

RnR: A System for Extracting Rationale from Online Reviews and Ratings

Dwi A.P. Rahayu¹, Shonali Krishnaswamy¹, Cyril Labbe², and Oshadi Alhakoon¹

¹ Centre of Distributed System and Software Engineering, Monash University, Australia

² University of Genoble, France

dwi.ap.rahayu@gmail.com, {Shonali.Krishnaswamy,
Oshadi.Alhakoon}@monash.edu.au, Cyril.Labbe@imag.fr

Abstract. In this paper we present a web based system as well as web service based application to summarise and extract the rationale that underpins online ratings and reviews. The web-based version of RnR system is available for testing from <http://rnrsystem.com/RnRSystem>. RnR system web service is available from <http://rnrsystem.com/axis2/services/RnRData?wsdl>.

Keywords: Review mining, Online Reviews/Ratings.

1 Introduction

The phenomenal growth of online social networking and Web 2.0 has led to an unprecedented increase in opinions/reviews on a wide range of topics, products and services being available and accessible both through websites or review service APIs [1]. In fact, these reviews and opinions are now a *de facto* basis and contributing factors for a range of daily activities such as buying products (e.g., electronic goods), choosing restaurants, booking hotels and planning holidays. Thus, there is an increasing reliance on online opinions for selection of product and services. This in turn is leading to an increasing focus in the area of opinion/review mining. The main aim of review/opinion analysis is to firstly identify the product/service and its key features and then to distill whether a review expresses positive/negative sentiments towards the object that is being reviewed.

In this paper, we present our RnR system that in addition to feature identification and sentiment analysis, focuses on explicating the rationale and reasoning that underpins an opinion expressed with respect to a product/service or its specific features. This can be easily justified as follows. Consider the example of a hotel which has very clean rooms, with a good view – but which is rather small in size. It is quite possible that for certain users, this hotel could be rated very positively because of the cleanliness or the view. However, it is also possible that some users have negative opinions based on the size of the room. Thus, it is important to understand what drives users to rate things differently since this makes selections based on such reviews more personalized and appropriate.

It is interesting to note that while many online reviews/opinions typically have a rating to specify the extent of positive or negative affinity for the product/service. We take the position those ratings (when available) along with textual descriptions provide a holistic representation of the opinion. Together, they combine an objective/directly measurable opinion along with a subjective/qualitative view which underpins the

rationale for the opinion [2]. We also take the view that in opinion analysis, “positive” opinions that indicate certain negative aspects or “negative” opinions that bring to the positive features are significant and worthy of highlighting. Finally, we take the position that in analyzing reviews/opinions, it is important to factor in the changing views over time. This temporal dimension captures the essential improvement or decline in the general perception of a product or service. In this paper, we present our RnR system for extracting rationale from online reviews/ratings. The system captures and summarizes the key rationale for positive and negative opinions expressed in a corpus of reviews. It highlights the negative features among positive reviews and vice versa. It also displays the changing perceptions of reviewers over time with respect to the entity that is being reviewed. We have developed as part of the RnR approach, innovative algorithms that leverage a new support metric in conjunction with a domain ontology to improve the computational overheads associated with sentiment identification. We have implemented the RnR system for a hotel review mining application. The RnR system uses reviews in the TripAdvisor.com as its corpus.

2 Architecture and Operation of the RnR System

RnR is implemented as web based system as well as web service application. Users can access the system using the web based application or embed RnR system in their websites using it as a web service. The user enters the product/service name, for which a performance summary based on online customer reviews is determined, either using RnR webpage or RnR service request. RnR main system connects to and accesses a corpus of previous reviews for the queried product/service. RnR system has a local cache for recently retrieved reviews. If the cached data is valid (in terms of being recent, as determined by a user threshold), and valid for the query, then the cached data is used rather than performing an online crawling and retrieval. Otherwise, the query is sent to an external site where online reviews are maintained (e.g. TripAdvisor.com, Expedia.com.au) via their service API. The retrieved set of reviews is then locally processed to extract the requisite rationale. The current implementation accesses TripAdvisor.com. An important feature of RnR is to identify “good” opinions in the negative group and “bad” opinions in the positive group. The RnR system highlights in grey negative adjectives within positive group and positive adjectives within negative group. The results are presented in four quadrants showing *an overall summary*, *a detailed summary of positive features (highlighting negative aspects within positive reviews)*, *a detailed summary of negative features (highlighting positive aspects within negative reviews)*, and *a chart showing temporal evolution of ratings*. Each point on the chart represents one rating given for a date of stay. The straight line within the scattered chart is the linear regression line showing the trend of performance.

References

- [1] Hu, M., Liu, B.: Mining opinion features in customer reviews. In: Proceedings of National Conference of Artificial Intelligent, pp. 755–760. AAAI Press, San Jose (2004)
- [2] Sherchan, W., Loke, S.W., Krishnaswamy, S.: Explanation aware service selection: Rationale and reputation. *Service Oriented Computing and Applications* 2(4), 203–218 (2008)