



Title	From More-Like-This to Better-Than-This: Hotel Recommendations from User Generated Reviews
Authors(s)	Dong, Ruihai, Smyth, Barry
Publication date	2016-07-17
Publication information	Dong, Ruihai, and Barry Smyth. "From More-Like-This to Better-Than-This: Hotel Recommendations from User Generated Reviews." ACM, July 17, 2016. https://doi.org/10.1145/2930238.2930276 .
Conference details	UMAP '16, Halifax, NS, Canada
Publisher	ACM
Item record/more information	http://hdl.handle.net/10197/9026
Publisher's statement	© ACM, 2016. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization. https://doi.org/10.1145/2930238.2930276
Publisher's version (DOI)	10.1145/2930238.2930276

Downloaded 2026-05-01 23:35:34

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

From More-Like-This to Better-Than-This: Hotel Recommendations from User Generated Reviews

Ruihai Dong
Insight Centre for Data Analytics
University College Dublin
Belfield, Dublin 4, Ireland
ruihai.dong@insight-centre.org

Barry Smyth
Insight Centre for Data Analytics
University College Dublin
Belfield, Dublin 4, Ireland
barry.smyth@insight-centre.org

ABSTRACT

To help users discover relevant products and items recommender systems must learn about the likes and dislikes of users and the pros and cons of items. In this paper we present a novel approach to building rich feature-based user profiles and item descriptions by mining user-generated reviews. We show how this information can be integrated into recommender systems to deliver better recommendations and an improved user experience.

Keywords

Sentimental Product Recommendation; Crowdsourcing

1. INTRODUCTION

Recommender systems help to provide users with the right information at the right time. They do this by learning about a user's interests and preferences over time and use this profile information to select and/or rank items for recommendation, preferring those that are similar to those the user has liked in the past. Traditional recommendation approaches, such as collaborative filtering and content-based techniques, rely on product ratings or meta-data. Recently however researchers have consider user-generated review content as a new type of recommendation knowledge [1].

Today's e-commerce sites are awash with reviews and sites like Amazon and TripAdvisor routinely list hundreds or even thousands of reviews alongside their products. The features mentioned in these reviews, and the sentiment expressed by real-users, represent a powerful source of objective evaluation. Dong et al. proposed [2] using this opinion information as the basis for hotel profiling. Intuitively, if a feature is mentioned by a user often, it may indicate that it is an important one. Likewise, Musat et al. [4] built a user interest profile for each user based on the topics mentioned.

In this paper we focus on a form of *personalised opinionated* recommendation by suggesting items that are not only *similar* to those a user has liked in the past, but that are *better* based on features that matter to the user. We focus on one particular use-case for this in the context of a hotel recommendation site: consider a user u_q who is traveling to a new city and wishes to book a hotel that

is *like* some query hotel h_q that she has stayed in previously. The main contribution of this work shows how we can combine similarity and sentiment to produce better personalized, recommendations than could be generated using similarity alone.

2. ITEM AND USER DESCRIPTIONS

This paper builds on recent work on mining features and opinions from user reviews for recommender system [2]. In this section, we will outline how to use this information to build rich feature-based user and item descriptions based on the features that users mention in their reviews and the polarity of their opinions.

Each item/hotel (h_i) is associated with a set of customer reviews $reviews(h_i) = \{r_1, \dots, r_n\}$ and the opinion mining process extracts a set of features, f_1, \dots, f_m , from these reviews, based on the techniques described in [2]. Each feature, f_j is associated with an *importance* score and a *sentiment* score as per Equations 2 and 3. An item description is composed of these features and scores as per Equation 1.

$$item(h_i) = \{(f_j, s(f_j, h_i), im(f_j, h_i)) : f_j \in reviews(h_i)\} \quad (1)$$

The importance score of f_j , $im(f_j, h_i)$, is the relative number of times that f_j is mentioned in the reviews of hotel h_i .

$$im(f_j, h_i) = \frac{count(f_j, h_i)}{|reviews(h_i)|} \quad (2)$$

The sentiment score of f_j , $s(f_j, h_i)$, is the degree to which f_j is mentioned positively or negatively in $reviews(h_i)$. Note, $pos(f_j, h_i)$ and $neg(f_j, h_i)$ denote the number of mentions of f_j labeled as positive or negative during the sentiment analysis phase.

$$s(f_j, h_i) = \frac{pos(f_j, h_i) - neg(f_j, h_i)}{pos(f_j, h_i) + neg(f_j, h_i)} \quad (3)$$

Similarly, we can generate a profile of a user u_q based on the reviews that they have written, by extracting features and importance information from these reviews as in Equation 4.

$$user(u_q) = \{(f_j, im(f_j, u_q)) : f_j \in reviews(u_q)\} \quad (4)$$

3. RECOMMENDATION RANKING

To begin with we implement a standard *more-like-this* approach in which we consider a query user u_q looking at some hotel h_q and requesting similar items h_1, \dots, h_c . We use h_q as the query and compare this to candidate items h_1, \dots, h_c , computing a similarity

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UMAP '16 July 13-17, 2016, Halifax, NS, Canada

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4370-1/16/07.

DOI: <http://dx.doi.org/10.1145/2930238.2930276>

score for each as the basis for ranking. Equation 5 demonstrates this for h_q and h_c , using the importance scores of shared features as the feature values.

$$Sim_h(h_q, h_c) = \frac{\sum_{f_i \in F(h_q) \cap F(h_c)} im(f_i, h_q) \times im(f_i, h_c)}{\sqrt{\sum_{f_i \in F(h_q)} im(f_i, h_q)^2} \times \sqrt{\sum_{f_i \in F(h_c)} im(f_i, h_c)^2}} \quad (5)$$

The above is a non-personalized similarity metric: the user’s profile has no bearing on the computation. We also implement a personalized version in which we use the importance weights from the query user u_q instead of the weights from h_q as in Equation 6.

$$Sim_u(u_q, h_c) = \frac{\sum_{f_i \in F(u_q) \cap F(h_c)} im(f_i, u_q) \times im(f_i, h_c)}{\sqrt{\sum_{f_i \in F(u_q)} im(f_i, u_q)^2} \times \sqrt{\sum_{f_i \in F(h_c)} im(f_i, h_c)^2}} \quad (6)$$

Next we implement a *better-than-this* approach in which we include the sentiment score for candidate item. As mentioned earlier, sentiment information is unusual in a recommendation context but it’s availability makes it possible to consider not only how similar an item is to some query but also whether it enjoys a better sentiment value; we want to recommend items that are not similar to the query and have also been positively reviewed. We do this based on a feature-by-feature sentiment comparison as per Equation 7. We can say that f_i is *better* in a candidate item h_c than the query item h_q ($better(f_i, h_q, h_c) > 0$) if f_i in h_c has a higher sentiment score than it does in h_q . Then we can calculate the sentiment score, $Sent(h_q, h_c)$ from the sum of these better scores for the features that are common to h_q and h_c as per Equation 8.

$$better(f_i, h_q, h_c) = sent(f_i, h_c) - sent(f_i, h_q) \quad (7)$$

$$Sent(h_q, h_c) = \frac{\sum_{f_i \in F(h_q) \cap F(h_c)} better(f_i, h_q, h_c) * im(f_i, h_c)}{|F(h_q) \cap F(h_c)|} \quad (8)$$

Accordingly we can implement two scoring functions based on the above as per Equation 9: (1) a non-personalized version combining Sim_h and $Sent$; and (2) a personalized version combining Sim_u and $Sent$. We can adjust the relative influence of similarity and sentiment by using the parameter w .

$$Score(q, i) = (1 - w) \times Sim(q, i) + w \times Sent(q, i) \quad (9)$$

4. EVALUATION

The dataset used in this work is based on the TripAdvisor dataset [3] which covers 148,575 users, 1,008,585 reviews, and 1,701 hotels. For the purpose of this work we use a subset of 1,000 users with at least 5 hotel reviews for a total of 11,993 reviews for 10,162 hotels. For each of these hotels, we collected their top 100 reviews to produce a larger set of 867,644 hotel reviews.

For each of these users and hotels we apply opinion mining to generate feature-based descriptions. On average our test users have written 12 reviews resulting in profiles containing an average of 91 different features. Likewise the hotels are associated with an average of 89 reviews each, and in 189 features per review.

To evaluate our recommendation approaches we produce 888 *test triples* of the form (u_q, h_q, h_t) corresponding to a query user u_q , a

query hotel from u_q ’s profile, and a target hotel visited and rated as 5-star by u_q . For each triple we use h_q (or h_q and u_q depending on approach) as an input and rank-order the other hotels in the same city as h_t , using one of the two scoring variations, varying w to adjust the mix of similarity and sentiment. We compute how often h_t is within the top-20 of these ranked hotels.

The results presented in Figure 1 show that as we increase w (that is, increase the influence of sentiment over similarity) the hit-rate of both the personalized and non-personalized versions improves. For example, at $w = 0$ sentiment is not included in the recommendation scoring and we can see that the hit-rate falls between 0.26 and 0.30; meaning that the target hotel is found in the top-20 recommendations 26%-30% of the time. As we increase w up to about 0.5-0.6 then this hit-rate increases to between 0.35 and 0.38. Beyond this value of w the hit-rate begins to fall again. This tells us that the introduction of sentiment has a positive impact on recommendation quality, up to a point. Furthermore, we can clearly see how the personalized variation outperforms the non-personalized variations, by about 20%, particularly for values of $w < 0.6$.

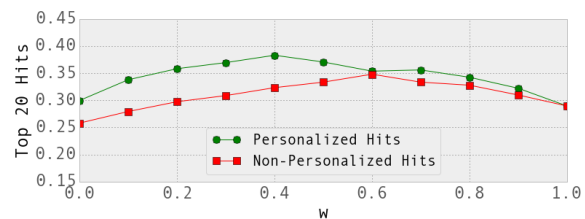


Figure 1: Recommendation hits for top 20 ranked items.

5. CONCLUSIONS

In this short paper we have outlined an approach to recommendation based on user profiles and item descriptions that are mined from user-generated reviews. We have described how this approach allows us to mix similarity and sentiment during recommendation to demonstrate the value of both factors during recommendation. Furthermore, we have shown how this approach can also be used in a personalized recommendation setting.

6. ACKNOWLEDGEMENTS

This work is supported by Science Foundation Ireland under Grant Number SFI/12/RC/2289.

7. REFERENCES

- [1] Chen, L., Chen, G., and Wang, F. Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction* 25, 2 (2015), 99–154.
- [2] Dong, R., O’Mahony, M. P., Schaal, M., McCarthy, K., and Smyth, B. Combining similarity and sentiment in opinion mining for product recommendation. *Journal of Intelligent Information Systems* (2015), 1–28.
- [3] Dong, R., O’Mahony, M. P., and Smyth, B. Further Experiments in Opinionated Product Recommendation. In *Proceedings of The 22nd International Conference on Case-Based Reasoning* (Cork, Ireland, Sept. 2014), 110–124.
- [4] Musat, C.-C., Liang, Y., and Faltings, B. Recommendation using textual opinions. In *IJCAI International Joint Conference on Artificial Intelligence, IJCAI ’13*, AAAI Press (2013), 2684–2690.