Information & Management 51 (2014) 964–971

Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im

Experimental evaluation of sequential bias in online customer reviews

Enes Eryarsoy [a], Selwyn Piramuthu [b,c,*]

[a] Management, Istanbul Sehir University, Istanbul, Turkey
[b] Information Systems and Operations Management, University of Florida, Gainesville, FL 32611-7169, USA
[c] RFID European Lab, Paris, France

ARTICLE INFO

Article history:
Received 31 March 2014
Received in revised form 25 June 2014
Accepted 8 September 2014
Available online 19 September 2014

Keywords:
Recommender systems
Online product reviews
Sequential bias

ABSTRACT

Explicit information used in the development of recommender systems includes online customer feedback. Such explicit input from customers suffers from several types of bias, which directly affect the quality of the resulting recommender system. Recent research has identified *sequential bias* to be present in online customer reviews. However, to our knowledge, no study to date has confirmed its existence in this context. Given the nature of sequential bias, confirming its existence necessitates a controlled experimental study, which is lacking in extant published literature. We attempt to address this gap. Results from our study show evidence for the existence of sequential bias.

1. Introduction

Before the existence of the Internet, a prospective customer's best option for recommendation was a (or, a few) trusted someone who has experience with an item of interest or is somehow a *maven* (an expert, relatively speaking) on such items. While this mode of operation worked reasonably well in the distant past, the existence of widely available product information over the Internet has completely changed the way in which customers seek out necessary information before they make their product/service purchase decisions.

Online customer product/service reviews are very widely used by customers and their influence cannot be overestimated (e.g., [16]). The reasons for their influence are multi-fold and include (1) their ease of availability, (2) multiple perspectives, all assembled in one location (e.g., Web page), (3) their seemingly impartial appearance, (4) the lack of other credible source(s) of information, among others. These reviews are in turn used by online retailers to generate recommendations through their recommender systems. In addition to recommending (seemingly) appropriate products/services, these recommender systems also help the customer discover new products/services through serendipity. Regardless, to a certain extent, online product/service reviews and recommender systems are slowly becoming indispensables for customers seeking information on products/services.

In an ideal world, online product reviews as well as recommendations generated by recommender systems would be perfect and free of bias of any kind. However, the reality is that both online reviews as well as the recommender systems that use them to generate recommendations are replete with several kinds of biases [21] and intentional manipulations (e.g., fake reviews that are generated by paid reviewers who may not necessarily have used the product/service before submitting their reviews online). For an excellent overview of several types of bias in general information processing, the interested reader is referred to Solloway [19] and Hu et al. [8,9].

We consider one type of bias that we believe is inherently present in online product reviews and, therefore, in recommendations generated by recommender systems that use such explicit online product reviews as a part of their input. Specifically, we consider *sequential bias* that is introduced as a direct consequence of the order sequence in which existing online reviews are perused by a new reviewer of a given product/service. Sequential bias, which is unintentional and results from existing reviews that are diametrically opposite to a new reviewer's opinion/experience of a product or service, precipitates in a reviewer submitting a review and/or review rating that is different from what was originally intended. To our knowledge, based on extant published literature, sequential bias has been studied (a) from an analytical perspective with necessary assumptions (e.g., [14]) and (b) through use of data collected at online retailer sites such as amazon.com (e.g., [17]). In the former, the authors support the possible existence of sequential bias through the existence of related phenomenon (e.g., first impression bias) and derive theoretical results based on this premise. Since the authors in the former do not show evidence

* Corresponding author at: Information Systems and Operations Management, University of Florida, Gainesville, FL 32611-7169, USA. Tel.: +1 3523928882.
E-mail addresses: eryarsoy@gmail.com (E. Eryarsoy), selwyn@ufl.edu (S. Piramuthu).

http://dx.doi.org/10.1016/j.im.2014.09.001
0378-7206/

of sequential bias in reality, the authors in the latter take it upon themselves to either prove or disprove the existence of sequential bias. They use customer review data collected at online retailer Web sites to operationalize their study. We argue that the use of such data to confirm the presence or absence of sequential bias is faulty simply because there is no control over the sequence in which existing reviews are read.

To determine the existence of sequential bias, the experimenter *has* to ensure that the perusal of reviews occur in a *sequential* manner before a new review is added. Data collected from online product review Web sites do not support this possibility. This is due to (a) some product/service review Web sites allowing users to sort reviews (based on date posted, most positive, least positive, etc.), (b) the users reading reviews in random order, and/or (c) inconsistencies in that not all new reviews are written after the perusal of existing reviews in sequential order. Clearly, based on these conditions, the methodology used in Sikora and Chauhan [17] *cannot* help evaluate the presence of sequential bias. They did *not* control for the *sequence* and their study may not necessarily be measuring sequential bias. This motivated our present study. We purport to address this gap in extant published literature by considering the dynamics of sequential bias in an online product review context through *controlled experiments*.

The remainder of the paper is organized as follows: We provide a brief discussion of some relevant publications in Section 2. This is followed by discussion of our experimental setup in Section 3. We discuss results from this study in Section 4 and discuss results from the paper in Section 5. We conclude the paper with a brief discussion on managerial and academic significance of this work in Section 6.

2. Background and related literature

We first discuss related background on recommender systems and how sequential bias could significantly affect their performance. We then discuss relevant literature that is tangentially as well as directly related to sequential bias.

2.1. Background

The knowledge-base in online recommender systems are developed using both explicit and implicit data (Fig. 1). Explicit data includes online customer reviews and implicit data are generally gathered from customers' online browsing behavior (e.g., time spent visiting an item or a Web page, previous Web page or item visited, next Web page or item visited, items product/service reviews read), items that were placed in their cart but later removed, previous purchase record, among others.

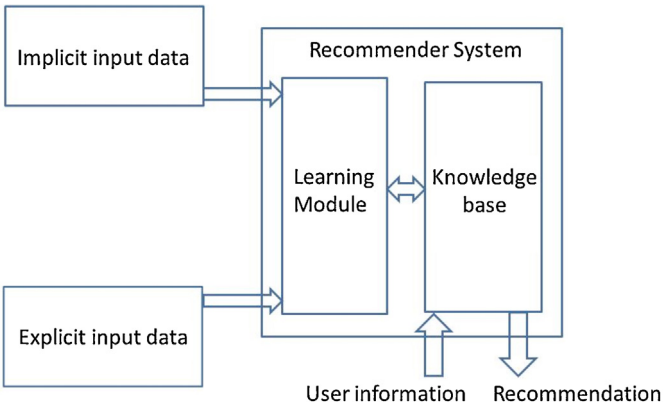

Fig. 1. Recommender system.

A recommender system consists of, at a minimum, a learning module and a knowledge-base. The learning module *learns* patterns present in input (implicit and/or explicit) data and the knowledge thus generated are stored in the knowledge-base. To prevent staleness, the knowledge-base is continually and incrementally updated with new/updated knowledge as new input data are received. Recommendations are then generated with the help of the knowledge-base and user-specific information. The user information is obtained either through knowledge of this user's past (purchase, browsing) behavior or information retrieved from the current online session (e.g., the mix of items that were considered, the amount of time spent on each considered item). Most online retailers allow the customer to create an online account, which can be used to track/trace specific customer online behaviors while at their online store. The information thus gathered can be used to fine-tune recommendations that are subsequently generated for this customer.

It is generally accepted that the more the recommender system knows about a customer and his/her interests and past behavior, the better it is in recommending appropriate products and/or services to this customer that would result in a purchase. While there are concerns with respect to privacy and security issues associated with recording online customer behavior when visiting an online retail store, these issues are completely under the retailer's control and are not as severe as in other related scenarios (e.g., the use of RFID tags to track/trace a customer by third parties even after the customer has left the Brick&Mortar store with a purchase). Nevertheless, the placement of cookies in the customer's computer has raised concerns in the past, and this continues to remain an issue. However, oftentimes there is an option on whether to accept placement of such cookies in the customer's computer. Recommender systems have other issues as well, including the cold start problem that arises primarily when a new product/service with no history in the recommender system is introduced, as well as when recommending a product/service to a new customer without knowledge of this customer's characteristics, interests, among others.

Since explicit input data in the form of online customer reviews comprise an important component of input to recommender systems, we are interested in considering online reviews written by reviewers and the process by which they write such reviews, including the inputs they implicitly and explicitly take into consideration. These inputs that go into the generation of an online review can be conscious as well as subconscious in the sense that the user may intentionally or unintentionally generate parts of the review – specifically, the numerical rating score that is assigned to a product or service of interest. In general, a new review is generated by a reviewer based on several different inputs (Fig. 2).

For example, these inputs could include the reviewer's first-hand experience with the product or service of interest, existing reviews by other reviewers, the numerical scores provided by these previous reviewers, the risk tolerance of the new reviewer, the significance of the product or service to the reviewer, any bias (e.g., sequential bias) that may be present, among others. The service/product significance indirectly refers to the degree of indifference of the customer toward this service or product. For example, while a customer may be indifferent when purchasing a pack of chewing gum, the same may not be true when this customer purchases a new car. Among these different sets of information the reviewer uses as input for providing a review, there are certainly some that evoke a stronger response from the reviewer while the reviewer may be indifferent to other sets of information. For example, when considering the purchase of a camera to take very high resolution pictures, the resolution specifications of a given camera as well as reviews and one's own