Inferring Shared Tastes with Network Topic Models

Anonymous Author(s)
Affiliation
Address
email

Abstract

It is widely assumed that people tend to gather in groups of shared interests, where such interests drive friendships and vice versa. Thanks to online social networking platforms, information about a users's friends as well the items he is interested in are available, but represent only an incomplete picture. We study probabilistic network topic models that distill common shared interests of friends from this data. So far, a popular choice is based on the mixed-membership stochastic blockmodel which draws inference on the absence of edges. In this paper we give theoretic and empirical evidence that absence of friendship does not refer to difference in taste. Rather, we present the shared taste model which is agnostic towards absent edges. The model's success over blockmodels is demonstrated on data sets from LibraryThing, Zune Social, and CiteSeer.

The shared taste model has many practical applications such as improving content subscription, visualizing the structure of contacts, or connecting people who otherwise would not have someone to share their interests with.

1 Introduction

Many online community platforms, such as Facebook, LibraryThing, Zune Social, last.fm, flickr, and CiteULike, store data about users, friend relationships between users, and items users interact with. Depending on the usage scenario, items may be status updates, books, songs, pictures, or scientific publications, respectively. It is widely assumed that people tend to gather in groups of shared interests, where such interests drive friendships and friendships drive interests.

Community platforms typically allow users to subscribe to items of their friends. This functionality may be unsatisfactory if the friend has diverse interests (e.g., likes rock and jazz), of which only some are shared with the subscribing user (who likes jazz and classical music). In this example, the common taste is jazz. The user will be frustrated if too many non-matching items are suggested. For instance, the user would expect to hear jazz and might be annoyed by rock songs. Re-weighting the friend's item list to match the common taste will improve the user experience. Unfortunately, users do not explicitly state the shared taste of each friendship.

This work focuses on unsupervised inference of shared tastes $\mathcal{T}$ given a content-enriched social network. A shared taste captures a common interest of two users. We use the terms taste, interest, and topic synonymously.

Definition (Content-enriched social network) $(\mathcal{N}, \mathcal{E}, \mathcal{C}, \mathcal{V})$: Let nodes $n \in \mathcal{N}$ in the social network refer to users and let edges $\mathcal{E}$ refer to online friendships between users. The edges are inherently undirected but can also be interpreted as bidirectional if required. In addition, each node is associated with a set of items $\mathcal{C}(n)$ from a common item vocabulary $\mathcal{V}$, e.g., tags or songs.

Problem: Given the content-enriched social network as input data, the goal is to learn a set of shared topics $\mathcal{T}$ that explain the common interests of the friends based on their item interaction and friendships with other users.

Recently, extensions of latent Dirichlet allocation [1] have been applied to different kinds of network data and studied in a broad range of prediction problems. For instance, the author topic model [2] is designed for a setting where one item set is shared among several nodes. Cohn, Hofman, and Erosheva et al. [3, 4] suggest models that identify groups of nodes sharing items and friends, but do not model topics shared via an edge. Mei et al. [5] discriminatively learn a topic model with a graph Laplacian as a regularizer, assuming that connected users should have the same topics, aiming at a scenario where the graph clusters into coherent components.

A popular choice for combining topic models with observed graph structure is the mixed-membership stochastic blockmodel [6]. Stochastic blockmodels associate each node with a topic mixture and explain the presence and absence of edges from the compatibility of topics. Pairwise Link-LDA [7] and the relational topic model [8] extend blockmodels to model node contents $\mathcal{C}$ together with the network structure $(\mathcal{N}, \mathcal{E})$.

The citation influence model [9] explains contents of nodes by borrowing topics of adjacent nodes, yielding topic mixtures that are influenced by several nodes simultaneously. Originally it was designed for directed acyclic graphs. In this paper we transfer it to undirected graphs with cycles for the identification of shared tastes.

The first main contribution of this work is the shared taste model[1] which associates each friendship with a topic mixture from which both nodes' contents are explained. It models the undirected graph without resorting to node duplication as required for the citation influence model. As in social networks most edges are absent because the users did not get to know each other, the shared taste model is agnostic towards absent edges. Thereby it is in contrast to blockmodels which treats absence of edges as evidence for incompatibility of the nodes.

The second main contribution is to study these differences theoretically and empirically. The goal is to identify which model assumptions are crucial to learn shared tastes from online social network data. Correctly inferred shared tastes will give rise to successful predictors for social interaction. For instance, shared tastes will correlate with membership in special interest groups. On data from a music platform, inferring shared tastes allows us to predict music a user will listen to in the near future, as inspired by his friends. Further the tastes correlate with music a user recommends to his friend. Such social interactions are held out from training and used for evaluation purposes only.

Outline. Section 2 introduces the shared taste model. After revising the citation influence model and stochastic blockmodel briefly, we discuss their application to the new task of identifying shared tastes. Section 3 reports on experimental results on data sets from LibraryThing, Zune Social, and CiteSeer. Concluding remarks and suggestions for practitioners are given in Section 4.

2 Topic Models for Content Enriched Social Networks

Latent Dirichlet allocation (LDA) [1] is a probabilistic model to distill topics from a corpus of text documents. The underlying assumption is that repeated co-occurrence of items across different collections $\mathcal{C}(n_1)$, $\mathcal{C}(n_2)$, ...$\mathcal{C}(n_{|\mathcal{N}|})$ is a strong indicator for these items to be grouped in one topic t.

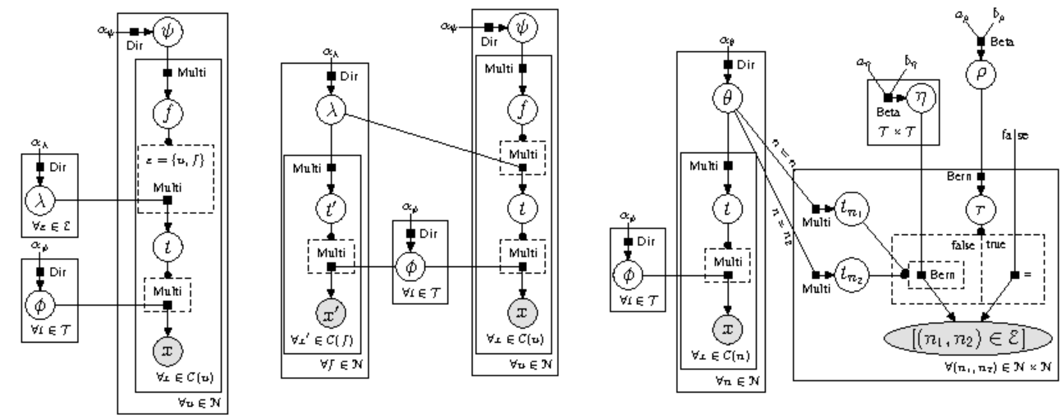
Applying latent Dirichlet allocation on content-enriched network data reveals tastes of users that are shared across the network. A post hoc heuristic can identify which topics are shared by two users, but the model does not exploit the friendship information to learn more suitable topics.

In the following, we study extensions to latent Dirichlet allocation which incorporate friendship information into the LDA model to learn better shared topics.

2.1 Shared Taste Model

The main idea is that each friendship potentially has shared interests, where the shared tastes are captured by a topic mixture λ. Via shared interests, the friends inspire a user to interact with items. Items are therefore explained by the topic mixtures of the respective friendships. The model allows for different friends having varying influence on the item interactions, some may even have no influence at all. Degree of influence is modeled by a multinomial parameter ψ ranging over the user's friends.

[1] Extended abstracts on this model were published as work in progress in [10, 11]. The work is part of the first author's PhD thesis.



(a) Shared taste model (b) Citation influence model (c) Extended stochastic blockmodel

Figure 1: Models in directed factor graph notation [12] with dashed boxed indicating gates [13].

Algorithm 1 Generative process of the shared taste model with own topics. Leaving out lines marked with * yields the plain shared taste model depicted in Figure 1a.

1 **for all** topics $t \in \mathcal{T}$ **do**
2 **draw** item distribution $\phi_t \sim$ Dirichlet(α_ϕ)
3 **for all** undirected edges $\{u, f\} \in \mathcal{E}$ **do**
4 **draw** shared topic mixture $\lambda_{\{u,f\}} \sim$ Dirichlet(α_λ).
5 **for all** nodes $u \in \mathcal{N}$ **do**
6 **draw** friend mixture $\psi_u \sim$ Dirichlet(α_ψ) ranging over friends of u.
7* **draw** own topic mixture $\Omega_u \sim$ Dirichlet(α_Ω).
8* **draw** coin $\epsilon_u \sim$ Beta(a_ϵ, b_ϵ).
9 **for all** items $x_{u,i} \in \mathcal{C}(u)$ **do**
10* **draw** decision $e_{u,i} \sim$ Bernoulli(ϵ_u).
11* **if** $e_{u,i} =$ **true then**
12* **draw** topic $t_{u,i} \sim$ Multinomial(Ω_u) from own topics.
13* **else**
14 **draw** friend $f_{u,i} \sim$ Multinomial(ψ_u).
15 **draw** topic $t_{u,i} \sim$ Multinomial($\lambda_{\{u,f_{u,i}\}}$) from shared taste.
16 **draw** item $x_{u,i} \sim$ Multinomial($\phi_{t_{u,i}}$) from the topic's item distribution.

The full generative process is given in Algorithm 1 (omitting lines with *); the corresponding directed factor graph is given in Figure 1a. A topic mixture θ for each user can be derived from friendship-wise topic mixtures λ by integrating out the graph structure and letting $\theta_u = \psi_u \lambda$.

During model estimation, configurations for hidden variables achieve high likelihood where items of two friends are associated with the same topic. The more items fit, the higher the friend's probability mass in the multinomial distribution ψ (cf. Algorithm 1, lines 6 and 14). It is a priori unknown which items of u's and f's content are to be explained by the shared interest. This introduces mutual dependency on estimation of shared topic mixtures λ and the friend mixtures ψ.

We assume that unconnected users do not influence each other's content, but similar tastes may develop independently. Therefore we devise the model to be agnostic towards absent edges: Topics of unconnected users are statistically independent, it is not assumed that they should be different.

Shared taste model with own topics. The shared taste model assumes that all items are inspired by friends. As users may develop tastes independently from friends, we extend the model as follows: The model gets the freedom to choose whether to explain an item via shared tastes ($\psi_u \lambda$) or by the user's own topic mixture Ω_u. The choice is modeled by a coin trial e. The resulting process is given in Algorithm 1. This is inspired by the "innovation topic mixture" introduced in [9].

The own topics Ω_u represents interests of a user that are not matched by his current friends. We might find new friends for him that cover these unmatched interests. In the example of Figure 2a, we might find new friends for the isolated "Poetry Fool" member in the far right.

2.2 Citation Influence Model for Social Networks

The citation influence model [9] was originally introduced to model citation networks, which are directed and acyclic. It explains documents in their roles of cited and citing publications. The authors suggest modeling cited publications as in latent Dirichlet allocation, while citing publications are explained by reusing topics associated with any of the publication it cites. The generative process for shared tastes is depicted in Figure 1b, we study the full model with the extension to innovation topic mixtures. As publications can take on both roles at the same time, the authors suggest to duplicate nodes in the citation graph to yield a bipartite graph with edges pointing from citing to cited nodes.

We transfer the model to social networks data, where graphs are undirected and contain loops and cliques (cf. supplementary material). We use the same trick of bipartite duplication: every node is represented by two copies: one for the active role of a user, and one for the passive role of a friend. The copy of the active role is associated with a distribution over friends ψ, the copy of the passive role is associated with a topic mixture λ. During model estimation, the topic mixture λ is estimated from all items of the user itself as well as items of all connected users. Intuitively, the topic mixture λ resembles the topics that describe the social role that the user plays within his friendship network, or the reason why people want to be friends with him. This is in contrast to latent Dirichlet allocation, where the topic mixture describes the items of the user in isolation.

Citation influence model versus shared taste model Both the shared taste and the citation influence model build on the same fundamental assumptions:

- Nodes in the network use topics from/across edges, preferring some neighbors over others.
- Absent edges do not reflect incompatible topics.

The differences are of subtle nature: The citation influence model uses topic mixtures for nodes together with its neighborhood. This will encourage topics of a broad nature. The shared taste model associates topic mixtures with edges. Each topic mixture represents the pair-wise interaction only. Thus, topics will be suitable for representing fine-grained aspects of sharing. A technical improvement is that shared taste model does not require duplication of any contents.

2.3 Stochastic Blockmodel

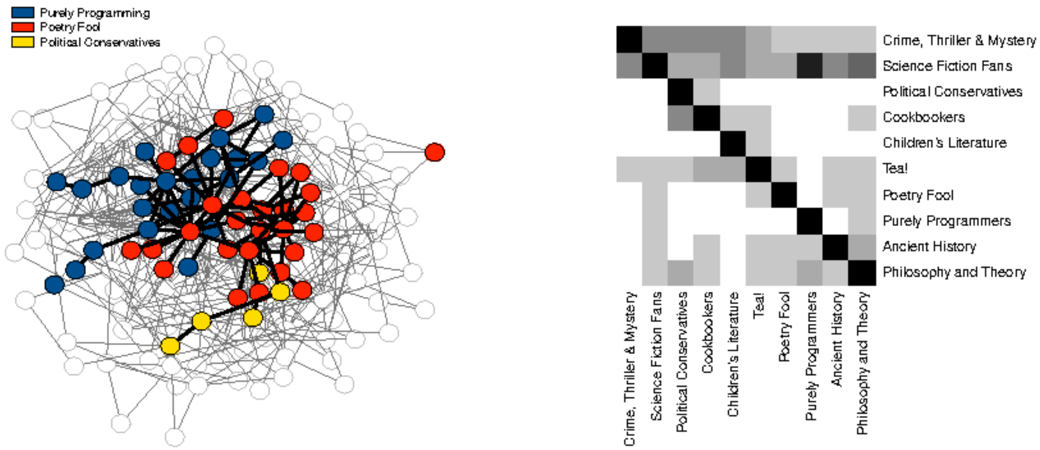
Many models for graph data are based on the stochastic blockmodel [14, 6, 15, 16], of which we focus on its mixed-membership variant. The model builds on the assumption that data can be explained by a set of communties. Nodes are members of communities, where this membership is of varying strength modeled by θ. Edges are formed between nodes of compatible communities. During inference, the observation of an edge is treated as evidence that the nodes are members of compatible communities. But not observing an edge is treated as evidence that the nodes are members of incompatible communities.

Stochastic blockmodels take absence of evidence as evidence for incompatibility. This assumption is valid in fully observed networks. We claim that this assumption is heavily violated in data from social networking platforms, where the set of friendships is incomplete.

Airoldi et al. [6] suggest to fix the model with an extension that corrects for rarity of interactions. For each absent edge, a coin trial $r \sim \rho$ decides whether the absence is due to incompatibility ($r =$ false) or due to rarity of interactions and therefore does not effect the estimation ($r =$ true).

Stochastic blockmodels for content-enriched networks. In order to use stochastic blockmodels to learn shared tastes from text content as well as graph structure, we follow on the approach taken by Nallapati et al. [7] in the Pairwise Link-LDA model. Here each community is associated with a distribution over items (in analogy to ϕ in LDA); each node is associated with a topic mixture θ that governs his community memberships and topics in his items.

We include the correction for rarity of interactions into Pairwise Link-LDA to arrive at the generative process (see supplement and Figure 1c). Since the task is focussed on shared tastes of friends we let the prior on compatibility η within a community $\eta_{t_1=t_2}$ be higher than across communities $\eta_{t_1\neq t_2}$.



(a) Selected user-friend graph in LibraryThing data. Colored users are members in three exemplary groups. Uncolored nodes participate only in the other seven groups.

(b) Conditional dependencies within group memberships $p(n \in g_{\text{Row}} | n \in g_{\text{Col}})$ on LibraryThing. (Black: $p = 1.0$, white: $p = 0.0$).

Figure 2: Statistics of the LibraryThing data set.

Blockmodel versus shared taste model The shared taste model and the stochastic blockmodels only share the underlying topic model, but employ it in a different way: The shared taste model associates edges with topic mixtures to capture shared tastes. The stochastic blockmodel associates a user with topic mixtures governing his items only. Further, friendship information is incorporated differently. The shared taste model models sharing of topics via edges. The stochastic blockmodel rather draws inference on topical compatibility of the users. But the key difference is that:

- Stochastic blockmodel treats absent edges as evidence for incompatibility.
- Shared taste model (and citation influence model) are agnostic towards absent edges.

Rarity of interactions is intrinsically addressed by the shared taste model—therefore not requiring correction for absent edges.

3 Experimental Evaluation

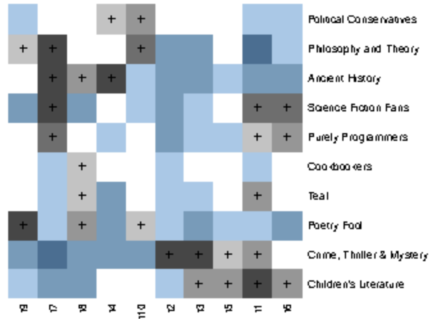
The models are evaluated on social network data from LibraryThing, Zune social, and citation data from CiteSeer. We empirically compare the three models with various extensions and latent Dirichlet allocation. Generative processes and implementation details for all methods are given in the supplementary material.

As LDA does not learn from the given network structure, any improvement gained by network topic models over LDA is due to the ability to exploit the network structure. To judge how much an underlying topic model contributes to the solution, we further study heuristic predictors.

All models and reference methods are implemented using Infer.NET [17] with variational message passing [18] as an inference algorithm. The resulting inference algorithm's memory consumption scales to reasonably sized graphs of several hundred nodes (each having content lists of thousands of items) even for the deeply nested models under study. This is sufficient for the motivated applications, which focus on supporting a particular user in predicting interaction with his friends.

Since labeled data for tuning is hard to get in a real application, we identify sensible choices of very weak prior parameters from literature and by inspection on a held out data set. We use uninformative Dirichlet priors for friend distribution and own topic mixtures $\alpha_\psi = 1$ and $\alpha_\Omega = 1$. For topic mixtures and item distributions we use recommendations from literature [6, 9, 19, 2] for sparse Dirichlets $\alpha_\theta = \alpha_\lambda = 0.1$, $\alpha_\phi = 0.01$ [9, 19]. Encouraging topic sharing over own topics by factor ten we set $a_\epsilon = 5.0$ (own tastes), $b_\epsilon = 45.0$ (shared tastes). For the blockmodels, we encourage links within each topic by using $a_\eta = 8$, $b_\eta = 2$ on the diagonal (average within topic compatibility $\eta_{t_1=t_2} = 0.8$) and $a_\eta = 2$, $b_\eta = 8$ otherwise (on average across topic compatibility is $\eta_{t_1 \neq t_2} = 0.2$, following our impression from Figure 2a); we model rarity of interactions with a vague Beta prior

Shared taste model with own topics:

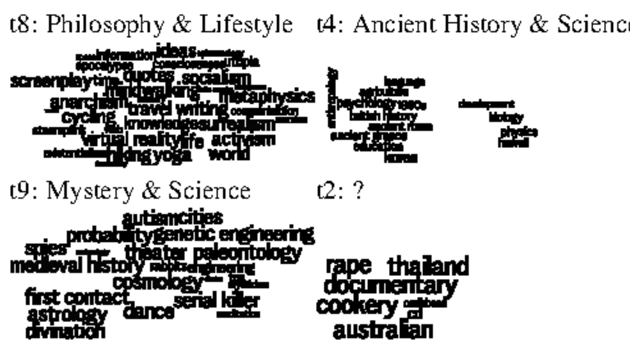


(a) Group-topic correlation

(b) Tag clouds according to ϕ_t

Stochastic blockmodel:

(c) Group-topic correlation



(d) Tag clouds according to ϕ_t for the strong correlating topics

Figure 3: The shared taste model and with own topics (top) find better topics than the stochastic blockmodel (bottom) on data from LibraryThing. Group-topic correlations $\rho_{t,g}$ (Black/+: positive correlation; blue: negative correlation; white: uncorrelated.) Tag clouds according to $\phi_t(x)$ with manually annotated headers.

$a_\rho = 3.25$, $b_\rho = 1.25$ which on average explains only 25% of absent edges by the blockmodel rather than rarity of interactions.

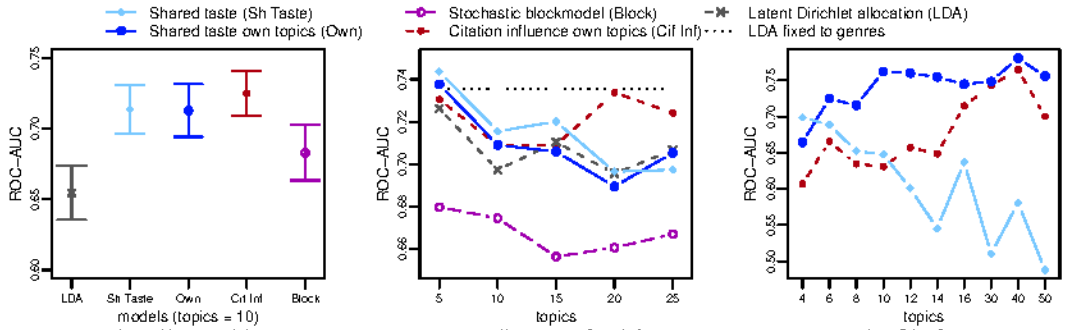
3.1 LibraryThing

LibraryThing[2] is a social networking platform centered around books. Users organize their virtual libraries using tags. The platform hosts user groups, providing a group-wise discussion forum and suggested reads centered around a common subject. We study whether the models can incorporate friend relationships and tags to learn topics that are well suited to explain social interaction. We evaluate the quality of topics by whether they correlate with membership in user groups, an information that is held out from training.

Data set. We choose a subgraph of $|\mathcal{N}| = 194$ users and a tag vocabulary $\mathcal{V}$ of size 748. The resulting friendship graph $(\mathcal{N}, \mathcal{E})$ is given in Figure 2a. For evaluation purposes only we select a held out ground truth of 10 user groups (listed in Figure 2b). In Figure 2a members of three user groups are highlighted, bold edges indicate friendships inside these three groups. The picture tells that the input data does not factor into graph clusters that follow the interest groups—violating an assumption often made in related work. However Figure 2a confirms that the data contains friendships with shared interests, which is the basic underlying assumption of the shared taste model.

The data set statistics indicate that some groups attract the same users. For instance, nearly all members of the "Purely Programming" group are also in "Science Fiction Fans". For that reason

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



(a) LibraryThing (b) Zune Social (c) CiteSeer

Figure 4: Prediction performance on held out interactions in ROC-AUC (higher is better).

these groups are not distinguishable from the graph structure alone. A successful model should distinguish these two groups based on the tag content. The correlation in group membership is visualized in Figure 2b.

As we are interested in finding the original ten user groups, we set the number of different topics $|\mathcal{T}|$ to 10.

Evaluation methods. We expect inferred topics to represent groups. We evaluate this by the Pearson correlation $\rho_{t,g}$ between the predicted topics $t \in \mathcal{T}$ and the true held out membership in groups g across all users. The correlation matrix ρ of all topics versus all groups are given in Figures 3a for the shared taste model and in Figure 3c for the stochastic blockmodel. Strong pro- and anticorrelation indicates good performance; no correlation (white) indicates bad performance. A strong diagonal would be ideal, but we do not expect to find a perfect match between unsupervised topics and a held out ground truth. Rather, we count one or two strong positive correlations per row as a success.

As it is difficult to compare correlation matrices of different models, we summarize them in one scalar number that represents the consistency of group-topic correlation across users: How well does weighting the topic mixture with the Pearson correlation coefficient resemble a user's true group membership? Results are given as averaged ROC-AUC values in Figure 4a. The standard error bars are given for reference, but as difficulty of predictions vary widely for all models (e.g., for Shared Taste model IQR=[0.55,0.9]) we test for significance with paired t-tests ($\alpha = 5\%$).

Further, items which are discriminative for each topic are visualized as tag clouds in Figures 3b, and 3d to qualitatively study the inferred topics.

Results. The shared taste and the citation influence model have the most consistent pro- and anticorrelation patterns in Figure 4a. The correlation matrix for the shared taste model in Figure 3a revealed that seven groups have strong positive correlation with one or two topics, demonstrating their success in identifying held out interests. This happened for six groups with the citation influence model. Both models are successful in distinguishing the related groups "Purely Programming" from "Science Fiction Fans" by content. Refering to the tag clouds in Figure 3b, topic t6 gathers computer terms and topic t7 science fiction terms.

In contrast, the performance of stochastic blockmodel approaches in Figure 4a are worse. In fact, they are similar to the performance of latent Dirichlet allocation, which ignores graph structure. Inspecting the compatibilities η and tag clouds in Figure 3d demonstrates the diminished performance, that even correcting for rarity of interactions did not overcome. Confer to the supplementary material for more details.

3.2 Zune Social

Microsoft's portable music player Zune comes with access to the platform Zune Social. The platform allows users to interact with other users, for instance by recommending songs to friends, and to synchronize the player for offline use. We selected a connected subgraph with $|\mathcal{N}| = 100$ users and vocabulary $\mathcal{V}$ of 1000 artists in playlists of one month. New artists of held out songs a user u listened to in the following month are predicted using the user's topic mixture. Further, new artists user u

recommends to his friend f are predicted using the shared topic mixture $\lambda_{\{u,f\}}$. Both rankings are evaluated by ROC-AUC versus held out data of true interactions. To evaluate generalization ability beyond overfitting, test items that were already included in the training set of the user are omitted from the evaluation. For comparison we include a reference method that uses predefined genres instead of estimated topics as a gold standard; genre information is held out from the models.

Results. The results are summarized in Figure 4b. For more results confer to the supplement material.

We verified that both latent Dirichlet allocation and the shared taste model with own topics identify topics that correlate well with genres. Both distinguish rock music from hiphop, soul, and contemporary R&B and identify a further cluster on electronic music.

In Figure 4b, the shared taste model predicts future items best. However, most models do not reliably outperform each other and meeting the gold standard for several topics. Only the stochastic blockmodels performs significantly worse, even worse than LDA.

All models fail to predict recommendations between users, i.e. AUC < 0.6, where to our surprise exploiting the genre information yields even worse performance. One speculative explanation for the results is that Zune users in our data do not socially interact with respect to common topic or genre. If this is the case, the edges do not inform us about the shared taste and the models' main assumptions are violated.

3.3 CiteSeer

The citation influence model was designed for unsupervised prediction of influences from cited publications on the citing publication. The authors use training data from CiteSeer and a held out set of manually labeled influences [9]. We repeat this experiment to study the performance of the shared taste model with reference to the citation influence topic model using the estimated friend distribution ψ as a measure of influence.

Results. The results are presented in Figure 4c. Although comparable to the citation influence topic model for $|\mathcal{T}| > 16$, the shared taste model with own topics gives improved performance on low topic dimensions. However, the plain variant of the shared taste model performs significantly worse, demonstrating the benefits of the "own topics" variation.

4 Conclusions

This paper introduces the shared taste model, a new generative model for inferring shared tastes in social networks with contents. We also apply the citation influence model to this new task. Both models are grounded on similar assumptions, but shared taste model avoids content duplication. Different variations of both models are studied in relation to a topic model based on the stochastic mixed-membership blockmodel, which is quite similar to the Pairwise Link-LDA model [7]. The stochastic blockmodel takes absent edges as evidence of incompatibility—an assumption that is violated by data form online social networks. The empirical study confirms this issue. The poor results raise concerns whether models that treat unobserved edges as absent edges are appropriate for data from online social networks.

In contrast, the more successful models are agnostic towards absent edges. Both the shared taste model and the citation influence model incorporate the same basic intuition: small groups of connected users share a topic mixture, where each user participates in multiple groups with varying degree.

In most cases, the prediction performance improves when models have the flexibility to explain some items by own topics or own items. In all experiments, the extra freedom never led to degraded performance. Modeling the "own topics" for each user captures interests that are not well matched by current friends—and the model can help with finding new friends to cover unmatched interests.

With the widespread use of social media, the Internet is becoming one large social network where ties such as friendships, one-way subscriptions, and web-page visits are formed according to common interests. With this increased use, an abundance of content is generated every day, and thus information overload is a major issue in our society. Inference of shared tastes allow us to measure the interestingness for a user and will help him to focus on information he is ultimately interested in.

References

[1] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 2003.

[2] Michal Rosen-Zvi, Thomas Griffiths, Mark Steyvers, and Padhraic Smyth. The author-topic model for authors and documents. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, 2004.

[3] David Cohn and Thomas Hofmann. The missing link - a probabilistic model of document content and hypertext connectivity. In *Advances in Neural Information Processing Systems*, 2000.

[4] Elena Erosheva, Stephen Fienberg, and John Lafferty. Mixed-membership models of scientific publications. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl 1), 2004.

[5] Qiaozhu Mei, Deng Cai, Duo Zhang, and ChengXiang Zhai. Topic modeling with network regularization. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, 2008.

[6] Edoardo M. Airoldi, David M. Blei, Stephen E. Fienberg, and Eric P. Xing. Mixed Membership Stochastic Blockmodels. *Journal of Machine Learning Research*, 9, 2008.

[7] Ramesh M. Nallapati, Amr Ahmed, Eric P. Xing, and William W. Cohen. Joint Latent Topic Models for Text and Citations. In *The 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2008.

[8] Jonathan Chang and David M. Blei. Hierarchical relational models for document networks. *The Annals of Applied Statistics*, 4(1), 2010.

[9] Laura Dietz, Steffen Bickel, and Tobias Scheffer. Unsupervised Prediction of Citation Influences. In *Proceedings of the 24th International Conference on Machine Learning*, 2007.

[10] Anonymous Author. Modeling Shared Tastes in Online Communities. In *NIPS Workshop on Applications for Topic Models: Text and Beyond*, 2009.

[11] Anonymous Author. Inferring Shared Interests from Social Networks. In *NIPS Workshop on Computational Social Science and the Wisdom of Crowds*, 2010.

[12] Laura Dietz. Directed Factor Graph Notation for Generative Models. Technical report, Max Planck Institute for Informatics, 2010.

[13] Tom Minka and John Winn. Gates. In *Advances in Neural Information Processing Systems*, 2008.

[14] Yuchung J. Wang and George Y. Wong. Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*, 82(397), 1987.

[15] Krzysztof Nowicki and Tom A. B. Snijders. Estimation and Prediction for Stochastic Blockstructures. *Journal of the American Statistical Association*, 96(455), 2001.

[16] Jake M. Hofman and Chris H. Wiggins. Bayesian Approach to Network Modularity. *Physical Review Letters*, 100(25), 2008.

[17] Tom Minka, John M. Winn, John P. Guiver, and Anitha Kannan. Infer.NET 2.3. http://research.microsoft.com/infernet, 2009.

[18] John Winn, Christopher M. Bishop, and Tommi Jaakkola. Variational message passing. *Journal of Machine Learning Research*, 6, 2005.

[19] Mark Steyvers and Tom Griffiths. Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, and W. Kintsch, editors, *Latent Semantic Analysis: A Road to Meaning*. Laurence Erlbaum, 2005.