Electronic Commerce Research and Applications 20 (2016) 87–104

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

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

The impact of profit incentives on the relevance of online recommendations

Umberto Panniello [a,*], Shawndra Hill [b], Michele Gorgoglione [a]

[a] Politecnico di Bari, Italy
[b] Wharton Business School, University of Pennsylvania, United States

ARTICLE INFO

Article history:
Received 12 May 2016
Received in revised form 7 October 2016
Accepted 7 October 2016
Available online 8 October 2016

Keywords:
Consumer trust
e-Commerce
Incentive-centered design
Recommender systems

ABSTRACT

Recommender systems are commonly used by firms to improve consumers' online shopping experiences, with the secondary benefit of increased sales and profits. Prior research has demonstrated that a trade-off between relevance and profit exists, and that recommendations' manipulations and biases may hurt the credibility of recommender systems, and thus reduces customer trust. While many of the proposed designs suggest simple heuristics to bias recommendations toward higher-margin items, very little is known about consumers' reactions (in terms of purchasing behavior and trust) to recommender algorithms that balance recommendations' relevance and profitability or the drivers of this behavior. We aim to fill this gap. Data from an online randomized field experiment showed that balancing recommendations' accuracy and profit has a positive effect on consumers' purchasing behavior and does not affect their trust. We also found that the profit made during our experiment was due to a balance of several variables.

1. Introduction

Recommender systems (RSs) attempt to deliver personalized recommendations to online users, allowing them to discover new products and sort through large choice sets. This functionality is currently widely used and increasingly considered indispensable by online retailers such as Amazon, Wal-Mart, and Netflix. RSs benefit both consumers and firms. RSs allow consumers to become aware of new products and influence the selection of desirable products from a myriad of choices (Pham and Healey, 2005). They have the potential to help firms increase profit by converting browsers into buyers, presenting cross-selling and upselling opportunities, and increasing customer loyalty (Schafer et al., 2001) and response rates (Ansari and Mela, 2003). Designing an effective RS, however, is not a trivial task and has attracted attention from the computer science, information systems, and marketing research communities. Typically, an RS must use limited information about consumers to infer their preferences. Prior research has proposed and analyzed various techniques of doing this (Adomavicius and Tuzhilin, 2005).

A gap exists between existing research on RSs and their real-world use though. The vast majority of researchers have evaluated designs that use consumer-centric measures such as the relevance of recommended items. Although these designs consider user satisfaction, researchers have only implied that firms deploying these systems benefit from increased customer satisfaction and higher purchase rates by users. In reality, RS are deployed by firms whose incentives range from increasing customer loyalty to increasing profitability. The case of Netflix (Shih et al., 2007) clearly demonstrates this point. The firm operated a mail-delivered movie-rental business and now it recommends movies for streaming to users through a web-based interface. The RS strives to recommend movies with the best fit for users. However, the system had a filter that avoided recommending new releases that were in high demand and had high carrying costs. While recommending these titles might increase the accuracy of the RS, it could also lower profit. Clearly, from the perspective of profit maximization, an RS should balance accuracy and profitability. E-commerce sites such as Amazon use a different strategy since popular items may have lower unit costs due to volume discounts or lower holding costs, making them desirable for recommendation. Though specific recommender policies may vary across firms, the notion of balancing the relevance and profitability of recommendations is central. RSs licensed by vendors such as Oracle and Adobe include features that allow deploying firms to assign higher weights to high-margin items.

In addition to balancing relevance and margin, firms also need to consider the trust that consumers place in the recommendations they receive when repeatedly interacting with the firm over time.

* Corresponding author.
E-mail address: umberto.panniello@poliba.it (U. Panniello).

http://dx.doi.org/10.1016/j.elerap.2016.10.003
1567-4223/

Academics and practitioners have argued that manipulations and biases hurt credibility and erode consumer trust, leading to a net reduction in firms' profits (Simonson, 2005). This results in an inter-temporal trade-off. If a firm and a consumer interact only once, an RS may recommend products that are somewhat less relevant but have high margins, which may compensate for low purchase probabilities. However, in a long-term scenario, if design modifications focus heavily on firm-centric measures, such as profit or margin, customers may detect bias and profitability will be reduced. Thus, any attempt to incorporate item profitability into the design of an RS should consider the impact on consumer trust in the system. Although the issue of consumer response to bias in RSs is relatively new, related issues have arisen in the past, such as current business models for search engines that provide sponsored results alongside organic results (despite the famous claim of the founders of Google that sponsored results are inherently biased and cause mixed users' reactions (Brin and Page, 1998)). However, search engines (including Google) have incorporated sponsored results with no drastic impact on consumer trust or usage. While it is unclear whether the experience of search engines is applicable to RSs, the seeming comparability begs examination.

In summary, the trade-off between relevance and profit margin and between short-term profitability and long-term trust has received limited attention in existing research on RS design. Further, to the best of our knowledge, no controlled experiments have evaluated these designs. As a result, little is known about how consumers would react (in terms of purchasing behavior and trust) to recommender algorithms that account for product profitability. We aim to fill this gap with our research. In particular, our research issue investigates whether and why an RS that balances the accuracy and profit of recommendations can lead to better performance compared to traditional recommendation engines. In this paper, we report the results of a randomized field experiment designed to test the impact of a firm-centric design on website usage, purchasing behavior, and trust. For the experiment we utilized a "profit-based RS" derived from the theoretical model proposed by (Hosanagar et al., 2008). In particular, this engine was designed on the assumption that optimal recommendation policies balance product relevance and profit margin.

Specifically, when trust in an RS is high, only the balance of relevance and profit margin is considered. The optimal recommendation typically maximizes current profit solely; however, when trust is low, the optimal recommendation should be aimed at restoring trust, even if this temporarily reduces profit. We found that a profit-based RS designed according to the aforementioned policies generates higher profits for firms than a content-based RS (based only on item relevance). In a post-experiment survey of users, we found no significant reduction in customers' trust in the profit-based RS. To investigate why profit-based RSs outperform traditional RSs and since prior works have demonstrated links between purchasing behavior, user trust, and recommendation characteristics (i.e., accuracy and diversity), we also investigated the role of diversity, relevance, and trust in driving customer purchasing behavior. Through linear regression models, we were able to confirm that diversity plays an important role in driving customer purchases and exclude diversity as the main, or sole, driver behind increases in customer purchases. Accordingly, we can confirm that the profit gains during our experiment were due to the profit-based design based on a balance of trust, diversity, and relevance instead of the single effect of one of these factors.

2. Literature review

Given the potential impact of personalized recommendations for both firms and consumers, the issue of recommender design and its impact on consumer choice has been a topic of great interest. In the following, we describe these two research streams and position our work within the literature.

2.1. Recommender system design

Designing effective RSs that can infer user preferences and recommend relevant items is a challenging task. Several fields, including computer science, information systems, and marketing, have produced several studies on the topic. Surveys utilizing various approaches have been proposed and can be classified as content-based, collaborative filtering, and hybrid (Adomavicius and Tuzhilin, 2005). Content-based systems recommend items with a high degree of similarity to users' preferred items (inferred through ratings or purchases) (Mooney and Roy, 2000; Pazzani and Billsus, 2007). An advantage of using content-based designs is that even small sets of users can be addressed effectively. However, a major limitation is that a machine must be used to parse the items or items' attributes must be manually assigned (Shardanand and Maes, 1995; Balabanovic and Shoham, 1997). Collaborative filtering systems recommend items based on historical information drawn from other users with similar preferences (Breese et al., 1998). Collaborative RSs overcome some of the limitations of content-based RSs but suffer from the "new-item problem" (i.e., the difficulty of generating recommendations for items that have never been rated by users). Hybrid approaches combine collaborative and content-based methods in various ways (Soboroff and Nicholas, 1999).

Researchers have also studied how other information besides customers' demographic data, past purchases, and past product ratings can be included to improve the accuracy of recommendations (Aamir and Bhusry, 2015). One study (Adomavicius et al., 2005) described a way to incorporate contextual information into an RS using a multidimensional approach in which the traditional user $\times$ item paradigm is extended to support additional contextual dimensions such as time and location. Previous studies have demonstrated that including context can improve an RS's ability to predict behavior (Palmisano et al., 2008) and have compared different approaches to incorporating contextual dimensions in an RS (Panniello and Gorgoglione, 2012; Panniello et al., 2014; Panniello et al., 2009). Similarly, researchers have studied how expert evaluations can be used in addition to traditional user evaluations and item characteristics (Ansari et al., 2000) and how to build a multidimensional RS using multicomponent ratings (Sahoo et al., 2012).

A notable research stream has studied how RS designs can be altered to increase the diversity of recommendations. One study demonstrated that diversity is an additional important item-selection criterion and that significant gains can be made if its introduction is carefully tuned (McGinty and Smyth, 2003). Studies in this field have demonstrated that RSs that discount item popularity when selecting recommendable items may increase sales more than RSs that do not (Fleder and Hosanagar, 2009) and that increased product variety (through electronic markets) can be a significantly larger source of surplus gain (Brynjolfsson et al., 2003). The combination of diversity and accuracy of recommendations affects customers' trust, which, in turn, affects customer purchasing behavior (Gorgoglione et al., 2011; Panniello et al., 2016). Perceived diversity significantly influences users' perceived ease of use and usefulness of an RS as well as behavioral intentions and positive attitudes toward the system (Hu and Pu, 2011). In addition, studies have shown that ranking recommendations according to predicted rating values provides good predictive accuracy but poor performance with respect to the diversity of recommendations and have proposed several recommendation ranking techniques that can improve diversity (Adomavicius and Kwon, 2012; McGinty and Smyth, 2006).

Although better predictive accuracy increases consumer purchase probability, many other factors serve firms' goal of