Hybrid Trust-Aware Model for Personalized Top-N Recommendation

Arpit Merchant
Center for Exact Humanities
IIIT, Hyderabad
arpit.merchant@research.iiit.ac.in

Navjyoti Singh
Center for Exact Humanities
IIIT, Hyderabad
navjyoti@iiit.ac.in

ABSTRACT

Due to the large quantity and diversity of content being easily available to users, recommender systems (RS) have become an integral part of nearly every online system. They allow users to resolve the information overload problem by proactively generating high-quality personalized recommendations. Trust metrics help leverage preferences of similar users and have led to improved predictive accuracy which is why they have become an important consideration in the design of RSs. We argue that there are additional aspects of trust as a human notion, that can be integrated with collaborative filtering techniques to suggest to users items that they might like. In this paper, we present an approach for the top-N recommendation task that computes prediction scores for items as a user specific combination of global and local trust models to capture differences in preferences. Our experiments show that the proposed method improves upon the standard trust model and outperforms competing top-N recommendation approaches on real world data by upto 19%.

Keywords

Recommender Systems, Trust, Collaborative Filtering

1. INTRODUCTION

In recent years, more content is being produced every hour than can be possibly consumed by any single individual. While this means that there is a really large variety of items for viewers to choose from, the sheer amount makes it extremely hard for them to easily find something that they would actually like watching. Recommender Systems (RS) have proved to be an important response to this information overload problem by helping to locate the sought content quickly. Top-N RSs provide users with a ranked list of N relevant items to encourage views and have thus become ubiquitous across domains.

It is widely established that Collaborative Filtering (CF) based strategies have been successfully deployed in solving the Top-N recommendation task ([12], [11]). Collaborative filtering is based on the notion that we rely on the community to provide suggestions. It is a method of processing or filtering information obtained through collaboration between various agents, information sources and other metadata. Neighborhood methods and latent space models represent users and items in a common latent feature space and identify similar users and items, and are common approaches used for collaborative filtering [3].

However, such RSs share an inherent drawback in choosing recommendation partners based on similar ratings histories with the active user and/or recommendation items with similar content features to the active item. Apart from similarity, trustworthiness of users is an important consideration in providing a high quality set of recommendations. Indeed trust-based RSs have proven to improve predictive accuracy of standard CF frameworks [8]. However, the previous trust-based models are limiting in that they estimate only a single model for all users based on a narrow definition of user trust. We argue, that there are additional aspects to trust that can further enhance the performance of these RSs. Specifically, *personal trust* that arises from the unique features and attributes of the individual users, *social trust* which comes from contexts, social groups and communal habits, and *functional trust* which is based on the role and method of functioning of the underlying system [9].

In this paper, we are interested in combining a holistic view of trust with collaborative filtering to improve the ability of the RS to make accurate and trustworthy predictions. We propose a Top-N recommendation approach that extends the sparse linear method [10] by combining context-aware global and local trust models in the personalized setting. To do this, we leverage the taxonomic framework in [9] to collect personal, social and functional trust information and solve an optimization problem following a tensor factorization process to strengthen the coupling between user and item features. Our experimental evaluation of the proposed method on a TED[1] dataset shows that it has upto 19% higher accuracy as compared to competing methods along with better diversity and popularity which are desirable for user satisfaction.

The rest of the chapter is organized as follows: Section 2 presents the related work. Section 3 describes the proposed method. Section 4 presents the experimental evaluation and Section 5 presents the performance analysis. Finally, Section 6 provides concluding remarks.

[1] https://www.ted.com/

CODS 2016, IIT Madras, India
© 2016 ACM. ISBN 123-4567-24-567/08/06...$15.00
DOI: 10.475/123_4

2. RELATED WORK

Trust has been defined across various categories and there are numerous methods on how to view and measure trust ([15], [7]). Trust-based systems are generating an increasing amount of interest and have made several notable advancements to the state-of-the-art in social network analysis, community detection and recommender systems [13]. Here we present a few notable works in the field.

Adali et al [1] measured the behavioural trust in social networks based on conversations and propagation of information. Trifunavic [14] proposed a trust model for opportunistic networks using explicit and implicit social trust. Based on the taxonomic framework by [9], we are interested in three categories of trust; Context-specific Social trust wherein a user A might trust user B when it comes to some specific situation, but not necessarily in some other. System-based Functional trust which describes the trust in the system as a whole and its interpersonal elements. And lastly the Personal trust by the user in the capabilities of the former two categories to provide quality recommendations.

O'Donovan and Smyth [11] devised algorithms for computing Profile-Level and Item-Level Trust based on trust statements provided by users about other users. Massa and Avesani [8] showed how trust weight, computed through a local trust metric obtained from trust values, can be integrated with collaborative filtering to improve accuracy. In her PhD thesis, Golbeck [5] studied online social networks and defined a trust metric called TidalTrust, that in a breadth-first manner computed trust propagation based on how much a user trusted other users' film ratings. While their work is based on explicit trust statements and/or values provided by users, in our work we derive trust values implicitly from comments and ratings (of the TED. It is interesting to note that our findings are similar to those obtained by aforementioned works and are reported in Section 5.

Pappas et al [12] compare Content Based (CB), Collaborative Filtering and hybrid methods based on semantic vector spaces to recommend TED talks in a personalized setting. They show that CB methods are outperformed by the CF methods, but hybrid methods do well too. They leave the impact of comments and user behaviours for future study. We utilize local and global trust metrics at the user and item level that outperform competing approaches to solving the personalized Top-N recommendation task.

3. PROPOSED APPROACH

3.1 Formalization

The goal is to recommend "talks" to the user based on knowledge obtained from the ones he/she has already viewed, and other trusted members of the community. Let $U = \{u_1, u_2, \ldots, u_n\}$ be a set of users and $T = \{t_1, t_2, \ldots, t_m\}$ be a set of talks. Let R be the *user-talk favourites* matrix of size $n \times m$ that shows the rating provided by users for talks that they have viewed and liked. If a user u marks a talk t as "favourite", then the corresponding entry r_{ut} of the matrix is 1 and 0 otherwise for the unrated talks. Let $\Re_u$ denote the set of items marked as favourite by user u.

3.2 Estimating Contextual Social Trust

Social trust between users depends on the nature of interactions between them and the context. For instance, Alice may trust Bob when it comes to recommending talks related to (say) Technology but may not trust Bob as much when it comes to talks related to (say) Psychology. To compute social trust between users u and v in the context of a particular talk t, we carry out sentiment analysis of their comments on t. We denote the sparse *user-user-talk* tensor $Y \in \mathcal{Y}^{n \times n \times m}$ by:

$$Y^t_{u,v} = \frac{p^t_{u,v} - n^t_{u,v}}{p^t_{u,v} + n^t_{u,v}} \tag{1}$$

where $p^t_{u,v}$ and $n^t_{u,v}$ are the number of comments of similar and opposite relative polarity respectively, made by u and v on t. Similar to Hu et al [6], we transform the counts $Y^t_{u,v}$ into *contextual social trust values* $S = (S^t_{u,v})_{u,v,t}$ given by the equation $S^t_{u,v} = 1 + \alpha \log(1 + Y^t_{u,v}/\epsilon)$. α and ϵ are scaling parameters which are empirically determined.

The observed values in S are formed by the information provided by the users and estimating the unknown values boils down to a Tensor Completion (TC) problem. Analogous to Matrix Factorization (MF), the aim is to factorize the tensor into three matrices $U \in \mathbb{R}^{n \times d_U}$, $V \in \mathbb{R}^{n \times d_V}$, $A \in \mathbb{R}^{n \times d_A}$ and a central tensor $J \in \mathbb{R}^{d_U \times d_V \times d_A}$ such that $F = J \otimes U \otimes V \otimes A$ approximates S. That is, it minimizes the loss function between the observed and estimated values. We define the loss function as:

$$L(F, S) := \frac{1}{||J||_1} \sum_{u,v,t} D^t_{u,v} L(F^t_{u,v}, S^t_{u,v}) \tag{2}$$

where $D \in \{0,1\}^{n \times n \times m}$ is a binary tensor taking the value 1 when $S^t_{u,v}$ is observed and L is the pointwise loss function given by $L(F^t_{u,v}, S^t_{u,v}) = \frac{1}{2}(F^t_{u,v} - S^t_{u,v})^2$. We also add the Frobenius norm as a regularization term for improving the generalization performance and to prevent overfitting and it is given by:

$$\Omega(U, V, A) := \frac{1}{2}[\lambda_U ||U||^2_{\mathrm{Frob}} + \lambda_V ||V||^2_{\mathrm{Frob}} + \lambda_A ||A||^2_{\mathrm{Frob}}] \tag{3}$$

and similarly also for tensor J. The regularization weights λ_U, λ_V and λ_A control the sparsity of the corresponding matrices U, V and A. Thus combining Equations (2) and (3), the overall minimization problem becomes:

$$H(U, V, A, S) := L(F, S) + \Omega(U, V, A) + \Omega(J) \tag{4}$$

The optimization problem of Equation (6) can be solved using HOSVD-decomposition [4] and stochastic gradient descent.

3.3 Computing Global Trust

To compute the global trust value of a talk t, we first construct a network graph $G = (V, E, W)$ of the underlying system, where each vertex $v \in V$ is a talk and there exists an edge between two talks if they share at least one tag in common. The weights assigned to each edge represent the total number of tags that two talks have in common. We now compute, for each talk t:

$$b(t) = \frac{1}{|V|(|V| - 1)} \sum_{\substack{(i,j) \in E \\ i \neq j}} \frac{\sigma_{ij}(t)}{\sigma_{ij}} \tag{5}$$

where σ_{ij} is the number of shortest paths from i to j and $\sigma_{ij}(t)$ is the number of shortest paths from i to j that are internal to t. $b(t)$ roughly denotes how easy it is to be able to reach from talk t to some talk t'. That is, (5) estimates the influence of t in the graph. We also calculate for each t, its PageRank score *c(t)* given by:

$$c(t) = \alpha \sum_{j} d_{jt} \frac{c(j)}{L(j)} + \frac{1-\alpha}{n} \quad (6)$$

where $L(j) = \sum_t d_{jt}$ is the number of neighbours of node j. The motivation behind it is that it provides an estimate of how important the node is and what the probability of visiting the node in a random walk is.

Combining the global importance and influence of talk t in the network graph, we define the global trust of t as:

$$G(t) = \frac{2 \times b(t) \times c(t)}{b(t) + c(t)} \quad (7)$$

3.4 Hybrid Trust Recommendation

In this subsection, we present our Hybrid Trust-Aware method for personalized top-N recommendation. We compute the social and global trust values as described in Subsections 3.2 and 3.3 respectively. Putting it all together, the predicted rating $\tilde{r}_{ut}$ for user u for talk t will be estimated by:

$$\tilde{r}_{ut} = \sum_{l \in \Re_u} g_u G(l) \times s_{lt} + \sum_{v \in \mathcal{T}_u} (1 - g_u) F^t_{u,v} \times \arg\min_{l \in \Re_v} s_{lt} \quad (8)$$

where g_u is the personal trust for user u which controls the interplay between the social and global terms. Its value lies in the interval $[0, 1]$ where 0 denotes that only the social part plays a role in the prediction and 1 denotes that only the global part is used. g_u is initially set to 0.5 to ensure equal weightage to both parts and it is updated during training to reflect the behaviour of the user. $G(l)$ represents the global trust for talk l as obtained from Equation (7) and $F^t_{u,v}$ represents the social trust between users u and v for talk t as seen in Section 3.2. And lastly, s_{lt} represents the item-item Jaccard similarity coefficient between the lth item rated by u and the target item t.

We then compute the predicted rating $\tilde{r}_{ut}$ for every unrated talk t for a user u according to Equation (8), sort the values and recommend the top-N talks with the highest ratings to the user.

4. EXPERIMENTAL EVALUATION

4.1 Dataset

We evaluate our method on a real-world TED dataset[2] collected by Pappas et al [12] which they have released under the same Creative Commons license as TED. The dataset is based on a snapshot of the TED website taken in September, 2012. It includes 1.2k talks, 69k users, over 100k instances of talks marked as favourites and over 200k comments. Each talk is characterized by a unique ID, title, number of views, name of the speaker, etc. Additionally, metadata such as related tags (e.g. "Technology", "Science") and related talks that have been manually assigned (by the TED staff) for each talk are also made available.

4.2 Evaluation Methodology

For the top-N personalized recommendation task, for each user u, we randomly select 80% of his/her favourites to place into the training set (M) and the remaining 20% form the test (E) set. The set of users is selected on the basis of the minimal number of favourites available. After these splits have been constructed, the model is trained on the training set and the recommendation quality is measured on the test set. We use the widely popular Precision and Recall measures for the top-N list.

$$\text{Precision}(E, N) = \frac{|\text{Top}(u, N) \cap \{t|(u, t) \in E\}|}{N} \quad (9)$$

$$\text{Recall}(E, N) = \frac{|\text{Top}(u, N) \cap \{t|(u, t) \in E\}|}{|\{t|(u, t) \in E\}|} \quad (10)$$

We repeat the experiments 10 times with new training and test samples and report average precision and recall values.

Additionally, we carry out a more subjective evaluation of user satisfiability using the following metrics to show the relative advantages and weaknesses of our model:

1. *Popularity* - computes the number of other users who have viewed the talk.
2. *Newness* - computes the average of the release dates of the recommended talks.
3. *Diversity* - measures the distribution of themes covered by the talks.

Furthermore, we compare our model with three baseline approaches namely, Random wherein talks recommended to the user are chosen uniformly at random from all available talks, Content Based (CB) algorithm that given a similarity function recommends N most similar items to those in the user's favourites list, and the standard Resnick prediction method. We also compare with related recommender methods from literature; Profile-Level Trust in [11], TaRS [8] and ESA [12].

4.3 Results

We construct the network graph $G = (V, E)$ defined in Section 3.3 for the TED dataset with $|V| = 1203$ and $|E| = 306977$. Its characteristic path length is 1.55 which is small compared to the size of the graph and its diameter is 3 which indicates that on average, it is possible to reach any talk from any talk in relatively few steps. The degree distribution of the graph yields an exponent of $\lambda = 1.03$ with an estimated $x_{\min} = 0.0008$. Since the distribution is exponential, the value of λ is invariant to the size of the network. And the average clustering coefficient is 0.702. Taken together, we can conclude that the graph is scale free and exhibits small-world properties [2].

Table 1 shows the mean average precision and recall values for the Hybrid model and the other baseline and competing methods for the personalized recommendation task. We only measure precision and recall for top-10 and top-20 talks since most users do not scroll beyond these many recommendations. As can be seen from Table 1, the Hybrid method outperforms the baseline methods as well as the competing methods in terms of accuracy. While CB is able to provide talks with similar content, it fails to identify those talks that are not similar that a user would like. On the other hand, while the CF-based Resnick method is able to suggest talks viewed as popular in the community, trusted users are found to better predictors and so it is worth the extra effort to find them. The Hybrid model conveys to the user that the talks were recommended to him/her by these specific users who have successfully recommended talks to other users based on

[2]https://www.idiap.ch/dataset/ted

Table 1: Accuracy of three baseline approaches, competing methods and the Hybrid model proposed in this work. These models are evaluated for the top-10 and top-20 recommendation tasks.

Model	Random	CB	Resnick	TaRS	ESA	Hybrid
Precision@10	0.028	0.026	0.043	0.085	0.132	**0.158**
Recall@10	0.019	0.022	0.041	0.078	0.036	0.047
Precision@20	0.016	0.019	0.032	0.117	0.060	0.172
Recall@20	0.026	0.043	0.078	0.082	0.076	0.065

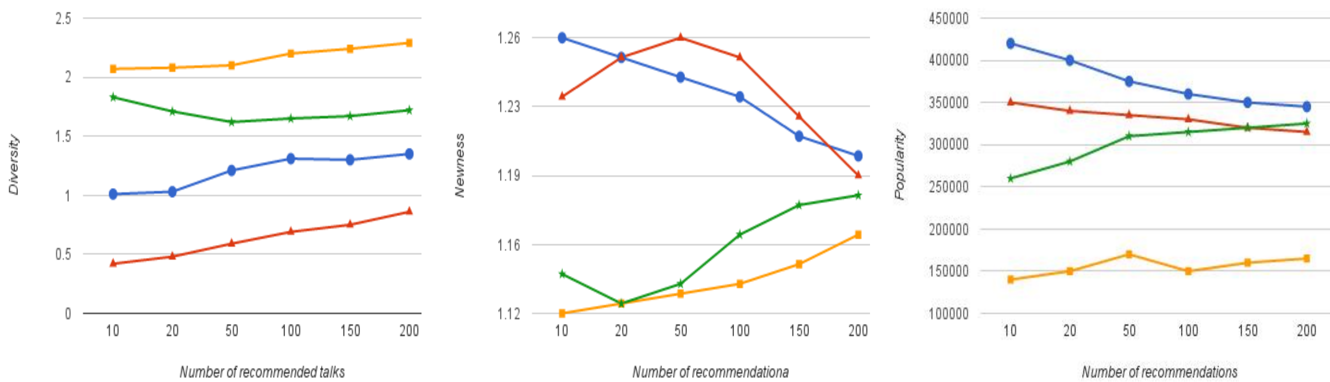

Figure 1: Subjective evaluation of Global Trust (Blue), Social Trust (Red), Random (Yellow) and Hybrid (Green) models based on popularity, diversity and freshness criteria to gauge user satisfaction.

their trustworthiness which in turn causes the user to trust future recommendations.

Figure 1 compares the popularity, newness and diversity of the recommendations and it is a more subjective evaluation which we argue is equally important as accuracy. The Hybrid model scores high on diversity and mediocre on newness and popularity. We compare this with models based on Global trust only, Social trust only and the Random method. The Random method outperforms the Hybrid model on diversity as can be expected but the Hybrid model performs better than the Global and Social trust metric only models. This suggests that Hybrid can discover talks to recommend even if they are not very popular or newly added.

5. CONCLUSIONS

In this paper, we address the personalized top-N recommendation problem. The model we propose combines a three-dimensional (personal, social and functional) view of trust with collaborative filtering to recommend trusted content suggested by trusted fellow users. Experimental results depict that the model outperforms standard CB and CF methods as well as competing trust-based approaches. While our model is evaluated only on a dataset of TED talks, we believe the results are promising enough to be applied to other domains as well.

6. REFERENCES

[1] S. Adali, R. Escriva, M. K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. K. Szymanski, W. Wallace, G. Williams, et al. Measuring behavioral trust in social networks. In *Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on*, pages 150–152. IEEE, 2010.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[3] E. Christakopoulou and G. Karypis. Local item-item models for top-n recommendation. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 67–74. ACM, 2016.

[4] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[5] J. Golbeck, B. Parsia, and J. Hendler. *Trust networks on the semantic web*. Springer, 2003.

[6] Y. Hu, Y. Koren, and C. Volinsky. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE International Conference on Data Mining*, pages 263–272. Ieee, 2008.

[7] S. P. Marsh. Formalising trust as a computational concept. 1994.

[8] P. Massa and P. Avesani. Trust-aware recommender systems. In *Proceedings of the 2007 ACM conference on Recommender systems*, pages 17–24. ACM, 2007.

[9] A. Merchant, T. Jha, and N. Singh. The use of trust in social machines. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 787–792. International World Wide Web Conferences Steering Committee, 2016.

[10] X. Ning and G. Karypis. Slim: Sparse linear methods for top-n recommender systems. In *2011 IEEE 11th International Conference on Data Mining*, pages 497–506. IEEE, 2011.

[11] J. O'Donovan and B. Smyth. Trust in recommender systems. In *Proceedings of the 10th international conference on Intelligent user interfaces*, pages 167–174. ACM, 2005.

[12] N. Pappas and A. Popescu-Belis. Combining content with user preferences for ted lecture recommendation. In *Content-Based Multimedia Indexing (CBMI), 2013 11th International Workshop on*, pages 47–52. IEEE, 2013.

[13] W. Sherchan, S. Nepal, and C. Paris. A survey of trust in social networks. *ACM Computing Surveys (CSUR)*, 45(4):47, 2013.

[14] S. Trifunovic, F. Legendre, and C. Anastasiades. Social trust in opportunistic networks. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pages 1–6. IEEE, 2010.

[15] C.-N. Ziegler. *Towards decentralized recommender systems*. PhD thesis, Institut für Informatik, 2005.