

Using Lexical Semantic Analysis to Derive Online Brand Positions: An Application to Retail Marketing Research

Praveen Aggarwal^{a,*}, Rajiv Vaidyanathan^{b,1}, Alladi Venkatesh^{c,2}

^a Labovitz School of Business & Economics, University of Minnesota Duluth, 385A LSBE, 1318 Kirby Drive, Duluth, MN 55812, United States

^b Labovitz School of Business & Economics, University of Minnesota Duluth, 385B LSBE, 1318 Kirby Drive, Duluth, MN 55812, United States

^c CRITO (Center for Research on Information Technology and Organizations) at the University of California, Irvine, CA 92697, United States

Abstract

This paper provides an innovative approach to brand tracking in the context of online retail shopping by deriving meaning from the vast amount of information stored in online search engine databases. The method draws upon research in lexical text analysis and computational linguistics to gain insights into the structural schema of online brand positions. The paper proposes a simple-to-use method that managers can utilize to assess their brand's positioning relative to that of their competitors' in the online environment.

© 2009 New York University. Published by Elsevier Inc. All rights reserved.

Keywords: User-generated content; Lexical semantics; Brand personality; Positioning

Recent developments in e-Commerce indicate a phenomenal growth in Internet retailing (Weathers, Sharma, and Wood 2007; Yadav and Varadarajan 2005). One area of increasing attention among retailing researchers and strategists is online branding and customer behavior (Ailawadi and Keller 2004). In this context, a typical question the brand manager might ask is, "What meanings do people associate with our brand?" As it is becoming increasingly difficult to exclusively claim an attribute or quality assertion for one brand ("Volvo makes safe cars") in a market crowded with a multitude of brands, the best that brand managers can do is to put some distance between their brands and those of their nearest competitors for such claims. From a managerial perspective, it would be helpful if researchers could provide an efficient and easy-to-use method for ascertaining a brand's association with key descriptors that could also be used to gauge brands' relative performance on those descriptors. This is a critical issue in online retailing and branding research.

The objective of this paper is to propose a simple, non-intrusive way of discovering brand–descriptor associations by

exploiting the vast and continuously expanding information databases of online search engines such as Google. Using such associations, we propose a method based on lexical semantic analysis for comparing a brand's positioning against its competitors, assessing differences in how the brand is perceived, and drawing inferences about a brand's personality. This method permits us to obtain a summary assessment of the online representation of a brand, based on a set of managerially relevant descriptors. The basic premise of this paper is that there is valuable information contained in these brand–descriptor associations that can be efficiently mined using simple, semantic search-algorithms.

Background and problem statement

If managers do a simple keyword search to assess what the online world says about their brands, they will usually get back thousands of hits from a search engine. They will then face the onerous task of sifting through the voluminous data generated by those hits. Currently, the usefulness of the hit information is limited by the fact that the listings contain no summary information about the content. Nor do the listings contain information on the "semantic" properties of hits containing textual data. This raises the question of how to derive meaning from text-oriented data, especially as it relates to a neutral target such as a brand name.

* Corresponding author. Tel.: +1 218 726 8971.

E-mail addresses: paggarwa@d.umn.edu (P. Aggarwal),
rvaidyan@d.umn.edu (R. Vaidyanathan), avenkate@uci.edu (A. Venkatesh).

¹ Tel.: +1 218 726 6817.

² Tel.: +1 949 824 6625.

To obtain managerially relevant insights through web searches, we draw upon research in lexical text analysis. We propose some critical indices that can be used by managers to glimpse the structural schema of their brands in the online world by examining links between brand names and key adjectives/descriptors in these massive online databases. We propose some techniques and measures for drawing on data stored in search engines. We also discuss three studies that report “real” applications of these techniques to generate information of interest to marketers.

A recent development motivating studies of this kind is the notion of a semantic web (Berners-Lee, Hendler, and Lassila 2001; Leuf 2006). Until recently, the architecture of the web had been viewed in syntactic terms; the web is examined merely as a repository of information without interpretation or meaning attribution. The focus of search system development has been primarily on the efficient storage and retrieval of information rather than on recognizing the meaning of content. A semantic vision of the web allows for machine-based inferences, and implies that information can be processed and understood by computers, which leads to “lexical semantic analysis.” In our research, we draw on the insight that the co-occurrence of adjectives (e.g., “reliable”) and nouns (e.g., “Sony”) is a strong indicator of subjectivity (Hatzivassiloglou and McKeown 1997). Subjectivity refers to the aspect of a language used to communicate an evaluation or an opinion (Banfield 1982; Wiebe 1994). When such co-occurrences are compounded over a vast amount of textual data, reliable inferences about a brand’s online position can be obtained.

To exploit the semantic potential of the web, this paper draws on theoretical developments in computational linguistics, lexical semantic analysis, and statistics. We provide an easy to use, comprehensive methodology for studying brand associations, and for drawing meaningful conclusions from the hit counts provided by popular search engines such as Google. This methodology will allow retailers and marketers to effectively use the vast dynamic data stored in search engine databases to develop online brand strategies.

Lexical semantics—some initial ideas

Typing the name of a brand or a travel destination into a search engine will often provide thousands of hits, but the searcher does not gain information on whether these sites have positive, negative, or neutral things to say about the brand or destination (Turney 2002). To get a better insight into the meaning in the text, researchers in the fields of artificial intelligence and natural language processing are beginning to focus on evaluating the content of large text-based corpora.

One of the early approaches to textual data in the field of marketing is content analysis (Kassarjian 1977). Content analysis has been used to analyze advertisements and textually based promotional materials (Arnold, Kozinets, and Handelman 2001; Voss and Seiders 2003), to determine knowledge structures of salespeople (Sharma, Levy, and Kumar 2000), and to understand communication through home shopping networks (Warden et al. 2008). Content analysis has not been embraced in a signif-

icant way by marketers to the extent it has been in fields such as communication research and journalism. Presumably marketing research has been deterred by problems associated with sampling and measurement, as well as the reliability and validity of content categories. Also, over the years, there has been a greater push toward research driven by numeric data. In addition, content analysis requires manual coding of data from diverse sources, and this task can be quite prohibitive.

Another recent online research method gaining attention is “netnography” (Kozinets 2002). Marketing scholars have studied online reviews to uncover the dimensions of online service quality (Yang and Fang 2004), as well as online conversations as word-of-mouth communication (Godes and Mayzlin 2004). A common theme running across these studies is that authors typically restrict their data to a small subset of available discourses (typically a few hundred pages) in order to keep the process of data parsing and analysis manageable. While manual analysis of these pages results in a rich and in-depth investigation of the restricted dataset, the obvious compromise involved in this approach is the exclusion of a vast majority of available pages.

Recent developments in computerized semantic analysis have opened new possibilities in text-data analysis, and are enabling researchers to overcome the shortcomings of traditional content analysis and netnographic approaches. With developments in machine intelligence and semantic analysis software, there has been a greater scope for exploring web-based information using new techniques. It is in this spirit that our paper examines the semantic properties of brand positions.

Lexical semantics—the method

For the purpose of our analysis, we utilize two key properties of textual data: (1) consumers attach meanings to words and (2) meanings are inherent in the text, or more specifically in the adjectival expressions used by the author of the text. Understanding the valence or semantic orientation of a word (Hatzivassiloglou and McKeown 1997) can help describe its associated noun. That is, we can infer the evaluative nature of a sentence describing a noun by examining the association between the noun and the adjectives modifying that noun.

Descriptors which identify the evaluative nature of brand-related text have another use. Prior research has demonstrated a positive relationship between the presence of adjectives and the subjectivity of the sentence (e.g., Bruce and Wiebe 1999; Hatzivassiloglou and Wiebe 2000). Within the context of brand information available on the web, we can infer the overall evaluative nature of text-based, brand-related information by examining the semantic orientation of the adjectives or descriptors used in association with the brand. The goal is to get a sense of the evaluation of brands in subjective sentences (as opposed to objective sentences which merely state non-evaluative facts) by examining the association between the brand and various carefully selected adjectives or descriptors.

Researchers have inferred the semantic orientation of phrases extracted from an online database using a pointwise mutual information (PMI) algorithm. PMI essentially takes the difference between the number of associations between the target and

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات