Chapter 9

e–WOM Analysis Methods

Ioannis Stivaktakis
University of Nicosia, Cyprus

Angelika Kokkinaki
University of Nicosia, Cyprus

ABSTRACT

Electronic word of mouth (e-WOM) is rapidly becoming an empowering tool for consumers to express their experiences on services or products, on social media or other platforms. Beyond the obvious implications of such content to potential consumers, interest is also high among researchers, industry players, and other stakeholders who strive to analyze before-and-after sales expectations, emotions, and perceptions of customers. The need to find efficient ways of extracting and then analyzing online content rendered the reuse of tools and methodologies initially applied in other fields as well as the development of new approaches. In this chapter, the authors identify high-impact scientific work related to e-WOM and point out the analytical methods for analyzing e-WOM content. Furthermore, this chapter refers to the most relevant studies employing such methods and their findings. More specifically, it discusses clustering, sentiment analysis, supervised and unsupervised machine learning, lexicon-based approaches, corpus-based approach, summarization and predicting, and regression analysis.

INTRODUCTION

Extensive use of social media (SM), defined as “a group of Internet-based applications that exist on the Web 2.0 platform and enable the Internet users from all over the world to interact, communicate, and share ideas, content, thoughts, experiences, perspectives, information, and relationships” (Perez-Aranda et al. 2017, pp. 53), lead to phenomenal increase of user-generated, online content. According to Munar and Jacobsen (2014), online content is mainly generated through wikis, blogs, microblogs, media-sharing sites, review and voting sites, and social media platforms.

A special subcategory of online content is the electronic word of mouth (eWOM), which can be traced to various channels as mention in Bickart (2005) and summarized below:
e-WOM Analysis Methods

1. Posted reviews, which are consumer opinions uploaded in public or private view by merchant websites, consumer review portals, or personal webpages like blogs.
2. Mailbags, which include comments posted online by customers and readers on news and magazines websites, and manufacturers and service providers pages.
3. Discussion forums, which consist of bulletin, message and discussion boards or Usenet groups.
4. Listservs, which consist of consumer opinions shared through group member email lists.
5. Personal emails. These are shared opinions directly from one person to one or more other, through electronic mail.
6. Chat rooms, which are areas on the internet or other computer networks where people and groups have real-time conversations.
7. Instant messaging, where people converse online in real-time, usually one-to-one.

The first three communicating forms (posted reviews, mailbags, and discussion forums) are maintained online for longer periods of time; as a result, they are accessible by a potentially higher number of readers (Bickart 2005). Listservs, personal emails, chat rooms and instant messaging, are not publicly available and therefore are accessible by far fewer users (Bickart 2005). The period a message is online and the number of readers who access it are two of factors that contribute to the impact of the message onto its readers. Other factors may also affect this process. Understanding, influencing or predicting human behavior through SM and eWOM analysis, is based on handling structured or unstructured user-generated content and derive value through its transformation into meaningful knowledge.

The focus of this chapter is to identify and discuss the main methods for analyzing eWOM content. To achieve this goal, we must systematically derive all relevant literature, as described in the following section.

EWOM: SYSTEMATIC LITERATURE REVIEW

In this section, we discuss the systematic literature review on impactful resources published up until October 2018. The systematic literature review focuses on regarding analysis methods of eWOM data. Based primarily on the methodology of Shaikh and Karjaluoto (2015) and the review approach of Sarma and Choudhury (2015), this endeavor start with identification of the keywords that will be used in our base search queries. The search terms include eWOM, online reviews, electronic word of mouth, online word of mouth, online consumer reviews. Then, we establish the inclusion criteria of our literature: we focus on resources published in English, peer-reviewed journals that are Scopus indexed. Based on the aforementioned criteria, the following search query has been formulated; it resulted in 1876 articles.

TITLE (((eWOM) OR (word-of-mouth) OR (online AND reviews) OR ("online word of mouth") OR ("online consumer reviews"))) AND (LIMIT-TO (SRCTYPE, "j")) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))

Each article was examined individually with regards to the Field-weighted Citation Impact (FWCI), which compares the article’s citation count vs one of similar articles in the same field and timeframe in the Scopus database. An excel file was created with all publications, including their total citations. After eliminating redundancies, we selected the top hundred most cited articles. To reduce the risk of leaving out important literature we also included all impactful publications (i.e. those with FWCI index with a score of 10 or more). Every article was read and checked for exclusion criteria, that is, thematic and data