

# Aspect-based Sentiment Analysis for Social Recommender Systems

Yoke Yie Chen

IDEAS Research Institute,  
Robert Gordon University,  
Aberdeen, Scotland  
{y.y.chen}@rgu.ac.uk

## 1 Introduction

Social recommender systems provide users with a list of recommended items by exploiting knowledge from social content. Representation, similarity and ranking algorithms from the Case-Based Reasoning (CBR) community have naturally made a significant contribution to social recommender systems research [1, 2]. Recent works in social recommender systems have been focused on learning implicit preferences of users from online consumer reviews. Most online reviews contain user opinion in the form of positive and negative sentiment on multiple aspects of the product. Since a product may have multiple aspects, we hypothesize that users purchase choices are based on comparison of products; which implicitly or explicitly involves comparison of aspects of these products. Therefore, our main research question is *“Does considering product aspects importance (weight) improve prediction accuracy of a product ranking algorithm?”*

## 2 Research Aim

This research aims to develop a novel aspect-based sentiment scoring algorithm for social recommender systems. Our particular focus will be on using social content to develop novel algorithms for different product domains. For this purpose, we intend to:

1. Develop an aspect extraction algorithm to extract product aspects.
2. Develop aspect weighting algorithms to extract product aspects weights from social content.
3. Study the effect of temporal dynamics on aspect weight.
4. Evaluate the performance of our proposed algorithms in performing a top-N recommendation task using standard performance metrics such as mean average precision.

### 3 Challenges

Social recommender system harness knowledge from product reviews to generate better recommendation. Key to this task is the need for a novel aspect based sentiment analysis approach to harness this large volume of information. However, this approach suffers three main challenges:

1. Aspects extracted from product reviews using NLP-based techniques rely on POS tagging and syntactic parsing which are known to be less robust when applied to informal text. As a result, it is not unusual to have a large numbers of spurious content to be extracted incorrectly as aspects.
2. A user's purchase decision hints at the aspects that are likely to have influenced their decision and as such be deemed more important. To understand the importance of an aspect to users, it is necessary to further reveal the importance weight that users placed on an aspect. Additionally, user preferences change over time. Term frequency (TF) is the naive approach for this task where the weight of an aspect is equal to the number of occurrences of that aspect in product reviews. However, this approach is not able to capture users' preferences that change over time.
3. The absence of ground truth data causes evaluating ranking algorithm a challenging task in recommender system. For example, Best Seller ranking in Amazon can be a straightforward reference to evaluate the ranking of system generated recommendation list. However, this ranking is biased towards old products in Amazon. Therefore, there is a need to study relevant knowledge sources to construct a reference ranking for evaluation purpose.

### 4 Proposed Plan of Research

To answer our research question, our proposed plan of research is:

1. Compare the performance of our proposed aspect extraction algorithm with key state-of-the-art algorithms to determine the impact of aspect quality on recommendation tasks. We will evaluate the performance of these algorithms through accuracy metrics in extracting genuine product aspects. Thereafter, we apply the extracted aspects from these approaches in our aspect based sentiment scoring algorithm and rank the products. We then compare the recommendation performance of aspect based sentiment scoring algorithm with a sentiment analysis algorithm that is agnostic of aspects.
2. Feature selection techniques in machine learning are known to enhance accuracy in supervised learning tasks such as text classification by identifying redundant and irrelevant features. We propose to explore different feature selection techniques (e.g. Information Gain and Chi-squared) to select aspects that are important to users.
3. Our initial approach in aspect weighting algorithm places individual product aspect with equal importance weight across all products. We intend to

explore other related approaches such as TF-IDF (Term Frequency Inverse Document Frequency) to represent the importance of a product aspect. TF-IDF has been widely used in Information Retrieval community to evaluate the importance of a word to a document in a corpus. We propose to augment our aspect weighting algorithm by evaluating the importance of a product aspect to a particular product.

4. To study the effect of temporal dynamics in aspect importance weight, we look into investigating aspect weights that are inferred by:
  - **Trending information.** We would like to analyse different trending patterns of aspects occurrence in product reviews over the years (e.g. upward, downward and recurring trend). Specifically, a higher weight should be given to aspects which have an upward and recurring trend, indicating that the importance of an aspect is growing. Likewise, a lower weight should be given to aspects having a downward trend.
  - **Recency of aspects.** Aspects which frequently appear in old product reviews will have a lower weight than aspects appearing in recent product reviews. This indicates that aspects that are frequently occurring in recent product reviews are deemed important.
5. To evaluate our ranking algorithm, we use users' ratings as the baseline to compare with our proposed ranking approach. This baseline ranks each product using the average users' rating. Products in the higher rank are thus recommended.

## 5 Current Progress

Designed and developed novel algorithms in the following areas:

- **Aspect extraction.** The proposed approach integrates semantic relationship and frequency cut-off. The proposed approach was evaluated against state-of-the-art techniques and obtained positive results.
- **Aspect selection.** We address the problem of selecting important aspects using feature selection heuristics based on frequency counts and Information Gain (IG) to rank and select the most useful aspects.
- **Aspect-based sentiment scoring.** The proposed algorithm incorporates aspect importance weight and sentiment distribution. We investigated two different resources that infer the importance of product aspects: preference and time.

## References

1. R. Dong, M. Schaal, M. OMahony, K. McCarthy, and B. Smyth. Opinionated product recommendation. In *Inter. Conf. on Case-Based Reasoning*. 2013.
2. L. Quijano-Sánchez, D. Bridge, B. Díaz-Agudo, and J. Recio-García. Case-based aggregation of preferences for group recommenders. In *Case-Based Reasoning Research and Development*, pages 327–341. 2012.