Social Regularisation in a BPR-based Venue Recommendation System

Siwei Liu, Iadh Ounis, and Craig Macdonald

University of Glasgow, Glasgow , UK
s.liu.4@research.gla.ac.uk
iadh.ounis@glasgow.ac.uk, craig.macdonald@glasgow.ac.uk

Abstract. Venue Recommendation is a key application for Location-Based Social Networks (LBSNs) such as Yelp and Foursquare. Bayesian Personalised Ranking (BPR) is a popular pairwise recommendation technique that can be used to recommend a ranked list of venues to users. Social information such as friendship plays an important role in venue recommendation since it can alleviate the data sparsity problem induced by *cold-start* users with few check-ins and rating information. Indeed, introducing social information into BPR is a promising way to further improve the effectiveness of the underlying venue recommendation system. In this paper, we propose a novel Bayesian Personalised Ranking Social Regularisation (BPRSoReg) approach that uses social information as a regularisation method to enhance the performance of BPR. Experiments are conducted on a large-scale dataset from Yelp. The experimental results show that the BPRSoReg approach can improve over BPR by up to 54.0% in terms of mean reciprocal rank.

Keywords: Recommender systems · Matrix Factorisation · Social Networks

1 Introduction

Matrix Factorisation (MF), a class of collaborative filtering algorithms, has developed extensively since it has been used in recommender systems in the Netflix Prize competition in 2009 [2]. Based on its simple yet effective intuition that similar users are likely to appreciate the same items, Matrix Factorisation has been widely deployed in many online commercial platforms including Amazon, or for music recommendation at iTunes [4]. However, traditional matrix factorisation techniques suffer from the data sparsity problem since the explicit information e.g. the explicit rating of items by users is intrinsically rare. Indeed, the density of available ratings in commercial recommender systems is usually less than 1% [8]. Location-Based Social Networks (LBSNs) represent an important scenario that can benefit from recommender systems. In LBSNs, the core task is to recommend venues of interest to users. Usually, these venue recommendation systems suffer from a data sparsity problem. For example, the Round-13 Yelp challenge dataset used in our experiments has a rating density of only 0.002%. Due to this data sparsity problem, recommender systems face a challenge in extracting the user's interests and representations of each item/venue, often resulting in less accurate recommendations [3].

To address the sparsity problem, various approaches have been proposed in the literature to incorporate different sources of information including social information [4] and the textual content of comments [5]. Moreover, negative sampling [6] – as used by the Bayesian Personalised Ranking (BPR) [7] family of approaches – is another important technique that tackles the sparsity problem, based on a realistic assumption that most of the venues are unobserved or not particularly interesting to users. BPR [7] takes advantage of implicit feedback, which is more abundant compared with explicit feedback. Indeed, users' implicit feedback including clicks, viewing times, purchases, or check-ins, is much easier to collect since users do not need to express their interests explicitly. Moreover, implicit feedback can be tracked automatically online.

As mentioned above, in the current literature, there are two main approaches to address the sparsity problem: either to incorporate different sources of information into the traditional MF technique or to leverage implicit feedback in BPR. Since BPR is based on matrix factorisation, in this paper we aim to investigate whether it is possible to integrate both approaches in a single recommender system. Our objective in this paper is therefore to propose a ranking recommender system that can incorporate social information in addition to implicit feedback. Hence, we propose the combination of the social regularisation, originally proposed for MF, into the BPR pairwise recommendation approach; Experiments on a large dataset from the Yelp LBSN demonstrate the benefit of our approach.

The remainder of this paper is structured as follows: In Section 2, we firstly describe the BPR rank-based recommender system then details of how Social Regularisation (SoReg) can incorporate social information i.e. friendship information. Since both BPR and SoReg are based on matrix factorisation, in Section 3, we show how to integrate SoReg and BPR to achieve our objective. Experimental Setup and Results are described in Sections 4 & 5. Concluding remarks follow in Section 6.

2 Related Work

Bayesian Personalised Ranking: While matrix factorisation is designed for the rating prediction task, it is not directly optimised for the *ranking* of venues. In venue recommendation, the core task is to generate a personalised ranking for users and hence users will not have to scan through all suggested venues. Instead, they usually focus on the top ranked venues. Therefore, in many real venues recommendation applications, giving users the best top ranked venues is more useful than making rating predictions. Building on matrix factorisation, BPR [7] is a pairwise ranking method that has been proposed to generate a personalised ranked list of venues for users. BPR takes advantage of the latent representation of users and venues generated by MF using venue pairs as training data and optimising for correctly ranking venue pairs instead of focusing on correctly predicting the ratings. Its optimisation criterion is based on the assumption that a user u prefers a venue v that has been viewed by this user over all other non-observed venues. The maximum posterior estimator is computed as follows:

$$BPR - OPT := \ln p(\theta| >_u) = \sum_{(u,i,j) \in D_s} \ln \sigma(\hat{x}_{uij}) - \lambda_\theta \|\theta\|^2 \tag{1}$$

where θ represents the parameter vector of an arbitrary model class, σ is the logistic sigmoid, $\hat{x}_{uij}$ is an arbitrary real-valued function of the model parameter vector θ that captures the special relationship between user u, venue i and venue j; λ_θ is a model-specific regularisation parameter; $\|.\|$ denotes the Frobenius norm; and $\ln(\cdot)$ is the natural logarithm. Therefore $p(\theta| >_u)$ is the posterior probability that θ produces the correct personalised ranking for user u, which BPR aims to maximise through stochastic gradient descent.

Social Regularisation: Social regularisation (SoReg) is a recommendation model proposed by Ma et al. [4]. It utilises social information to improve the performance of traditional matrix factorisation recommender systems by introducing a social regularisation term to constrain the matrix factorisation objective function. The traditional objective function of matrix factorisation is:

$$L(U,V) = \min_{U,V} \frac{1}{2} \sum_{m}^{i=1} \sum_{n}^{j=1} I_{i,j} * (R_{i,j} - U_i^T V_j)^2 \tag{2}$$

where $I_{i,j}$ is an indicator variable that is 1 if user i rated venue j, otherwise 0. A regularisation term is added to Equation (2) to avoid over-fitting, as follows:

$$\underset{MF}{L(U,V)} = L(U,V) + \frac{\lambda}{2}(\|U\|_F^2 + \|V\|_F^2) \tag{3}$$

where λ is a parameter controlling the amount of regularisation.

Users' decisions may be influenced by their friends. For example users often recommend venues to their friends. It is also common for users to ask their friends for suggestions especially when they visit unfamiliar places. Based on the assumption that users are likely to be influenced by their friends, the SoReg model introduces social information i.e. friendship information, as a regularisation term to minimise the distance between the latent representations of target user U_i and his/her friends, U_f, which further modifies Equation (3) as:

$$\underset{SoReg}{L(U,V)} = \underset{MF}{L(U,V)} + \frac{\alpha}{2} \sum_{i=1}^{m} \sum_{f \in \mathcal{F}(i)} pcc(i,f) \|U_i - U_f\|_F^2 \tag{4}$$

where $\mathcal{F}(.)$ is the set of friends of user i and α is a parameter that controls social influences; *pcc()* estimates the similarity between the ratings of two users using the Pearson Correlation Coefficient (PCC):

$$pcc(i,f) = \frac{\sum\limits_{j \in V_r(i) \cap V_r(f)()} (R_{ij} - \overline{R_i})(R_{fj} - \overline{R_f})}{\sqrt{\sum\limits_{j \in V_r(i) \cap V_r(f)()} (R_{ij} - \overline{R_i})^2} . \sqrt{\sum\limits_{j \in V_r(i) \cap V_r(f)()} (R_{fj} - \overline{R_f})^2}} \tag{5}$$

where $V_r(i)$ are the venues that user i has rated and $\overline{R_i}$ is their average rating.

In Equation (4), the loss function of the social regularisation model is the sum of the loss function of MF and the social regularisation term. The Frobenius norm of the difference between the latent representations of the user and his/her friends is then multiplied by the PCC, which scales according to how similar the user and his/her friends are, based on their rating history.

The functionality of SoReg is to ensure that the user-friend pair who share similar interests are predicted to rate similarly other venues. The more similar a user is to his/her friend, the greater the correlation. By applying the SoReg model, not only a user and his/her friends become closer based on their similarity level in the latent space, but so do the friends' friends, who become closer at the same time. This means that not only a user is recommended venues by his/her friends, but also by his/her friends' friends, although indirectly. The range of Pearson Correlation Coefficient is [-1,1]. We apply a mapping function $f(x) = (x+1)/2$ to bound the range of correlations to [0,1].

3 Pairwise Social Regularisation

While BPR is based on matrix factorisation, the objective function is Equation (1) introduced above. It can be noted in Equation (4) that the objective function of SoReg is the objective function of MF augmented by the social regularisation. Therefore, to combine BPR with SoReg, we fit Equation (1) into Equation (4) to replace the $\underset{MF}{L(U,V)}$ part, thereby making the objective function of our proposed BPRSoReg approach as follows:

$$\underset{BPRSoReg}{L(U,V)} = \sum_{(u,i,j)\in D_s} \ln \sigma(\hat{x}_{uij}) - \lambda_\theta \left\|\theta\right\|^2 + \frac{\alpha}{2}\sum_{i=1}^{m}\sum_{f\in\mathcal{F}(i)} pcc(i,f)\left\|U_i - U_f\right\|_F^2 \tag{6}$$

Using Equation (6), a pairwise recommender system that takes into account social information can therefore be established.

4 Dataset and Experiments

We use the Yelp challenge Round-13 dataset to conduct all our experiments. This dataset has 6,685,900 rating records, 1,637,138 unique users and 192,606 unique venues. Hence, the data sparsity is 0.002120%. Of *all* these users, there are 122,824 *regular* users (users with $\geq$ 10 rating records) and 1,514,314 *cold-start* users (user with $<$ 10 rating records). On average, the *regular* users have 36.1 friends, while the *cold-start* users have 6.8 friends. While no filtering is applied on the dataset, we analyse separately the *regular* users' and *cold-start* users' results, to determine how differently our proposed model performs on different types of users.

We use Spotlight, a Python recommendation platform that includes an implementation of a BPR implicit factorisation, to conduct our experiments[1]. Following previous work in [1] and [4], we set the latent dimension d to 10 and λ to

[1] https://github.com/maciejkula/spotlight

0.001. The dataset is randomly split into 80% for training and 20% for testing the ranking results given by the model.

Following the existing literature in recommender systems, to measure the effectiveness of recommendations, mean reciprocal rank (MRR) is used where the reciprocal rank is the multiplicative inverse of the ranked venues for a specific user. The average value is taken for all users in the dataset as an estimation of the performance of the system. MRR is computed as follows:

$$MRR = \frac{1}{U} \sum_{i=1}^{|U|} \frac{1}{rank_i} \tag{7}$$

where $rank_i$ is the rank position of the first relevant venue retrieved for user i.

In our experiments, we vary the social parameter α (see Equation (4)), which controls the influence of the social regularisation, from 10^{-7} to 1, multiplying α by 10 at each step. In the Yelp challenge Round-13 dataset, 45.4% of users have a *friend-list* through which we can find their friends' information according to the user-id(s) provided. With this information, the Pearson Correlation Coefficient between users is pre-computed before training commences, to increase training efficiency. We compare the effectiveness – in terms of MRR – of BPRSoReg with different amounts of social regularisation to the BPR baseline. Finally, significance tests are conducted using a paired t-test with $p < 0.01$.

5 Results

Table 1 reports the effectiveness of BPRSoReg and BPR in terms of MRR for different groups of users. From the left hand group of Table 1, it can been seen that the BPRSoReg model can consistently outperform BPR in terms of MRR across *all* users - indeed, the best performance occurs when α=10^{-6}, significantly improving MRR by 54.0%. This demonstrates that social regularisation can greatly improve the effectiveness of BPR. As α increases, the model's effectiveness starts to saturate, which means that the social regularisation is penalising models that produce results that are too different from each given user's friends. Conversely, for a very small α, social regularisation gives a very small penalty to models where users differ widely from their friends, and the performance naturally returns to that of the baseline. Such a variance can be observed clearly in Figure 1, which shows how MRR varies with α.

Next, the centre and right-hand groups of Table 1 report the obtained results for the *regular* and *cold-start* users, respectively. From the table, we observe that introducing social information leads to a consistent improvement for *cold-start* users but for *regular* users the ranking performance drops markedly when $\alpha \geq 10^{-4}$. This might be because for *regular* users, their check-ins are enough for the BPR model to capture their interests so the model does not benefit from receiving additional regularisation for those users' friends. However, the *cold-start* users have very few check-ins, hence the model will benefit when additional social information is available. As noted in Section 4, *regular* users have, on average, 6 times as many friends are *cold-start* users. While the *cold-start* users and *regular* users share the same social parameter α during training, it is possible that the

Table 1. Ranking performances in terms of MRR of BPRSoReg and BPR, while varying α for *all* users, as well as *regular* users ($\geq$ 10 check-ins) and *cold-start* users separately. Improvements (Δ %) are calculated with respect to the BPR baseline. Best result is highlighted.

	All Users		Regular Users		Cold-start Users	
α	MRR	Δ	MRR	Δ %	MRR	Δ %
BPR	0.00163	-	0.00111	-	0.00167	-
10^{-7}	0.00193*	(18.4%)	0.00114	(2.70%)	0.00199	(19.2%)
10^{-6}	**0.00251**	**(54.0)**	**0.00129**	**(16.2%)**	**0.00261**	**(56.3%)**
10^{-5}	0.00232*	(42.3)	0.00112	(0.9%)	0.00241	(44.3%)
10^{-4}	0.00190*	(16.6)	0.00078	(-29.7%)	0.00199	(19.2%)
10^{-3}	0.00201*	(23.3)	0.00076	(-31.5%)	0.00211	(26.3%)
10^{-2}	0.00195*	(19.6)	0.00066	(-40.5%)	0.00206	(23.3%)
10^{-1}	0.00188*	(15.3)	0.00065	(-41.4%)	0.00198	(18.6%)
1	0.00188*	(15.3)	0.00069	(-37.8%)	0.00197	(18.0%)

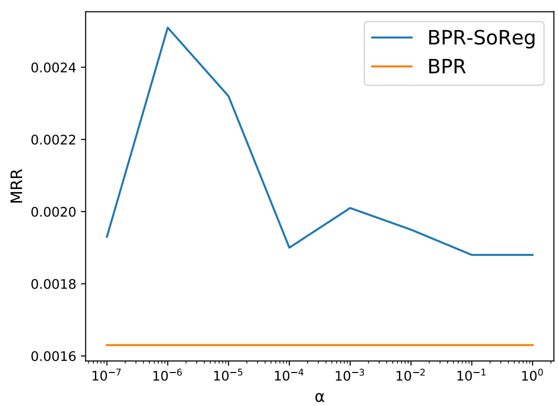

Fig. 1. Plot of prediction performances in terms of MRR of BPRSoReg with α varying from 10^{-7} to 1 in comparison to the BPR baseline.

large numbers of friends mean that forcing the *regular* users' latent factors to be similar to all of their friends is too aggressive in penalising models that would otherwise be effective. Therefore, the best effectiveness is obtained when α is small.

6 Conclusions

In this paper, we explored whether social information can be leveraged to improve Bayesian Personalised Ranking (BPR). We proposed BPRSoReg, which incorporates social information as a regularisation for BPR. Our experiments on a large LBSN dataset demonstrate that we can significantly enhance the effectiveness of BPR. Furthermore, our experiments show that BPRSoReg particularly benefits *cold-start* users. In the future, we plan to add more ranking metrics including Hit Ratio and Normalised Discounted Cumulative Gain. In addition, we plan to integrate social information to different frameworks of recommender systems such as those based on neural networks to examine the generalisation of social information in recommender systems.

Acknowledgements. We thank ACM SIGIR for providing the first author with an FDIA 2019 travel grant.

References

1. Guo, G., Zhang, J., Yorke-Smith, N.: TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In: Proceedings of AAAI. pp. 123–129 (2015)
2. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
3. Loni, B., Pagano, R., Larson, M., Hanjalic, A.: Bayesian personalized ranking with multi-channel user feedback. In: Proceedings of RecSys. pp. 361–364 (2016)
4. Ma, H., Zhou, D., Liu, C., Lyu, M.R., King, I.: Recommender systems with social regularization. In: Proceedings of WSDM. pp. 287–296 (2011)
5. Manotumruksa, J., Macdonald, C., Ounis, I.: Regularising factorised models for venue recommendation using friends and their comments. In: Proceedings of CIKM. pp. 1981–1984 (2016)
6. Pan, W., Zhong, H., Xu, C., Ming, Z.: Adaptive bayesian personalized ranking for heterogeneous implicit feedbacks. Know.-Based Syst. **73**(1), 173–180 (2015)
7. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: BPR: Bayesian personalized ranking from implicit feedback. In: Proceedings of UAI. pp. 452–461 (2009)
8. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of WWW. pp. 285–295 (2001)