ARTICLE IN PRESS

The effect of user-controllable filters on the prediction of online hotel reviews

Ya-Han Hu[a], Kuanchin Chen[b], Pei-Ju Lee[c],*

[a] Department of Information Management, National Chung Cheng University, Chiayi, 62102, Taiwan, ROC
[b] Department of Business Information Systems, Western Michigan University, 3344 Schneider Hall, Kalamazoo, MI 49008-5412, United States
[c] Department of Information Management, National Chung Cheng University, Chiayi, 62102, Taiwan, ROC

A R T I C L E   I N F O

A B S T R A C T

Product reviews have gained much popularity in recent years. This study examines the theoretical foundation of review helpfulness and reports how the interactions among three user-controllable filters together with three groups of predictors affect review helpfulness. Reviews from TripAdvisor.com were analyzed against three analytical models. The results show that these groups of variables have a varying effect on different user-controllable filters. Review rating and number of words are key predictors of helpfulness across all three filters. The recency, frequency, and monetary (RFM) model has received a consistent support across all filters as well. Managerial implications are provided.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

With the rapid development of the Internet, online social media has become a popular platform for users to share their personal experiences. The shared information, termed user-generated content (UGC) in the academic literature, is the first source of information for many people to make their decisions [1,2]. Of all types of UGC, online consumer reviews represent the majority for purchase decisions. Online review websites are a major channel of communication that provides valuable information to consumers [3,4]. These websites also tap into online reviews for the opportunity of promotions, customer service, and other reve-nue-generating activities [5]. Studies have shown that travel websites greatly influence the tourism industry; 62% of travelers search the Internet for their upcoming travel activities and 43% of visitors read online reviews written by other travelers [6–9].

Although the growth in the volume of online hotel reviews is a welcomed trend for consumers, it also likely causes information overload for those who wish to meaningfully use it. Travelers must manually filter helpful reviews on travel websites, which considerably increases the search cost to locate hotel reviews helpful to meet their goals. This is the reason many product websites offer review helpfulness to help readers sort through a sea of reviews. Hotel reviews are not an exception. Most hotel review websites also provide some form of helpfulness indicators as well. Assessment of helpful reviews is, therefore, an important and essential task for consumers [10–12]. Review helpfulness typically refers to the total number or percentage of positive votes a product review has received; it represents a consumer's analysis of how the review matches the expectations for the trip in mind [13].

Online tourism websites with more helpful reviews can provide more valuable information to potential customers. Therefore, the development of an automatic review evaluation system to identify high-quality reviews on tourism websites can both reduce the search time for a consumer to locate the desired information, and facilitate the creation of diversified services compared with those of the existing websites. Therefore, online review helpfulness has become a key variable of interest in the product review literature that spans across multiple disciplines.

Mudambi and Schuff [12] were among the first to provide a theoretically grounded explanation of review helpfulness. They concluded with a model where the construct relationships vary between search goods and experience goods. Therefore, the requirements for a review to be considered helpful are not quite the same between the two types of goods. This is consistent with the literature where experience goods are defined as goods that require consumers to sample or "experience" the product before

* Corresponding author.
  E-mail address: pjlee@mis.ccu.edu.tw (P.-J. Lee).

formulating their own quality assessment, but quality assessment could be conducted for search goods before they are even purchased [14]. Similarly, readers of hotel reviews expect to learn from other people's experience with the hotel, which makes hotels an experience good. Although Mudambi and Schuff's model that predicts helpfulness with review rating, word count, and total votes are a great foundation model, it was not specifically designed for experience goods or more specifically for hotel reviews.

In fact, their theoretical basis of review helpfulness was information diagnosticity from Jiang and Benbasat [15] and others, which suggest that diagnosticity is highly desired when the salient product attributes are better assessed through experiences. Additionally, the accessibility–diagnosticity model indicates that "accessible information is not used as an input for judgement and choice when more diagnostic or probative information is available" ([16] Herr et al., 1991; p. 457). Therefore, factors in addition to review rating, word count, and total votes are also cues to enrich Mudambi and Schuff's model for hotel reviews.

Moreover, studies of product reviews have traditionally focused on searching for an optimal set of predictors of review helpfulness, but neglecting the fact that even the predictors may interact with each other. For example, climate and seasonal shifts may affect tourism demand [17–21], which points to a possible interaction between travel season and geographic location of hotels. This is the reason several travel websites have offered filters of hotel reviews based on these characteristics. Because of the availability of these filters, the visibility of a review may be altered through the selection of a filter. In the end, it affects a review's opportunity to be voted on for review helpfulness [11]. If interactions among predictor variables are not accommodated in a theoretical model

for hotel reviews, the predictive power or even the accuracy of the model may be hampered.

Based on the above assessment, the present study is designed with the following objectives:

1. To enrich the theoretical model of Mudambi and Schuff with additional predictor variables from the relevant literature.
2. To provide empirical evidence of interaction effects for the common filters of hotel reviews (i.e., travel regions, travel seasons, and travel types on review helpfulness).
3. To improve the performance of review helpfulness prediction models by considering the above two objectives.

2. Related work

2.1. Factors affecting review helpfulness

Table 1 summarizes the predictors of review helpfulness from the literature. The predictors used in these studies can be divided into the following three categories: review quality (i.e., review content and review readability), review polarity (i.e., review sentiment and review subjectivity), and reviewer (i.e., reviewer characteristics and RFM (recency, frequency, and monetary) features). Such a classification is rooted in both the diagnosticity and electronic word-of-mouth (eWOM) literatures. For example, Wang et al.'s [49] finding of informant credibility supports that the characteristics of the information provider (i.e., product reviewer) are related to acceptance of a product. Similarly, Li et al. [50] also

Table 1
Previous studies on review helpfulness.

Work	Data source	Search (S)/Experience (E) goods	Review quality		Review polarity		Reviewer	
			Review content	Readability	Sentiment	Subjectivity	Reviewer characteristics	RFM
Kim et al. [22]	Amazon	S/E	✔		✔			
Liu et al. [23]	Amazon	S	✔	✔	✔			
Forman et al. [24]	Amazon	E	✔	✔		✔		
Zhang [25]	Amazon	S/E	✔			✔		
Liu et al. [26]	IMDB	E	✔				✔	
Otterbacher [27]	Amazon	S/E	✔				✔	
O'Mahony and Smyth [28]	TripAdvisor	E	✔				✔	
Mudambi and Schuff [12]	Amazon	S/E	✔					
Chen and Tseng [29]	Amazon	S	✔		✔		✔	
Ghose and Ipeirotis [30]	Amazon	S/E	✔	✔		✔	✔	
Yu et al. [31]	IMDB	E	✔		✔			
Ngo-Ye and Sinha [32]	Amazon	E	✔					
Liu et al. [33]	Amazon	S	✔		✔	✔		
Dong et al. [34]	Amazon	S	✔	✔	✔			
Ngo-Ye and Sinha [35]	Amazon/Yelp	E / E	✔					✔
Hu et al. [36]	Amazon	E	✔		✔			
Hwang et al. [37]	TripAdvisor	E	✔		✔			
Yin et al. [38]	Yelp	E	✔				✔	
Lee and Choeh [39]	Amazon	S	✔				✔	
Martin and Pu [40]	Amazon/Yelp/TripAdvisor	S / E / E	✔	✔	✔			
Zhu et al. [41]	Yelp	E	✔	✔			✔	
Liu and Park [42]	Yelp	E		✔			✔	
Weathers et al. [43]	Amazon	S/E	✔				✔	
Huang et al. [44]	Amazon	S	✔				✔	
Ahmad and Laroche [45]	Amazon	S	✔		✔			
Chua and Banerjee [46]	Amazon	S/E	✔	✔				
Fang et al. [47]	TripAdvisor	E	✔	✔	✔		✔	
Hu and Chen [11]	TripAdvisor	E	✔	✔	✔	✔	✔	
Qazi et al. [48]	TripAdvisor	E	✔		✔	✔	✔	