part of Y by means of a *p-core*[9]. Similarly to [2,4], we used the 5-core for BibSonomy and the 10-core for Last.fm. Table 1 summarizes the characteristics of the datasets we used. For convenience, let $|E_{\text{res}}|$, $|E_{\text{user}}|$ and $|E_{\text{user}}^{\text{res}}|$ denote the number or edges according to the $\mathcal{R}_{\text{res}}$, $\mathcal{R}_{\text{user}}$, $\mathcal{R}_{\text{user}}^{\text{res}}$ relations respectively.

5.2 Experiment Setting

We analyzed two situations, one where only new users, and a second where new users and new resources were present in the data. For the first situation (new user problem), the test set is composed by only new users (we sampled 30% of the users in U to be in the test set, the rest is used for training) but no new resources. For the second scenario, i.e., new user and new resource problem, we sampled for each user a percentage of test resources to be new. We evaluated our methods on data where 1% and 10% of the test posts contain new resources. In reality many users have untagged resources in their profiles but those untagged posts are usually removed from the standard datasets, thus we simulated this situation. We use the standard *LeavePostOut* [2] protocol, but additionally exploit the untagged posts, i.e., while recommending tags for one post, the other sampled posts are used as untagged posts (their tags are removed).

We used the standard $F1$ measure on top-5 tag lists, similar to [5] we estimate the optimal number of tags to be recommended (i.e., we do not always recommend 5 tags). As in [1] we rewarded the best relation by a weight of c^{10}. We optimized c as well as

[7] Parameters $d := 0.4$, #iterations:=10.

[8] We have used the same data snapshots as in [2,4].

[9] A p-core of X is the largest subset of X where each user, resource and tag must occur in at least p posts.

[10] BibSonomy $c = 2.5$ for $\mathcal{R}_{\text{res}}$, Last.fm $c = 2.5$ for $\mathcal{R}_{\text{user}}$.

the hyperparameters for PITF on a holdout set. Moreover, we restricted the maximum number of iterations for the *PWA** algorithm to 75.

5.3 Results and Discussion

New User Problem. Figure 3 shows the results achieved with *PWA**, *WAOneShot*, *PWAOneShot*, FolkRank, PITF, WA and the most popular [11] algorithm for various numbers of recommendations (1-5) applied to BibSonomy. All users in the test set are new but all resources are known in the system already. This is reflected in the results, the difference among the results is small, *PWAOneShot*, *WAOneShot* and *WA* perform very similar. *WA* performs well since every test post is connected to at least one training post, thus recommendation of tags is possible for each test post. But even in this situation where there are more connections between test and training set, *PWA** outperforms the other algorithms, i.e., can still profit from label propagation and initialization of test posts. The situation changes for Last.fm (see Figure 4). Here both FolkRank and PITF outperform all other methods. *PWA** performs best among the semi-supervised methods, but seems not to profit so much from label propagation and initialization of test posts. *PWAOneShot*, *WAOneShot* and *WA* achieve again very similar results. The reason lies in the nature of this dataset, here the user relation contains the more valuable information, the same phenomena was observed in [2]. So, in this case, proposing tags that the user already used in the past instead of tags other users attached to the resource, may provide a better chance to suggest the tags the user finally chose. Since, the most valuable tag information is contained in $\mathcal{R}_{\text{user}}$, but all the users in the test set are new and $\mathcal{R}_{\text{res}}$ yields low quality recommendations, unpromising labels are propagated, hence leading to poor results. FolkRank and PITF on the other hand, does not suffer from ill propagated labels and moreover, FolkRank can explore other users, resources and tags that are only indirectly connected to the test posts, which in this dataset at least, seems to yield a great benefit.

New User/Resource. Figure 5 shows the results for the second scenario on the BibSonomy dataset, where both the new user and new resource problem occurs. In general one can see that as the number of new resources increases the results deteriorate, but in both cases (1% and 10% of new resources) the semi-supervised relational methods outperform the state-of-the-art methods, while *PWA** achieves the best result. As expected *WA* does not perform very well since some test posts are only connected to other test posts, so that in some cases it cannot recommend any tags. The reason why *PWA** performs particularly good in this dataset, is because here the $\mathcal{R}_{\text{res}}$ relation already contains usefull tag assignments, and since this relation is the only training information available, it leads to the propagation of promising labels. Folkrank performs similar to *WA* and *WAOneShot* but is slightly outperformed by PITF. Again, for Last.fm (see figure 6) things are a little different. For both situations (1% and 10% of the resources are new), *PWA** still achieves the best result. In the 1% case *PWAOneShot* too, achieves better results than the state-of-the-art. *PWA** and *PWAOneShot* perform better because they use probabilities, which is not the case for *WAOneShot*. *WA* performs poorly, since, for many

[11] This baseline refers to the *most popular tags* of the folksonomy, i.e., it recommends, for any user $u \in U$ and any resource $r \in R$, the same set: $\hat{T}(u,r) := \text{argmax}^{n}_{t \in T}(|Y_t|)$.

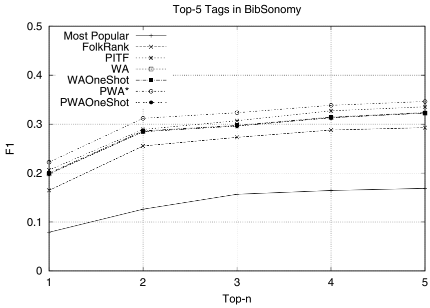

Fig. 3. New user problem BibSonomy

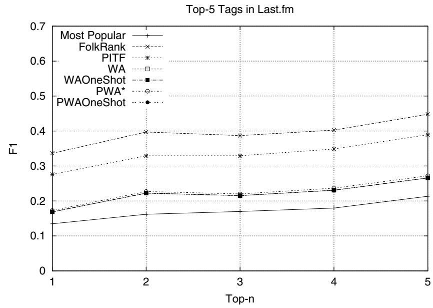

Fig. 4. New user problem Last.fm

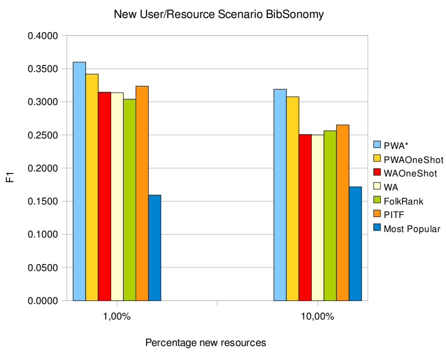

Fig. 5. New user/resource problem BibSonomy

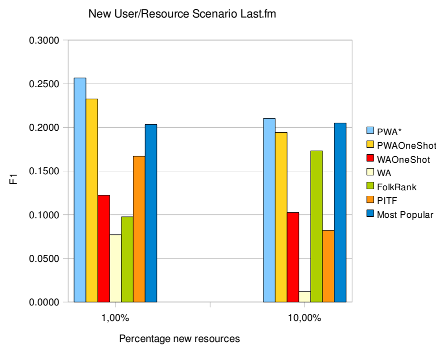

Fig. 6. New user/resource problem Last.fm

test posts no tags can be recommended since many test instances are only connected to test posts. FolkRank and PITF on the other hand, do not maintain the same performance achieved on the new user problem. This happens because the number of connected tag assignments to the test posts decreases proportionally to the number of new resources inside the posts of a new user, thus making it difficult to compute a good set of tags. In this particular situation, the simple most popular method performs surprisingly good, showing that in special cold-start cases, like this, it is a good alternative.

In general the re-estimation and propagation of labels in the post graph as well as the initialization of test posts seems to be the main reason for the good results of *PWA** (since *PWA** performed better than *PWAOneShot*). Furthermore, we see that the new user problem is easier to handle than situations where both new users and new resources occur, since the graph is less sparse, and therefore supervised methods work almost as well (or better) as semi-supervised methods.

6 Conclusions and Future Work

In this paper we have introduced an approach for tag recommendations that is particularly suitable for the cold-start problem. Our model is based on semi-supervised relational classification, that allows to exploit the structural information of untagged posts. We evaluated our approach against state-of-the-art methods in two real world datasets. We showed that semi-supervised relational methods which are based on label propagation are achieving very good results. In some special cases though, where the available training relations are of low quality, unpromising labels can be propagated thus deteriorating the results. In future work we want to investigate automatic ways of detecting more informative relations as well as other semi-supervised methods and new kinds of relations between the posts (e.g. content-based) for further improvement of cold-start related issues.

Acknowledgements

This work was funded by the X-Media project (www.x-media-project.org) sponsored by the European Commission as part of the Information Society Technologies (IST) programme under EC grant number IST-FP6-026978. The authors gratefully acknowledge the partial co-funding of their work through the European Commission FP7 project MyMedia (www.mymediaproject.org) under the grant agreement no. 215006. For your inquiries please contact info@mymediaproject.org.

References

1. Marinho, L.B., Preisach, C., Schmidt-Thieme, L.: Relational classification for personilized tag recommendation. In: ECML/PKDD Discovery Challenge 2009 at ECML/PKDD (2009)
2. Jaeschke, R., Marinho, L., Hotho, A., Schmidt-Thieme, L., Stumme, G.: Tag recommendations in social bookmarking systems. AI Communications, 231–247 (2008)
3. Hotho, A., Jaeschke, R., Schmitz, C., Stumme, G.: Information retrieval in folksonomies: Search and ranking. In: Sure, Y., Domingue, J. (eds.) ESWC 2006. LNCS, vol. 4011, pp. 411–426. Springer, Heidelberg (2006)
4. Rendle, S., Marinho, L.B., Nanopoulos, A., Schimdt-Thieme, L.: Learning optimal ranking with tensor factorization for tag recommendation. In: KDD 2009, pp. 727–736. ACM, New York (2009)
5. Rendle, S., Schmidt-Thieme, L.: Factor models for tag recommendation in bibsonomy. In: ECML/PKDD Discovery Challenge 2009 at ECML/PKDD (2009)
6. Rendle, S., Schmidt-Thieme, L.: Pairwise interaction tensor factorization. In: WSDM 2010 (2010)
7. Gunawardana, A., Meek, C.: Tied boltzmann machines for cold start recommendations. In: RecSys 2008, pp. 19–26. ACM, New York (2008)
8. Chakrabarti, S., Dom, B.E., Indyk, P.: Enhanced hypertext categorization using hyperlinks. In: Haas, L.M., Tiwary, A. (eds.) SIGMOD 1998, pp. 307–318. ACM Press, New York (1998)
9. Jensen, D., Neville, J., Gallagher, B.: Why collective inference improves relational classification. In: KDD 2004, pp. 593–598. ACM, New York (2004)
10. Sen, P., Namata, G., Bilgic, M., Getoor, L., Gallagher, B., Eliassi-Rad, T.: Collective classification in network data. AI Magazine (2008)
11. Lu, Q., Getoor, L.: Link-based classification using labeled and unlabeled data. In: Workshop on the continuum from labeled to unlabeled data, ICML 2003 (2003)
12. Preisach, C., Schmidt-Thieme, L.: Relational ensemble classification. In: ICDM 2006, pp. 499–509 (2006)
13. Macskassy, S.A., Provost, F.: Simple models and classification in networked data. In: CeDER Working Paper 03-04, Stern School of Business (2004)
14. Macskassy, S., Provost, F.: A simple relational classifier. In: Workshop on Multi-Relational Data Mining, KDD 2003, pp. 64–76 (2003)