Integrating Topic Model and Heterogeneous Information Network for Aspect Mining with Rating Bias

Yugang Ji[1], Chuan Shi[1(✉)], Fuzhen Zhuang[2,3], and Philip S. Yu[4]

[1] Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing, China
{jiyugang,shichuan}@bupt.edu.cn
[2] Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China
zhuangfuzhen@ict.ac.cn
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[4] University of Illinois at Chicago, Chicago, USA
psyu@uic.edu

Abstract. Recently, there is a surge of research on aspect mining, where the goal is to predict aspect ratings of shops with reviews and overall ratings. Traditional methods assumed that aspect ratings in a specific review text are of the same level, which equal to the corresponding overall rating. However, recent research reveals a different phenomenon: there is an obvious rating bias between aspect ratings and overall ratings. Moreover, these methods usually analyze aspect ratings of reviews with topic models at textual level, while totally ignore potentially structural information among multiple entities (users, shops, reviews), which can be captured by a Heterogeneous Information Network (HIN). In this paper, we present a novel model integrating Topic model and HIN for Aspect Mining with rating bias (called THAM). Firstly, a phrase-level LDA model is designed to extract topic distributions of reviews by using textual information. Secondly, making full use of structural information, we constructs a topic propagation network, and propagate topic distributions in this heterogeneous network. Finally, by setting review as the sharing factor, the two parts are integrated into a uniform optimization framework. Experimental results on two real datasets demonstrate that THAM achieves significant performance improvement, compared to the state of the arts.

Keywords: Aspect mining · Rating bias · Topic model · Topic propagation network · Heterogeneous information network

1 Introduction

With the rapid development of E-commerce, a large number of opinion reviews and ratings have been accumulated on the Web in the past decade [3,9,14].

Q. Yang et al. (Eds.): PAKDD 2019, LNAI 11439, pp. 160–171, 2019.
https://doi.org/10.1007/978-3-030-16148-4_13

These reviews and ratings have played an important role which can not only help people make more favorable purchase decisions, but also give valuable advice to the shops [1,10]. For instance, users may pay attention to both of overall ratings and reviews of a shop before making purchase decisions. The owner of a shop can learn the positive and negative feedback embedded in users' reviews as well.

In recent years, there is a surge of research on aspect mining and the main goal of aspect mining is tox effectively discover the aspect distribution and the aspect ratings of entities [16]. To address this problem, the earlier studies prefer to take advantage of Probabilistic Latent Semantic Analysis (PLSA). For example, both of Lu et al. [5] and Luo et al. [7] regarded reviews as several opinion phrases and respectively designed two PLSA-based models. However, these two models ignored the influence of ratings to reviews. Recently, many researchers [6,8,13,15] took the influence of ratings into consideration and utilized Latent Dirichlet Allocation (LDA) to describe the generation of reviews in details. Luo et al. [6] paid attention to the latent distribution of overall ratings and designed an LDA-based method for aspect rating prediction. Laddha et al. [2] integrated both discriminative conditional random field, regression, LDA to simultaneously extract phrases and predict ratings.

Almost all models for aspect mining usually have a basic assumption that the overall rating could be close to aspect ratings or the average score of aspect ratings. Thus, these methods preferred to directly associate review phrases or terms with the corresponding overall rating. However, recent research [4] found an insightful observation that there is an obvious rating bias between overall rating and aspect ratings. For example, in Dianping, the bias between overall rating and Environment is often $+0.54$, while in TripAdvisor, the bias between overall rating and Food is -0.09. This phenomenon indicates that review phrases or terms are more likely rated by latent aspect ratings rather than overall rating. Furthermore, Li et al. [4] proposed the RABI model to handle aspect rating prediction considering rating bias. Although the RABI obtains performance improvement on aspect rating prediction compared to previous models, there are several weaknesses existing in RABI. On the one hand, this model is based on PLSA without considering some other latent dependence, for example, the topic of modifier. This may restrict performance improvement. On the other hand, in the RABI model, it assumes that overall rating is on the center of the model, where it determines the reviews and aspects. Although this assumption may simplify the model, it is a little against our common sense.

Besides, contemporary aspect mining methods all focus on making use of textual information and overall rating, but ignore abundant structural information existing on review networks among the multi-typed entities, such as users, shops, and reviews. However, these structural information may be useful for aspect mining. For example, the reviews given by a user can describe his profile and the generation of a review is influenced by the quality of a shop as well as the corresponding user profile. In order to utilize the rich structural information of these multi-typed entities (e.g., users, reviews, and shops) and the various relations (e.g., writing and evaluating) among them, it is naturally to form the review network as a Heterogeneous Information Network (HIN) [11,12]. In a review HIN, both of user profile and shop profile can be easily described by propagating topic distribution of review texts to neighbour entities. Similarly, the topic distribution of review texts can be influenced by the profiles of their neighbour entities.

Motivated by these observations, we propose a novel method integrating Topic model and Heterogeneous information network for Aspect Mining with rating bias (THAM for short). To overcome the weaknesses of RABI [4] and describe the process of generating reviews more reasonably, THAM designs a LDA-based topic model at phrase-level to describe the generation of reviews and mine the aspect rating distribution of each review text. In this topic model, the modifier term of a phrase is associated with the sampled aspect rating rather than directly rated by overall rating because of the existing rating bias. Moreover, taking the abundant structural information into consideration, we propose a topic propagation network based on HIN to propagate topic distribution among users, shops and reviews for keeping the consistency of topic distributions of neighbour entities. Furthermore, in order to effectively fuse textual information and structural information, we design a uniform optimization framework through setting reviews as the sharing factor to integrate topic model and topic propagation network. An iterative optimization algorithm is proposed for this optimization framework.

2 Preliminary

Here we introduce the relevant concepts and the problem of aspect mining with rating bias.

Review: A review d is the text to express the user's opinion of a shop, and there are $|\boldsymbol{D}|$ reviews in total.

Phrase: A phrase $l = < h, m >$ consists of a head term h and its modifier term m, for example, $< food, delicious >$. There are $|\boldsymbol{L}|$ phrases in all.

Aspect/Topic: An aspect z is a specific topic of a shop. There are K aspects/topics. Note that, "topic" and "aspect" are used interchangeable in this paper.

Overall rating: An overall rating r is the quantified overall opinion of a review d. There are R levels of overall ratings and R is usually 5.

Aspect rating: An aspect rating $r_{s,z}$ is a numerical rating on the aspect z of the shop s. There are R levels of aspect ratings too.

Rating bias: The rating bias is the gap between the average of overall ratings and the average of aspect ratings.

Heterogeneous information network: Heterogeneous information network (HIN) is a special information network containing multiple entities and various relations [11]. For instance, the review network shown in left box of Fig. 1 is such a network, which contains three types of entities: user (u), shop (s), and review (d), and each edge represents a specific relation (e.g., "writing" for u to d).

Aspect rating prediction with rating bias: This problem is to predict ratings on each aspect for each shop with the bias prior information. Given a set of reviews $\boldsymbol{D}$ written by users $\boldsymbol{U}$ to evaluate shops $\boldsymbol{S}$, the task is to identify the aspect of each phrase and predict the aspect ratings of shops considering rating bias.

Since aspect ratings are always missing in real applications but very valuable to users and shops, aspect rating prediction is an effective way to repair the missing information. Moreover, rating bias plays an significant role to improve the accuracy of aspect ratings [4]. Therefore, it is meaningful to study the problem of aspect mining with rating bias.

3 The THAM Model

In this section, we propose the THAM model, which makes full use of textual information and structural information for addressing the problem of aspect identification and aspect rating prediction with rating bias.

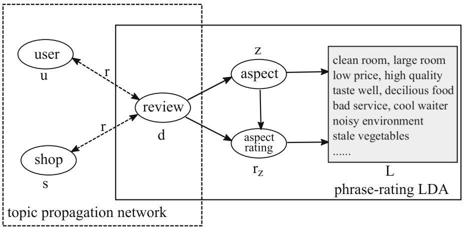

Fig. 1. The framework of THAM model. The dotted-line box is a network schema of topic propagation network, and the solid-line box describes the phrase-rating LDA.

3.1 The Phrase-Rating LDA

Here we design the phrase-rating LDA to more effectively learn topic distributions of reviews at textual level. In Fig. 2, both of aspect and aspect rating of a review are assumed as latent factor respectively sampled by aspect distributions and aspect rating distributions. Moreover, each review consists of several opinion phrases, in which head terms are generated by aspects while modifier terms are dependent on the sampled aspect ratings. Different from related methods, we consider the sampled aspect rating is associated with not only the observed overall rating but also the corresponding rating bias. Obviously, r_z, the sampled aspect rating of modifier term m, plays quite significant role in this model. Taking rating bias into consideration, we make two basic assumptions about aspect ratings. On the one hand, r_z is sampled by the aspect rating distribution of d. On the other hand, the mean of aspect rating distribution could be similar to

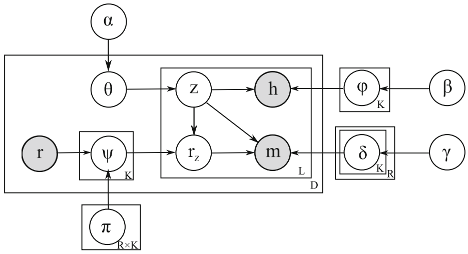

Fig. 2. The graphical model of the phrase-rating LDA.

but not equal to overall rating r because of the rating bias. Therefore, we design an aspect rating distribution ψ and regard that the Dirichlet prior parameter for aspect rating distribution, π, is related to the overall rating r. Given the observed overall rating r, the π_{r,k,r_z} is defined as follows:

$$\pi_{r,k,r_z} = B(r'_z|\omega(1-r'),\omega(r')), \tag{1}$$

where $B(\cdot)$ is the beta probability distribution, $0 < r' < 1$, $0 < r'_z < 1$ respectively represents the small scaled value of r and r_z, ω is the prior parameter. By using Eq. 1, we cleverly utilize overall rating to constrain the corresponding aspect rating distribution. Moreover, taking the rating bias into consideration, we set the aspect z's rating levels as $\{1-b_z, 2-b_z, ..., R-b_z\}$.

Given a set of review texts and overall ratings, both of r and $< h, m >$ are the observed variable, α, β, γ, π are the Dirichlet prior parameters, and the main latent parameters learnt are θ, ψ, ϕ, δ, z, and r_z. Given the model parameters and overall rating, the probability of observing the review text (i.e., the likelihood) is:

$$L_1 = -log(\prod_d \prod_l \sum_z \sum_{r_z} p(z|\theta_d)p(r_z|\psi_{d,z})p(h_l|\phi_z)p(m_l|\delta_{r_z,z})). \tag{2}$$

We employ Gibbs sampling to estimate the posterior probability given the observed phrases.

It is noteworthy that the phrase-rating LDA does not describe the dependence between aspect distribution θ and overall rating r, because the dependence is closely associated with user profile and shop profile.

3.2 Topic Propagation on Review Network

In order to make full use of structural information for aspect mining, we design a HIN-based topic propagation network shown in Fig. 3, to propagate topic distribution among neighbour entities so as to describe user profile and shop profile.

In the topic propagation network, the topic distribution of each entity should be related to its neighbour entities. Furthermore, we constrain the topic propagation must under the same overall rating. This constraint is reasonable because

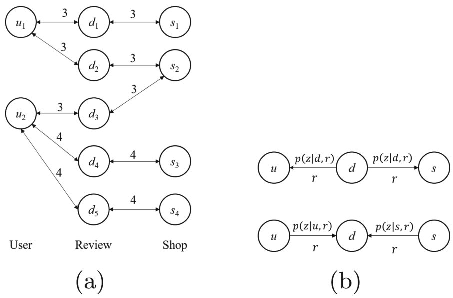

Fig. 3. An example of topic propagation network and the topic propagation strategy.

the aspect distributions under different overall ratings represent different meanings. For instance, if a user gives high overall rating to a shop, he is likely to write many positive phrases to describe the aspect he cares, and vice versa.

Therefore, given the topic distribution of a review $p(z|d, r)$ where r is the observed overall rating of the corresponding review d, we design the topic propagation strategy as shown in Fig. 3(b), and the topic distribution of a user u is constructed by his/her reviews, denoted as:

$$p(z|u, r) = \sum_{d_{u,r} \in \boldsymbol{D}_{u,r}} p(z|d_{u,r})p(d_{u,r}|\boldsymbol{D}_{u,r}) = \sum_{d_{u,r} \in \boldsymbol{D}_{u,r}} \frac{p(z|d_{u,r})}{|\boldsymbol{D}_{u,r}|}, \tag{3}$$

where $\boldsymbol{D}_{u,r}$ is the set of reviews under overall rating r belonging to user u, $|\boldsymbol{D}_{u,r}|$ is the number of these reviews. Similarly, the topic distribution for a shop is constructed by its reviews, calculated by:

$$p(z|s, r) = \sum_{d_{s,r} \in \boldsymbol{D}_{s,r}} p(z|d_{s,r})p(d_{s,r}|\boldsymbol{D}_{s,r}) = \sum_{d_{s,r} \in \boldsymbol{D}_{s,r}} \frac{p(z|d_{s,r})}{|\boldsymbol{D}_{s,r}|}, \tag{4}$$

where $\boldsymbol{D}_{s,r}$ is the set of reviews under overall rating r belonging to shop s, $|\boldsymbol{D}_{s,r}|$ is the number of these reviews.

Furthermore, the review d should have similar topic distribution with its author u and its shop s. Therefore, aiming at obtaining effective topic distribution of reviews, we design two functions, one of which is to calculate the similarity of d and u, and the other is to calculate the similarity of d and s. The two functions are shown as follows:

$$L_2 = \frac{1}{2} \sum_d \sum_z [p(z|d, r) - p(z|u_d, r)]^2, \tag{5}$$

$$L_3 = \frac{1}{2} \sum_d \sum_z [p(z|d, r) - p(z|s_d, r)]^2, \tag{6}$$

where u_d is the user who writes the review d and s_d is the shop whom the review d evaluates.

3.3 Uniform Optimization Framework

To make full use of textual information and structural information at the same time, the THAM model incorporates the topic model and the topic propagation into a uniform optimization framework.

In this framework, we consider the review d as the sharing factor, which plays significant role not only in topic propagation but also in topic modelling. To ensure the optimization process of the model, we design a combined loss function here:

$$Loss = L_1 + \frac{\lambda}{2}(L_2 + L_3), \tag{7}$$

where $\lambda \geq 0$ is to control the balance between topic modelling and topic propagation. Obviously, if $\lambda = 0$, we only take into account the loss of phrase-rating LDA. With the increase of λ, the loss from topic propagation will be paid more and more attention.

There are two main steps for learning the algorithm. In topic modelling, we sample the distribution of reviews on phrase-rating LDA to reduce L_1 where L_2 and L_3 are fixed. In topic propagation, we get rid of the Newton-Raphson updating formula, which decreases function $f(x)$ by updating $x_{t+1} = x_t - \xi \frac{f'(x_t)}{f''(x_t)}$, to decrease L_2 and L_3. $p(z|d,r)$ in topic propagation is updated by:

$$p(z|d,r)_{t+1} = (1-\xi)p(z|d,r)_t + \frac{\xi}{2}(p(z|u_d,r)_t + p(z|s_d,r)_t), \tag{8}$$

where ξ is a step parameter. Then, the corresponding topic distribution of u_d and s_d can also be updated by in Eqs. (3) and (4) respectively. In this step, we also take L_1 into consideration because the updated $p(z|d,r)$ (i.e., θ) can also influence the value of L_1.

3.4 Aspect Identification and Rating Prediction

Based on the obtained aspect-head distribution ϕ and aspect-modifier distribution δ, we can identify the aspect which phrase $l =< h, m >$ should be assigned to by using Eq. (9):

$$g(l) = \underset{z'}{\operatorname{argmax}} \sum_{r_z} \delta_{r_z,z',m} \phi_{z',h}, \tag{9}$$

and the corresponding rating of l is:

$$r_l = \frac{\sum_{r_z} \delta_{r_z,z,m} \phi_{z,h} r_z}{\sum_{r_z} \delta_{r_z,z,m} \phi_{z,h}}, \tag{10}$$

where $z = g(l)$. And the predicted aspect rating of each shop is calculated by:

$$\widehat{r}_{s,z} = \frac{\sum_{d \in \boldsymbol{D}_s} \sum_{r_z} \psi_{d,z,r_z} r_z}{\sum_{d \in \boldsymbol{D}_s} \sum_{r_z} \psi_{d,z,r_z}}, \tag{11}$$

where $\widehat{r}_{s,z}$ is the rating on aspect k of shop s, $\boldsymbol{D}_s$ is the set of reviews of shop s.

4 Experiments

4.1 Dataset

There are two real datasets in different languages for conducting experiments: Dianping in Chinese [4] and TripAdvisor in English. The review information in Dianping dataset consists of a Chinese review text and three aspect ratings on Taste, Service, and Environment. Similarly, the TripAdvisor dataset crawled from the TripAdvisor website, is a set of English reviews and each review includes English comments, an overall rating and three aspect ratings on Value, Service, and Food. In addition, the range of ratings in the two datasets are in $[1, 5]$. The statistics of the two datasets are shown in Table 1. Note that, b_1, b_2, and b_3 respectively represents the rating bias of Taste, Service and Environment (on Dianping) or Value, Service, and Food (on TripAdvisor).

4.2 Preparation

To obtain phrases from reviews, the dataset is preprocessed via the process similar to that in RABI [4] Besides, the number of aspects K is set as 3 for both of Dianping and TripAdvisor. The prior parameters α, β, and γ in the phrase-rating LDA are set as $50/K, 0.01, 0.01$ respectively. ω is set as 2.5 and ξ is set as 0.1. The max iteration is 1000. The controlling parameter λ is adjusted to 9000 on both the two datasets by parameter analysis.

Since the performance of an aspect mining method may be affected by the size of the training dataset [4], we sample four subsets for Dianping and TripAdvisor with different scales of reviews (i.e., 25%, 50%, 75%, 100% of review data). To ensure that the latent aspects correspond to the given aspects, we also select several head terms as prior for each latent aspect.

We select Root Mean Square Error (RMSE) and Pearson Correlation Coefficient (PCC) as evaluation metrics. RMSE is to measure the average difference between real ratings and predicted ratings on all aspects. The smaller the value of RMSE, the better the algorithm performs. Considering that rating prediction are often used for ranking-based recommendation, we also measure the linear relation of the predicted results and the real results by the PCC metric. The larger value of PCC represents the better performance.

4.3 Comparison Methods

To demonstrate the effectiveness of THAM model, four representative methods, including QPLSA [7], SATM [13], AIR [3], and RABI [4], are adopted

Table 1. The statistic information of Dianping and TripAdvisor

Dataset	# Users	# Shops	# Reviews	# Phrases	Avg. Overall Rating	b_1	b_2	b_3
Dianping	14519	1097	216291	696608	3.97	+0.28	+0.48	+0.54
TripAdvisor	107368	5178	243186	2544148	4.10	+0.27	−0.05	−0.09

for comparisons. Since neither QPLSA and SATM nor AIR takes into account the rating bias, we adjust the results of the three baselines through subtracting/adding the rating bias for fair comparison and mark the adjusted method with "*".

Furthermore, in order to validate the effectiveness of rating bias and structure information, we also test three simplified versions of our THAM. First, we remove the assumption of rating bias from THAM, and call this version THAM\B. Second, we only use the textual information without topic propagation. We call it THAM\H. Third, we remove both of rating bias and topic propagation form THAM, and call this version THAM\HB.

Table 2. Top 10 rated phrases for different aspects of the two datasets

Datasets	Aspects	Phrases (Ratings)
Dianping	Taste	great taste (4.35), good mouth-feel (3.61), delicious dish (3.54), *suitable price (3.52)*, delicious drink (3.33), *good restaurant (3.28)*, *high taste (3.17)*, light flavour (3.04), common flavour (2.58), few dish (2.51)
	Service	enjoyable service (3.95), enthusiastic service (3.92), comfortable service (3.78), *nice shop (3.72)*, *delicious service (3.43)*, handsome waiter (3.43), good impression (3.28), good attitude (3.25), cold waiter (1.95), not enthusiastic service (1.85)
	Environment	elegant style (4.29), *cheap price (4.08)*, nice inside (3.56), *good feeling (3.32)*, suitable position (3.32), *easy to find (3.20)*, good traffic (3.09), common environment (2.76), small room (2.31), unreasonable design(2.05)
TripAdvisor	Value	unbeatable price (4.48), best quality (4.19), cheap price (3.89), reasonable price (3.57), pricey fare (3.44), *great place (3.41)*, *big price (3.29)*, *good place (3.29)*, *good selection (3.20)*, poor value (1.81),
	Service	fantastic waitress (4.12), friendly service (4.03), courteous waiter (3.89), great experience (3.53), interesting waitress (3.18), *good drink (3.31)*, *good meal (3.24)*, first experience (3.09), slowest service (1.82), disgusting service (1.57)
	Food	amazing food (4.54), wholesome food (4.20), excellent dishes (4.03), *nice location (3.98)*, rich menu (3.31), good food (3.06), *good atmosphere (3.02)*, small dish (2.91), *small restaurant (2.77)*, not fresh dish (2.59)

4.4 Aspect Identification

Since the opinion phrases are unlabelled, it is hard to quantitatively validate the effectiveness of aspect identification. Therefore, we list some representative rated phrases for each aspect on the two datasets respectively for illustration. The most possible phrases for each aspect are automatically mined and shown in Table 2. In addition, we rank these phrases by their ratings and the meaningless phrases are marked in italic type.

Here we find that most of the extracted phrases in both English and Chinese can accurately express users' feelings about specific aspects and these frequent opinion phrases are effectively assigned to the related aspects which they describe. On the one hand, the head term of a phrase can indicate the aspect which the user describes, such as "mouth-feel" for taste and "room" for environment. On the other hand, a positive modifier term can express a positive evaluation while a negative modifier term may indicate a lower rating. It is obvious that some phrases with positive modifier terms, like "great" and "unbeatable" can get high ratings while those with negative modifier terms, like "poor" and "slowest" get the lowest ratings.

4.5 Effectiveness Experiments

In this section, we present the results of predicted aspect ratings on reviews with overall ratings and measure the performances of the different methods in terms of RMSE and PCC. In addition, each method here is run ten times and the average results (RMSE and PCC) are recorded in Tables 3 and 4 respectively.

RMSE Performance. To evaluate the accuracy of these methods on predicting aspect ratings, we calculate all RMSE values of these results. As is shown in Table 3, we can clearly find the following observations:

Table 3. RMSE performances of different methods on two datasets.

Dataset	Dianping				TripAdvisor			
	25 %	50 %	75 %	100 %	25 %	50 %	75 %	100 %
QPLSA	0.5806	0.5750	0.5724	0.5705	0.5805	0.4628	0.4005	0.3876
QPLSA*	0.3639	0.3518	0.3483	0.3460	0.5798	0.4486	0.3944	0.3833
SATM	0.5783	0.5754	0.5698	0.5601	0.6101	0.4886	0.4203	0.4064
SATM*	0.3818	0.3783	0.3704	0.3642	0.6012	0.4737	0.4136	0.3822
AIR	0.5369	0.5307	0.5157	0.5112	0.6517	0.5572	0.4778	0.4408
AIR*	0.3363	0.3207	0.3055	0.3034	0.6446	0.5475	0.4546	0.4380
RABI	0.3228	0.3150	0.3024	0.2951	0.5286	0.4388	0.3771	0.3695
THAM\HB	0.5064	0.4910	0.4873	0.4855	0.5027	0.4128	0.3614	0.3247
THAM\H	0.3089	0.2897	0.2833	0.2798	0.4920	**0.4024**	0.3477	0.3191
THAM\B	0.5060	0.4906	0.4869	0.4843	0.4985	0.4160	0.3610	0.3261
THAM	**0.3078**	**0.2891**	**0.2822**	**0.2789**	**0.4889**	0.4048	**0.3475**	**0.3101**

Compared with baselines, our THAM achieves the best performances on all subsets. There are two main advantages of our THAM. On the one hand, we utilize the bias prior information more effectively by designing a reasonable LDA-based topic model, and this topic model can overcome the weaknesses of RABI. On the other hand, we take into account not only textual information but also structural information contained in the review network, while all of baselines only focus on review texts and ratings. Comparing THAM with its variant versions, we find that THAM performs the best on most situations.

We can make two main conclusions as follows: (1) Bias prior information is effectively utilized in THAM, by comparing THAM and THAM\B; (2) The topic propagation strategy can improve the performance of our method, by comparing THAM and THAM\H.

Table 4. PCC performances of different methods on two datasets.

Dataset	Dianping				TripAdvisor			
	25 %	50 %	75 %	100 %	25 %	50 %	75 %	100 %
QPLSA	0.5689	0.5766	0.5752	0.5837	0.5715	0.5855	0.5860	0.5918
SATM	0.3503	0.3656	0.3735	0.3984	0.5535	0.5919	0.6279	0.6471
AIR	0.5670	0.5707	0.5875	0.5949	0.6643	0.6667	0.6979	0.7218
RABI	0.6130	0.6245	0.6378	0.6398	0.6582	0.6621	0.6740	0.6801
THAM\H	0.6691	0.6956	0.6999	0.7060	0.7563	**0.7696**	0.7901	0.8030
THAM	**0.6721**	**0.6971**	**0.7031**	**0.7093**	**0.7594**	0.7665	**0.7907**	**0.8075**

PCC Performance. To evaluate the ability of these models to maintain relative order among shops, we also calculate all PCC performance of these models on datasets, and show the results in Table 4. Because rating bias rarely affects the order of shops, we only compare these original methods and our THAM\H, THAM.

As is shown in Table 4, obviously, the proposed THAM obtains the best performances on almost all datasets than other baselines. We can also observe that RABI performs better than AIR on Dianping dataset while AIR does better than RABI on TripAdvisor dataset. These observations once again validate that THAM is stable and robust enough to predict aspect ratings of shops. Therefore, THAM is proved as a better choice when recommending Top-N aspect ranking orders than other baselines.

5 Conclusion

In this paper, we have proposed THAM to integrate topic model and heterogeneous information network for aspect mining with rating bias. Taking advantage of both textual and structural information, THAM designs a phrase-level LDA model and the topic propagation strategy for aspect mining. In order to integrate the two parts for optimization, THAM sets the reviews as the sharing factor and proposes a uniform iterative optimization model. By comparing the performances of baselines, THAM performs better on the two datasets for aspect mining. In the future, we can make use of heterogeneous information network more effectively for aspect mining by taking the user attributes and shop attributes into consideration.

Acknowledgements. This work is supported by the National Key Research and Development Program of China (2017YFB0803304) and the National Natural Science Foundation of China (No. 61772082, U1836206, 61702296, 61806020, 61375058).

References

1. Bauman, K., Liu, B., Tuzhilin, A.: Aspect based recommendations: recommending items with the most valuable aspects based on user reviews. In: The ACM SIGKDD International Conference, pp. 717–725 (2017)
2. Laddha, A., Mukherjee, A.: Aspect opinion expression and rating prediction via LDA-CRF hybrid. Nat. Lang. Eng. **24**, 1–29 (2018)
3. Li, H., Lin, R., Hong, R., Ge, Y.: Generative models for mining latent aspects and their ratings from short reviews. In: 2015 IEEE International Conference on Data Mining, ICDM 2015, Atlantic City, NJ, USA, 14–17 November 2015, pp. 241–250 (2015)
4. Li, Y., Shi, C., Zhao, H., Zhuang, F., Wu, B.: Aspect mining with rating bias. In: Frasconi, P., Landwehr, N., Manco, G., Vreeken, J. (eds.) ECML PKDD 2016. LNCS (LNAI), vol. 9852, pp. 458–474. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46227-1_29
5. Lu, Y., Zhai, C., Sundaresan, N.: Rated aspect summarization of short comments. In: Proceedings of the 18th International Conference on World Wide Web, WWW 2009, Madrid, Spain, 20–24 April 2009, pp. 131–140 (2009)
6. Luo, W., Zhuang, F., Cheng, X., He, Q., Shi, Z.: Ratable aspects over sentiments: predicting ratings for unrated reviews. In: 2014 IEEE International Conference on Data Mining, ICDM 2014, Shenzhen, China, 14–17 December 2014, pp. 380–389 (2014)
7. Luo, W., Zhuang, F., Zhao, W., He, Q., Shi, Z.: QPLSA: utilizing quad-tuples for aspect identification and rating. Inf. Process. Manag. **51**(1), 25–41 (2015)
8. Moghaddam, S., Ester, M.: The FLDA model for aspect-based opinion mining: addressing the cold start problem. In: International Conference on World Wide Web, pp. 909–918 (2013)
9. Pecar, S.: Towards opinion summarization of customer reviews. In: Proceedings of ACL 2018, Student Research Workshop, pp. 1–8 (2018)
10. Schouten, K., van der Weijde, O., Frasincar, F., Dekker, R.: Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. IEEE Trans. Cybern. **48**(4), 1263–1275 (2018)
11. Shi, C., Li, Y., Zhang, J., Sun, Y., Yu, P.S.: A survey of heterogeneous information network analysis. IEEE Trans. Knowl. Data Eng. **29**(1), 17–37 (2017)
12. Sun, Y., Han, J., Zhao, P., Yin, Z., Cheng, H., Wu, T.: RankClus: integrating clustering with ranking for heterogeneous information network analysis. In: ACM SIGKDD 2009, pp. 565–576 (2009)
13. Wang, H., Ester, M.: A sentiment-aligned topic model for product aspect rating prediction. In: Conference on Empirical Methods in Natural Language Processing, pp. 1192–1202 (2014)
14. Xiao, D., Ji, Y., Li, Y., Zhuang, F., Shi, C.: Coupled matrix factorization and topic modeling for aspect mining. Inf. Process. Manag. **54**(6), 861–873 (2018)
15. Yu, D., Mu, Y., Jin, Y.: Rating prediction using review texts with underlying sentiments. Inf. Process. Lett. **117**, 10–18 (2017)
16. Zhang, L., Liu, B.: Aspect and entity extraction for opinion mining. In: Chu, W.W. (ed.) Data Mining and Knowledge Discovery for Big Data. SBD, vol. 1, pp. 1–40. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-40837-3_1