



Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus



Andreas Munzel

Toulouse University, 1 CRM CNRS (Center for Research in Management, UMR 5303), 2 Rue du Doyen Gabriel Marty, 31042 Toulouse Cedex 9, France

ARTICLE INFO

Article history:

Received 10 March 2016

Received in revised form

15 May 2016

Accepted 1 June 2016

Keywords:

Online deception

Fake reviews

Consensus information

Identity disclosure

Trustworthiness

ABSTRACT

Current discussions in academia and in the press increase consumers' awareness of potentially deceptive online reviews. The increasing practice of fake reviews posted online not only jeopardizes the credibility of review sites as important information sources for individuals but also endangers a valuable source of information for service providers. Two studies shed further light on the role of consensus and identity-related information in assisting consumers detect potentially faked reviews. In one preliminary study, a sample of 4826 rejected and 4881 published online reviews was analyzed to investigate the differences in the disclosure of author-related information such as name and age as well as star ratings across those reviews. In the main study, a 3 (identity disclosure) x 2 (consensus) x 2 (priming of fake reviews) experiment was carried out with 390 respondents. The results highlight the relevance of the review's consensus in relation to the overall rating of previous reviews and corroborate the results of the preliminary study from the perspective of an internet user: the value of the amount of available information on the review's author in assisting individuals detect potential fake reviews. This study complements research in computer science by highlighting the relevance of contextual—in addition to textual—indicators that assist internet users in detecting potentially deceptive online reviews.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In a recent post, Forbes contributor Cheryl Conner (2013) discusses the bright and dark sides of a sector that has become a \$5 billion industry: online reputation management (ORM). As consumers increasingly interact with each other over the internet and exchange opinions and experiences on various service encounters, firms are increasingly devoting resources to monitoring, analyzing, repairing, and improving their reputations online. While many ORM firms are legitimate, providers are increasingly adopting deceptive practices, such as publishing bogus reviews. The literature refers to the practice of fake reviews—or deceptive opinion spam—as “fictitious reviews that have been deliberately written to sound authentic, to deceive the reader” (Ott, Cardie and Hancock, 2012, p. 201); estimations hold that up to one-third of all online reviews are fake (Streitfeld, 2012). Those alarming numbers and the considerable threat to the trustworthiness of review sites for consumers have led to several initiatives by regulatory institutions, review sites, and independent associations to fight the increasing prevalence of deceptive opinion spam. In September 2013, for example, 19 companies that engaged in generating and

publishing fake reviews on a New York-based fictitious yoghurt shop were fined a total of \$350,000 (Rushe, 2013). In October 2015, Amazon announced it would sue more than 1000 fake reviewers in a Seattle lawsuit (Gani, 2015). From a broader perspective, the World Economic Forum (WEC) identifies digital wildfires in a hyperconnected world as a timely major concern in its 2013 Global Risks report (Howell, 2013): characteristics such as anonymity, speed, and scale of online communication could “enable the rapid viral spread of information that is either intentionally or unintentionally misleading [...], with serious consequences” (Howell, 2013, p. 23) and deception (i.e., intentional) or misinformation (i.e., unintentional) can be deliberately propagated by those “who stand and reap some kind of benefits” (Howell, 2013, p. 25).

The perspective on different initiatives to report and fight against deceptive practices on the internet in the form of fake reviews provides evidence for the pervasiveness of opinion spam. Scholars from computer science recently increased their efforts to distinguish features of fake reviews as opposed to truthful ones. The identified characteristics included temporal patterns (Xie et al., 2012), content-related aspects such as word use (Ott et al., 2011), word duplication and number of sentences (Bhattarai, Rus and Dasgupta, 2009), and different reviewer- and behavior-related characteristics based on machine-learning methods (Jindal and Liu, 2008; Li et al., 2011; Lim et al., 2010). Ott and his colleagues

E-mail address: andreas.munzel@iae-toulouse.fr

developed algorithms and models based on their findings, transferring knowledge from linguistics to fake review detection, and published the tool Review Skeptic (www.reviewsskeptic.com). Internet users can copy and paste hotel reviews on the website to test whether the review is real or fake—with 90% accuracy.

Where these sophisticated approaches from computer science are extremely helpful for review site operators to update and improve their opinion spam detection algorithms, ORM firms are able to adapt to these recent findings and include them with the aim of generating fake—but authentic sounding—reviews for their clients.

While numerous institutions and associations attempt to reduce the number of fake reviews and fine deceptive behaviors, a considerable amount of deceptive opinion spam is still present today—Yelp recently admitted that up to 25% of reviews on its site are at least suspicious (Roberts, 2013)—which highlights the importance for consumers to develop skills to detect potentially deceitful reviews. However, whereas consumers adopt a deception-aware mindset in other market-related contexts and are aware of persuasion attempts by marketers in the case of traditional advertising, they do not adopt this mindset in the context of online reviews. As internet users expect to interact with other consumers in the customer-to-customer environment of review sites, they usually do not wear the same shields of self-protection and skepticism.

Consequently, the objective of the present study is to complement recent findings from related disciplines that adopt review-centric approaches from a consumer-centric perspective. To do so, different cues that potentially assist readers of online reviews in detecting potentially deceptive opinion spam are investigated. The amount of available information on the author of a review is a potential additional cue that helps internet users shape their perception of the review author's credibility. The disclosure of information on the author of a review has attracted interest from scholars investigating the effects of covert persuasion attempts (Campbell, Mohr and Verlegh, 2013). A large sample of 9707 reviews submitted to a multi-services review site was analyzed in an initial study. The sample contained two types of reviews: a first set of reviews ($n = 4826$) that were rejected by the site manager based on text-related reasons such as the length of the text (i.e., too short) or the inclusion of inappropriate terms. Human detectors performed the assessment of those reasons. A second set of reviews ($n = 4881$) underwent verification by site managers and were published on the review site. The results show that binary factors related to the author's identity (i.e., disclosure of age, name, and phone number) as well as the amount of disclosed identity-related information significantly separated the two sub-samples such that reviews that were actually published included a higher level of identity-disclosure than rejected reviews.

Based on the findings of the preliminary study that underline the practical relevance and validity of the cue, a second study was conducted to experimentally investigate the role of identity-related information about the review's author in readers' assessments of its trustworthiness. In addition to the available information about the review's author, two additional potential cues were investigated. Inspired by the theoretical underpinnings of the Persuasion Knowledge Model (Friestad and Wright, 1994) as well as increasing media coverage informing consumers about deceptive practices in the form of fake reviews (Tuttle, 2012), an activation of the deception-aware mindset through exposure to a news article on fake reviews was included. Second, the consensus of a reviewer's opinion with the majority is increasingly given as an indicator by the press and consumer associations to differentiate true from deceptive reviews (Johnston, 2013). Additionally, consensus information in online reviews is also attracting increasing attention from marketing research (Benedickus, 2011).

The findings highlight the importance of consensus information on the perceived trustworthiness of a review's source. Furthermore, the disclosure of information about the reviewer influences the perceived trustworthiness directly and is also moderated by consensus. Finally, the investigation of potential mediation of different deception detection cues on behavioral intentions—the individual's intention to purchase as well as his or her intention to avoid the service provider in the future—through the source's trustworthiness further supports the relevance of trust in public online environments such as review sites.

The present study contributes to the marketing literature in several ways. First, the effects of available cues to evaluate the trustworthiness of online review contributors are empirically assessed. Second, the findings complement research in computer science by adopting a consumer-centric perspective in the evaluation of online reviews. Thus far, the marketing discipline has largely neglected research on deceptive practices in the form of fake reviews. Third, the results show the relevance of perceived trustworthiness as a mediator and its effects on positively connoted outcomes (i.e., purchase intention) and negative variables (i.e., avoidance behavior). The present paper further calls for future interdisciplinary research on the effects of deceptive communication in online environments to motivate service providers to cease and desist from such malicious practices.

2. Background and hypothesis development

2.1. Persuasion and deceptive communication in online environments

The internet has not only added new potential ways for marketers to communicate to and interact with their customers, it also enables consumers to exchange information, experiences, and opinions in various domains with fellow consumers. The consequence of this opportunity to conveniently acquire credible information from sources that are at least perceived to have no commercial interest in recommending brands, services, products, and companies, is bad news for marketers: individuals trust peers much more than marketing-related sources such as traditional or digital advertising efforts (Nielsen, 2013). They even tend to avoid or block out firms' persuasion attempts via marketing communications (Kaikati and Kaikati 2004; Kimmel, 2010; Rumbo 2002). Research on today's pervasive forms of customer interactions online—usually referred to as electronic word-of-mouth (eWOM)—confirms its effects on customer behavior and attitudes (Gupta and Harris, 2010; Senecal and Nantel 2004; Zhu and Zhang 2010). However, the power of positive online recommendations from peer consumers, the potential threat of negative articulations posted in public online environments, and interactions that are largely characterized by anonymity between interlocutors have led to numerous deception attempts by marketers (Magnini, 2011; Xiao and Benbasat 2011). In an aim to circumvent consumers' avoidance and tendency to block out marketing communications, marketers pose as prior customers and publish fake positive reviews on their own services (Darke and Ritchie, 2007; Xiao and Benbasat 2011). This form of online deception fulfills the three main conditions of deceptive communication as it is currently defined (Masip, Garrido and Herrero, 2004). First, the generation and publication of fake reviews can be considered a deliberate act by the marketer and can therefore be distinguished from misinformation that is usually understood as non-intentional distortion of the truth (Xiao and Benbasat, 2011). Second, fake online reviews are the result of manipulation of information about the marketplace by impersonating a usual customer (Boush, Friestad and Wright, 2009; Xiao and Benbasat, 2011). Third, fake positive

Download English Version:

<https://daneshyari.com/en/article/7433753>

Download Persian Version:

<https://daneshyari.com/article/7433753>

[Daneshyari.com](https://daneshyari.com)