Online Information Review

Competitive intelligence in social media Twitter: iPhone 6 vs. Galaxy S5

Yoosin Kim, Rahul Dwivedi, Jie Zhang, Seung Ryul Jeong,

OIR
40,1

Competitive intelligence in social media Twitter: iPhone 6 vs. Galaxy S5

Received 1 March 2015
Revised 20 July 2015
16 September 2015
26 October 2015
Accepted 5 November 2015

Yoosin Kim
Chungbuk National University, Cheongju, South Korea
Rahul Dwivedi and Jie Zhang
Department of Information Systems and Operations Management, University of Texas at Arlington, Arlington, Texas, USA, and
Seung Ryul Jeong
School of MIS, Kookmin University, Seoul, South Korea

Abstract
Purpose – The purpose of this paper is to mine competitive intelligence in social media to find the market insight by comparing consumer opinions and sales performance of a business and one of its competitors by analyzing the public social media data.
Design/methodology/approach – An exploratory test using a multiple case study approach was used to compare two competing smartphone manufacturers. Opinion mining and sentiment analysis are conducted first, followed by further validation of results using statistical analysis. A total of 229,948 tweets mentioning the iPhone6 or the GalaxyS5 have been collected for four months following the release of the iPhone6; these have been analyzed using natural language processing, lexicon-based sentiment analysis, and purchase intention classification.
Findings – The analysis showed that social media data contain competitive intelligence. The volume of tweets revealed a significant gap between the market leader and one follower; the purchase intention data also reflected this gap, but to a less pronounced extent. In addition, the authors assessed whether social opinion could explain the sales performance gap between the competitors, and found that the social opinion gap was similar to the shipment gap.
Research limitations/implications – This study compared the social media opinion and the shipment gap between two rival smart phones. A business can take the consumers' opinions toward not only its own product but also toward the product of competitors through social media analytics. Furthermore, the business can predict market sales performance and estimate the gap with competing products. As a result, decision makers can adjust the market strategy rapidly and compensate the weakness contrasting with the rivals as well.
Originality/value – This paper's main contribution is to demonstrat the competitive intelligence via the consumer opinion mining of social media data. Researchers, business analysts, and practitioners can adopt this method of social media analysis to achieve their objectives and to implement practical procedures for data collection, spam elimination, machine learning classification, sentiment analysis, feature categorization, and result visualization.

Keywords Smartphone, Sentiment analysis, Opinion mining, Competitive intelligence, Market sharing, Social media analytics
Paper type Research paper

Introduction
Social media networks such as Twitter and Facebook have become an important channel for communication throughout society. They contain a huge quantity of user-generated content (UGC), including consumer opinions and sentiments. Social networking service (SNS) sites in particular are rapidly developing into crucial online information sources. People develop relationships through social media, sharing

Online Information Review
Vol. 40 No. 1, 2016
pp. 42-61

1468-4527
DOI 10.1108/OIR-03-2015-0068

sentiments, complaints, and opinions of all kinds. Building an online reputation through social media is a powerful tool for most organizations, including businesses. For this reason, many organizations have adopted social media to communicate with users, and social media analytics are widely used for research and business purposes. In particular, opinion mining and sentiment analysis, which classify people's thoughts and attitudes within text data, are frequently used to gather business intelligence from SNSs.

Researchers have undertaken analyses using natural language processing, sentiment lexicons, and machine learning algorithms in a wide variety of industries including the film industry (Kim *et al.*, 2015; Rui *et al.*, 2013), hotels (Wu *et al.*, 2010; Ye *et al.*, 2011), restaurants (Lu *et al.*, 2013; Zhang *et al.*, 2010), retail businesses (Chen, 2010), and the food industry (Kim and Jeong, 2015). Nevertheless, these studies have left several questions about competitive intelligence in social media dynamics unanswered. Chen (2010) used sentiment analysis to study the internal and external stakeholders of a business organization, but excluded competitive market conditions from their research. Although B. Liu (2010) attempted to compare two brands of cellular phone using features including pictures, the battery, the camera, and the camera's size and weight, this study remained a visual comparison of feature-based sentiment polarity. Many studies have dealt with a single case (Chau and Xu, 2012; Chen, 2010) or a specific phenomenon (Hua *et al.*, 2013; Kim *et al.*, 2014; Shen *et al.*, 2013).

Business decision makers are not merely interested in what consumers think about their own products and services; they also need to know what consumers think about the products and services offered by their business rivals. Businesses want to know how their products are perceived by consumers, and whether they have a good reputation in the market. They also want to understand their own competitive advantage, in comparison with rivals. This study employs a multiple case study approach (Yin, 2002) with two competitors in the same market. To ensure that our research questions properly compare their disparities, we suggest a hybrid opinion mining method combining lexicon-based sentiment analysis and machine learning classification, and thus investigate three measurements: social media volume (Chen and Zimbra, 2010; Mangold and Faulds, 2009), purchase intention (Rui *et al.*, 2013), and consumer sentiment (Chen and Zimbra, 2010; Pang and Lee, 2008; Rui *et al.*, 2013). The results are validated using statistical analysis with quantitative measurements to demonstrate the differences between the two cases.

We expect to make a significant contribution that can assist researchers, marketers, and business decision makers. First, this study demonstrates the feasibility of using social media mining to acquire competitive intelligence, using the multiple case study method to compare opinions about the business rivals. Next, this paper provides a guide for practitioners and researchers who wish to mine social media for competitive intelligence; the results will help us to reveal the composition of a competitive market within social media dynamics. In addition, this paper indicates ways in which future studies can discover competitive intelligence on the web by using social media as well as other types of online communication.

Related work

Social media competitive analytics

Social media and Web 2.0 technologies provide an easy and friendly way for end users to communicate (Chau and Xu, 2012; Chen, 2010). Social media is gradually being recognized as a significant resource for understanding customer sentiments, the reputation of products, willingness to buy, and customer satisfaction with products and/or services. A huge volume of online information is generated by online users, which has strong influence on the market performances of companies and their products. For this reason, many organizations have adopted social media to communicate with customers

(Kietzmann *et al.*, 2011) and researchers have studied ways to use and influence social media to improve sales performance, consumer reputation, and benefits (Fan and Gordon, 2014; Zeng *et al.*, 2010).

According to Zeng *et al.* (2010), "social media analytics is concerned with developing and evaluating informatics tools and frameworks to collect, monitor, analyze, summarize, and visualize social media data" (Zeng *et al.*, 2010, p. 14). In comparison with traditional data, social media content is much richer and contains diverse range of information. In this regard, business intelligence gleaned from social media can enable business analysts and decision makers to develop market insights into consumer behavior, discover new marketing ideas, improve customer satisfaction, and ultimately increase returns on business investments (Chau and Xu, 2012; Chen *et al.*, 2012).

Many researchers have investigated social media analytics for business intelligence. Sonnier *et al.* (2011) have tried to identify the relationship between daily sales revenue and social media by collecting online comments about the target firm, classifying user-sentiments as positive, negative, or neutral, and comparing these sentiments. Rui *et al.* (2013) studied the influence of tweets on movie sales. They showed that the valence of social media data has a significant influence on a consumer's willingness to watch a particular movie, concluding that social media analyses should eventually be used to forecast movie sales. Li and Li (2014) demonstrated that the context of tweets can reveal the relationship between consumers and product brands. They emphasized that such relationships are effective in increasing brand awareness of existing or new products, and can help to build a strong brand community. In the same vein, a multiple case study of the UK football clubs highlighted the potentiality of social media on brand management (McCarthy *et al.*, 2014).

In addition, analyzing social media data can help to obtain first-hand information about not only the market reception of own products and services but also competitors'. According to Kahaner (1998), "Competitive intelligence is a process of monitoring the competitive environment, with a goal to provide actionable intelligence that will provide a competitive edge to the organization." Competitive analysis of social media uses the publicly available data in social media, and mine strength, weakness, and opportunities in comparing with competitors. Through a case study in the pizza industry, researchers found that pizza chains actively engaged their customers in social media. They recommended monitoring of social media including competitors social media activities and establish competitive benchmarking tools (He *et al.*, 2013). Another similar study suggested a novel social media competitive analytics framework using case study involving Costco. This study provides sentiment benchmarks to provide industry-specific marketing intelligence using framework and tools for sentiment analyses (He *et al.*, 2015). As a result, most studies using social media analytics support predictive power of social media (Evangelos *et al.*, 2014).

Opinion mining and sentiment analysis

Opinion mining and sentiment analysis, as part of social media analytics, have been used to mine text data about end users (Cambria *et al.*, 2013). Techniques for mining social opinion have generated better and more accurate assessments of customer responses, enabling companies to fine tune their market strategies (Chen and Zimbra, 2010; Liu *et al.*, 2010; Lusch *et al.*, 2010). Opinion mining (including sentiment analysis) is a series of processes used to identify sentiments, affect, subjectivity, and other emotional states in text content and to turn these into meaningful information that can support decision making (Cambria *et al.*, 2013; Chen and Zimbra, 2010; Pang and Lee, 2008).

Opinion mining is usually conducted in one of two ways: either via a lexicon-based approach or a machine learning approach (Lim and Kim, 2014). The lexicon-based approach uses a linguistic resource known as a sentiment dictionary or a sentiment lexicon. In lexicon-based studies, existing lexicons, such as SentiWordNet, are frequently used, as their extensive coverage of terms (linking large number of words with human sentiments) avoids the reliability problems that characterize the manual creation of custom-made sentiment dictionaries (Hung and Lin, 2013). For example, Chen (2010) employed SentiWordNet to build a market intelligence framework for evaluating stakeholder opinions of Wal-Mart. Liu *et al.* (2010) employed sentiment measures using SentiWordNet and OpinionFinder in their study of movies and social media (Liu *et al.*, 2010). Reliable sentiment lexicons could reduce the risk of manual sentiment dictionaries (Hung and Lin, 2013).

The other approach uses machine learning methods and generates classification algorithms with linguistic features in the training data set, followed by validating the performance of the algorithms within the test data set. A classification algorithm optimized through learning processes can be used to enhance the real classifier of operational systems. The machine learning methods generally provide higher performance in classification. Various machine learning algorithms such as SVM, ANN, and naïve Bayes, have been used to classify, optimize, and make predictions based on the text document (Lim and Kim, 2014; Pang *et al.*, 2002; Rui *et al.*, 2013). The SVM method, in particular, is frequently used in opinion mining because it generally achieves better text categorization results in cases where two categories are compared.

These two approaches are not exclusive and some studies have used both techniques at the same time. For instance, a combination study of lexicon-based sentiment analysis and machine learning methods evaluated the sentiment of Facebook users and recognize the emotional change (Ortigosa *et al.*, 2014). In this hybrid approach, the researchers extended the existing sentiment lexicon for sentiment analysis and generated parameters of classifiers to detect the users' sentiment patterns. On the other hand, Rui *et al.* (2013) built an intention lexicon to select intention tweets in manual way and applied it to the support vector machines algorithm. They generated the naïve Bayesian classifier using sentiment words/phrases as well. In their research, they gathered millions of tweets about 63 movies from June 2009 to February 2010 and classified the tweets into four types: intention, positive, negative, and neutral. Another research testing various machine learning classifiers showed that SVM classification performed much better than other classification algorithms. They also proved that a small size with 2,633 unigram features achieved 81.4 percent accuracy, which is similar in performance with the best result (82.9 percent) using 16,165 terms (Pang *et al.*, 2002).

Since both approaches have their own pros and cons, we decided to combine them in this research. First, in the sentiment analysis, we employed the lexicon-based method and used the sentiment lexicon of Hu and Liu (2004) as a reliable sentiment lexicon. They identify the semantic orientations of the adjectives in WordNet and generated a sentiment lexicon currently including 2,006 positive words and 4,783 negative words. The lexicon has been released on the web for public use (www.cs.uic.edu/~liub/FBS/sentiment-analysis.html). Second, for classification of purchasing intention to purchase a smartphone, we applied an SVM classification method. We randomly selected sample tweets and manually labeled a purchasing intention to make linguistic features of machine learning algorithms. We believe that our hybrid method in this study is a novel innovation for mining competitive intelligence using text mining and sentiment analysis.

Competitive intelligence in social media

Competitive intelligence in social media is still a challenging area. Fan and Gordon (2014) have explained how social media analytics can support competitive intelligence by helping organizations understand their suppliers, competitors, environments, and overall business trends. In particular, the comparison with rival companies can help a firm discover a rich, in-depth competitive strategy hidden within social dynamics. Hu and Liu (2004) conducted a study that mined and summarized customer reviews of digital cameras, phones, MP3 players, and DVD players, identifying the opinion words used. Liu *et al.* (2005) proposed an opinion comparison system that applied sentiment analysis and revealed opinion gaps between competing digital cameras. These articles used various measurements to demonstrate the benefits of comparing opinions and sentiments related to competitor firms. Our focus in this research is on three social media analytical factors: social media volume, purchase intention, and consumer sentiment. The volume of UGC is the simplest and clearest variable in social media analysis, the purchase intention involves a consumer's willingness to buy a product/service and sentiments reveal the polarity between users' assessments of the reputation of a target object. Thus we propose the following three research questions to explore competitive intelligence for competing with rival firms in the same industry:

RQ1. Is there significant gap in social media volume between competitors with different market shares?

The market share gap between revenue and shipment amounts maybe reflected in a social media content disparity expressed through the direct comments of consumers. We verify whether a volume gap exists between competitors, reflecting market share:

RQ2. Is there a significant discrepancy in levels of purchase intention between competing brands?

In consumer-generated social media content, many clues reveal authors' opinions; these include their interests, willingness to buy, sentiments, and complaints. In particular, willingness to buy is a critical piece of information for a business; identifying instances of purchase intention mentioned in consumer posts, and comparing them to competitor data could provide useful insights. To identify a customer's willingness to buy, we have used a machine learning method: support vector machines classification:

RQ3. Is there a significant gap in consumer sentiments about the two rivals?

By examining and comparing consumers' sentiments toward the two rivals, we hope to gain invaluable insights into the nature of consumer reputation. For this question, we employ lexicon-based sentiment analysis.

To answer these three questions, we will apply both opinion mining and statistical analysis (see Figure 1).

In the proposed approach, we have followed the guidelines for using a multiple case study prescribed by (Yin, 2002), and have selected two competing companies and their flagship products in the US smartphone market: the Apple iPhone6 and the Samsung GalaxyS5. The mobile market in the USA has already reached 74.9 percent penetration, but the high-end product known as a "smartphone" is rapidly overtaking customers' old-fashioned mobiles (ComScore, 2015). The smartphone market has therefore become one of the world's most competitive fields, and sales performance is influenced not only by product quality, but also by online reputation provided by customers. For this reason, a huge volume of word-of-mouth promotions have been generated on the web,

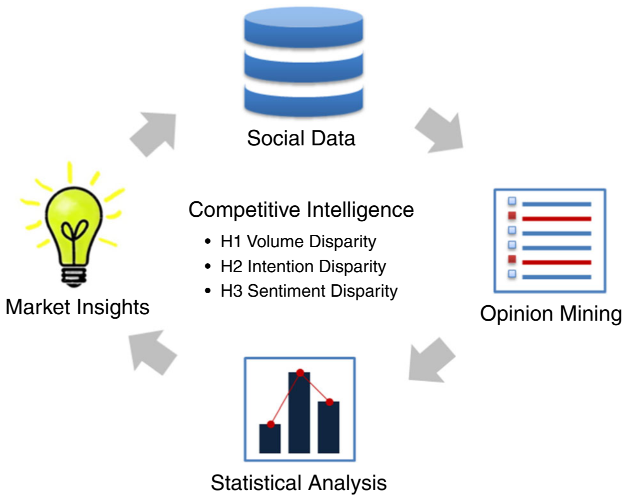

Figure 1.
Overview

with consumers asking, answering, recommending, judging, criticizing, and complaining. This demonstrates why the smartphone market is now considered an online information source as well as a research subject.

To investigate competitive intelligence in the smartphone market, we analyzed two competing high-end smartphones in the US smartphone market: the Apple iPhone6 and the Samsung GalaxyS5. According to the US smartphone subscriber market share report (ComScore, 2015), 182 million people in the USA had already acquired smartphones at the end of December 2014, reaching 74.9 percent penetration of the mobile market. This report also mentions that Apple is ranked as the top smartphone manufacturer in USA with 41.6 percent of US smartphone subscribers, while Samsung is in second place, with a 29.7 percent market share although globally Samsung is No. 1, with 24.7 percent of market share in the fourth quarter of 2014, as compared with Apple's 15.5 percent share (*Strategy Analytics*, 2015). Samsung is competing strongly both in the US market and globally. Apple supplies a small number of smartphones such as the iPhone6, iPhone6 plus, and iPhone 5, while Samsung's smartphones including the Galaxy series and Note series constitute less than 30 percent of the company's product portfolio.

Data

There are different kinds of social media platforms, with diverse UGC published and disseminated by people reading, sharing, and adding their opinions on these platforms. In particular, Twitter is one of the most popular social network services, with 255 million monthly active users worldwide sending 500 million tweets per day. It also appears to be a very credible source of information about people who are buying things (Rui *et al.*, 2013). Twitter has become an effective communication channel for spreading WOM content about brands among various types of consumer (Kim *et al.*, 2014). For this reason, we chose Twitter as our target social media source, and collected relevant tweets.

After targeting social media and various competitors, we became concerned with the problem of how to gather genuinely meaningful tweets, directly related to the iPhone6 and GalaxyS5. In a previous study, which collected four million movie tweets using a movie title query, 90 percent of the total tweets included irrelevant mentions (such as "suffering from a hangover" when the film was, *The Hangover* (Rui *et al.*, 2013). It was therefore necessary to improve this method in order to avoid irrelevant tweets. As a solution, we applied the hashtag "#" as a more direct way of finding tweets mentioning the iPhone6 and GalaxyS5. Using the hashtag enabled us to select more accurate and relevant tweets directly mentioning the target keywords in the Twitter. We set up the search keywords with "#iPhone6" and "#GalaxyS5." In order to collect tweets, we employed Twitter API, TwitteR package, in the R software program. This API can be obtained by executing the command, install.packages("twitteR") in the command console of R. In addition, in order to use this API, a researcher needs Twitter Apps Keys and Access Tokens, which can be obtained via creating developer account at the Twitter's Developer Center (https://dev.twitter.com/). Our program ran every day to collect relevant tweets. A tweet data include various pieces of information such as the author id, statusSource, screenName, created time, isRetweet, retweeted, and favorite status. This information can help researchers recognize critical clues and retrieve further data qualifications. Using this twitteR API, we collected a total of 229,948 tweets about the two smartphone rivals over a period of four months, from August to December 2014.

Pre-processing data

After collecting large number of tweets, we randomly selected and performed manual check on some of the tweets to test their usability. Despite using hashtag, collected data had many irrelevant tweets including promotions and advertisements. To find real consumers opinion, we removed commercial tweets from twitter data. For this purpose, we therefore devised a data cleansing process to select only qualified tweets containing pure consumer opinion. As shown in Figure 2, the pre-processing procedure of tweets involves removing retweets, tweets from marketing agencies and commercial accounts, and tweets containing advertising terms.

Since most retweeted tweets do not involve users' own opinions, the very first step in the data cleansing process is to remove retweet data. To accomplish this step, we queried whether the "isRetweet" data column was "False" and retained only these tweets. Our original data consisted of 229,948 tweets with 142,489 references to the iPhone6 and 87,459 to the GalaxyS5. Once retweets had been removed, 197,208 tweets remained.

The next task is removing commercial tweets by a predefined procedure. To determine whether a tweet was irrelevant, or not we checked three columns of tweet data: statusSource, screenName, and text. Under "statusSource," we found that some sites, such as Hootsuite, twitterfeed, and IFTTT were repeatedly publishing the same commercial tweets (see Table I).

By identifying the URLs for these tweets, we could tell that these were social marketing agencies supporting their respective business clients. We made a list of them, loaded it into the R program, compared each tweet's "statusSource" with the list,

Figure 2.
The procedure for pre-processing tweets

and eliminated such tweets from marketing agencies. The below R source code is a part of the cleansing program.

```
adMkt = load (agencyList)
for (stopwords in adMkt)
{
      tweet_step = tweet[!grepl(stopwords, tweet$satusSource, ignore.case = T),]
      tweet = tweet_step
}
```

The screenName also provided a clue which helped us remove garbage tweets. We found that there were a number of company names (e.g. NxtiPhoneNews, mobilephones2go, VZWDeals), official accounts (e.g. SamsungMobileUS, UK_Telecom, ATTFanPage), and media content makers (e.g. ARY_NEWS_4U, EasyAccOfficial). Since these did not express the willingness to buy or the opinions of real consumers, we removed tweets connected with these names. The last task involved detecting tweets that contained advertising words and phrases, such as "give away," "win an iPhone6," or "visit my shop." Similar to the process mentioned above, we made commercial content as stop-words and eliminated these tweets. After taking these four rigorous steps for data cleaning, we obtained 24,026 tweets written by real consumers. Table II summarizes how the irrelevant tweets were removed and how many tweets were left. As a result, 90 percent of the gathered tweets were eliminated by our filtering process, illustrating the fact that too many commercial tweets are circulating through social media.

The visualization output corresponding to Table II, as shown in Figure 3, presents the volume of irrelevant tweet types in each phase of the removals. We can also see that a high proportion of tweets are from social media marketers about the two smartphones.

In summary, data pre-processing consists of a set of activities to remove garbage data and access genuine social media data from real consumers. We defined a procedure for pre-processing tweets, and finally reduced the total number of 229,948-24,206 tweets containing 18,283 iPhone6 tweets and 5,743 GalaxyS5 tweets.

statusSource	Tweet mention content
twitterfeed	#APPLE #IPHONE6 – Shockproof Armor Back Protection Case Cover for iPhone6 Plus Silver flash: […] http://t.co/PZkrjfmzoO #Deals Ebay CA
Hootsuite	New Ghost Camera Halloween App http://t.co/24b4kQiOOq https://t.co/vEc1WpRHOH #scary #app #horror #iphone6 #04203631
IFTTT	Samsung Galaxy S5 Real Bavarian #Deerskin Sleeve ǀ XAVER ǀ http://t.co/uijxGlZGA2 ǀ #GalaxyS5 #GalaxyS5Sleeve #GalaxyS5Leather #GalaxyS5Case

Table I.
Sample tweets from social media marketing agencies

Phones	Total tweets	Removed retweets	Removed advertising sources	Removed advertising names	Removed advertising terms	Proportion of tweets remaining (%)
No. of iPhone6	142,489	113,705	22,282	21,784	18,283	12.83
No. of GalaxyS5	87,459	83,503	7,901	6,508	5,743	6.57
Sum	229,948	197,208	30,183	28,292	24,026	10.45

Table II.
Number of tweets remaining after pre-processing

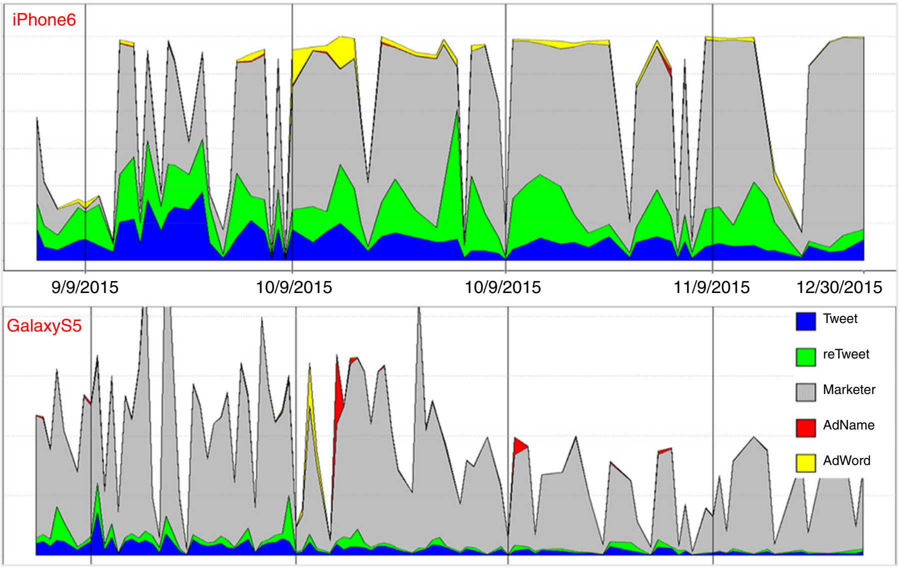

Figure 3. Tweet volume flow through pre-processing

Opinion mining and results

As mentioned previously, opinion mining that includes sentiment analysis involves recognizing and extracting an author's interests, emotions, and sentiment polarity from a piece of text (Cambria *et al.*, 2013; Chen and Zimbra, 2010; Pang and Lee, 2008).

Lexicon-based sentiment analysis

In the lexicon-based approach, sentiment analysis requires linguistic resources such as a sentiment dictionary. For the sentiment classification of smartphone tweets, we used the sentiment lexicon of Hu and Liu (2004), since it involves a relatively small number of categories that are reliable, and suitable for applying to a huge volume of tweets. We first loaded the lexicon onto our sentiment classification program in R and then matched every linguistic feature of each tweet with the sentiment terms. The appearance of sentiment features in each tweet was counted as a positive/negative point, and the sentiment of a tweet was determined by identifying the prevailing opinion, whether positive, negative, or neutral.

IF Sum(appr_Pos_term) > Sum(appr_Neg_term) THEN
Twt(i)senti = Positive(1);
ELSEIF Sum(appr_Pos_term) < (appr_Neg_term) THEN
Twt(i)senti = Negative(0);
ELSEIF Sum(appr_Pos_term) = (appr_Neg_term) THEN
Twt(i)senti = Neutral(0.5);
END IF;

Table III shows sample tweets with a positive/negative sentiment. A positive tweet is a tweet in which the author has expressed a positive sentiment toward the phone and a negative tweet is a tweet in which the author has expressed a negative sentiment toward the phone. Neutral tweets do not belong to either category.

Table IV summarizes the results of the lexicon-based sentiment classification. About 30 percent of the tweets are generally positive, and 12 percent negative. The proportion of neutral tweets is therefore relatively high; perhaps, as tweets can only be 140 letters long, users do not have much space to express sentiments.

In addition, we generated a daily sentiment point, merging the sentiment of each tweet so that it can be used for purchase intention classification and statistical analysis for conducting further analysis in the next steps.

Purchase intention classification

During this phase, we tried to determine whether each tweet expresses an intention to purchase the iPhone6 or a GalaxyS5. For example, users mentioned their willingness to buy in tweet content such as "I can't wait to get a new phone. #GalaxyS5," "finally ordered my new iphone6!" and "getting my new #iPhone6 Friday finally!" Other tweets revealed the satisfaction of users: "just switched to the #GalaxyS5 […] Loving it so far!" and "Happy! I got this iPhone6." We defined their purchase intention by identifying their willingness to buy, as well as their pleasure after buying.

To identify purchase intentions, we applied an SVM classification method. We first randomly sampled 400 tweets and manually tagged them with a purchasing intention label: true (1) or false (0). Thereafter, we broke up the text content into a word unit matrix using a parser of "TM," in the NLP package of R project after selecting the most frequently used terms to set the SVM features. The featured words included very favorable explicit words such as "amazing," "awesome," "beautiful," and "brilliant" as well as implicit terms such as "can't wait," "waiting," "dreaming," and even "hahaha." We demonstrated the tenfold cross-validation of SVM classifications in various conditions, including the varying sizes of feature terms, their frequency, and the frequency of sentiment words. In this experiment, we applied the SVM algorithm package "e1091" in R. As a result, our classification achieved the best possible performance: 74 percent accuracy in relation to 131 features and their frequency. Thus, we classified the purchasing intention of these tweets using the classifier. Finally, we recognized 6,014 intention tweets (true), including 4,479 intention tweets for iPhone6 and 1,535 intention tweets for the GalaxyS5 (see Table V).

Sentiment	Tweet mention
Positive	Maybe I'm mad but I'm now the proud owner of a potentially #bendy #iPhone6, it's so much bigger than the #4s
	Finally got to see an iPhone 6 today. Not revolutionary at all but it's absolutely gorgeous. (And I want one). #iPhone6
Negative	I'm not sure I want it. It's too big to fit in my back pocket! lol #iphone6
	I'm really disappointed with the #iPhone6. It took them 2 years to change the screen & size. Let down

Table III. Sample tweets of sentiment analysis

September 1-December 31, 2014	iPhone6 ($n = 18,283$)	GalaxyS5 ($n = 5,743$)	Total ($n = 24,026$)
Sentiment			
Positive	5,413 (29.61%)	1,919 (33.41%)	7,332 (30.52%)
Neutral	10,591 (57.93%)	3,172 (55.13%)	13,763 (57.28%)
Negative	2,279 (12.47%)	652 (11.35%)	2,931 (12.20%)

Table IV. Summary of the opinion mining results

Categorizing smartphone features

We categorized smartphone features including the display, battery, and camera because we anticipate that these specific smartphone features could help us to understand and provide in-depth insights for end users. Liu (2010) compared two cell phones that shared various features, such as pictures, batteries, cameras, size, and weight. Wu *et al.* (2010) categorized major hotel features, including rooms, service, and price, providing visualizations. Phone size, display, design, memory and storage, battery, camera, OS, prices, and advanced technologies were all mentioned as critical features of the iPhone6 and GalaxyS5 (Komando, 2013; Riaga, 2014). We defined six categories: display, battery, camera, storage, OS, and price. To determine the categories relevant to each tweet, we used a smartphone feature lexicon containing words such as big, small, charger, selfie, gigabyte, and expensive (see Table VI).

Since a tweet can mention several features at the same time, we allowed for the duplication of features in a tweet. Table VII shows the result of this sub-categorizing and reveals that online users mentioned OS, display, and cameras more frequently than other topics.

Statistical validation

In this section, we describe statistical analysis corresponding to our three research questions for determining whether iPhone6 and GalaxyS5 have significant disparities with respect to volume, purchase intention, and sentiment.

Table V. Result of intention classification

September 1-December 31, 2014	iPhone6 ($n = 18,283$)	GalaxyS5 ($n = 5,743$)	Total ($n = 24,026$)
Purchase intention			
True (1)	4,479 (24.50%)	1,535 (26.73%)	6,014 (25.03%)
False (0)	13,804 (75.50%)	4,208 (73.27%)	18,012 (74.97%)

Table VI. Sample of a smartphone feature lexicon

Feature	Lexicon
Display	Screen, display, size, big, small, slim, large
Battery	Battery, charger, charging, no charging
Camera	Camera, picture, photo, selfie, filter
Storage	Gigabyte, memory, storage, 16 gb, 64 gb, 128 gb
OS	ios, ios8, ios801, ios81, ios7, android, lollipop
Price	Price, prices, pricing, expensive, cheap, cost

Table VII. Summary of sub-categorizing results

September 1-December 31, 2014	iPhone6 ($n = 18,283$)	GalaxyS5 ($n = 5,743$)
Phone features		
Display	718 (3.93%)	191 (3.33%)
Battery	254 (1.39%)	179 (3.12%)
Camera	620 (3.39%)	255 (4.44%)
Storage	224 (1.23%)	21 (0.37%)
OS	1,027 (2.62%)	429 (7.47%)
Price	178 (0.97%)	37 (0.64%)

Volume disparity

To test our first research question of whether the rivals had a significantly different volume of tweets, we used a *t*-test to evaluate the daily volume of each competitor. In Table VIII, we can see that the daily volume mean for the iPhone6 ($m = 268.87$) was almost four times larger than that for the GalaxyS5 ($m = 61.10$), creating a large gap.

The time-series visualization graph in Figure 4 clearly presents the difference between the two companies. The graph shows that the tweet volume related to the iPhone6 is generally higher than that of the GalaxyS5, having increased rapidly after the launch of the iPhone6. During this time, a number of events affected the results; obviously, the iPhone6 was a newsmaker for a while.

To confirm our results, we conducted other investigations to confirm that our tweet volume comparison was reasonable. In Figure 5, Google analytics showed a similar result in web search trends about the iPhone6 and GalaxyS5.

The movement of stock prices for Apple Inc. as shown in Figure 6 shows high positive shifts in response to the announcement of a brand new Apple product. An explosion of interest in the iPhone6, as shown by the volume of tweets and by search trends, seems closely related to the stock price ascent of Apple Inc. This graph also shows that Bendgate, a very negative event, may have caused stock prices for Apple to decline for a short time. Fortunately for Apple, this event has been viewed as a caution for users rather than an iPhone6 defect and it faded quickly.

Interestingly, the rumor that the iPhone6 was bending emerged on September 22 and was exaggerated until September 24; Google search trends shows that "Bendgate" occurred on September 23 and the stock market price for Apple fell on September 24

Volume	Mean	SD	t	p-Value
iPhone6	268.87	190.80	−8.748	0.000**
GalaxyS5	61.10	51.94		

Note: **Significant at 1 percent level

Table VIII. Results of the volume *t*-test

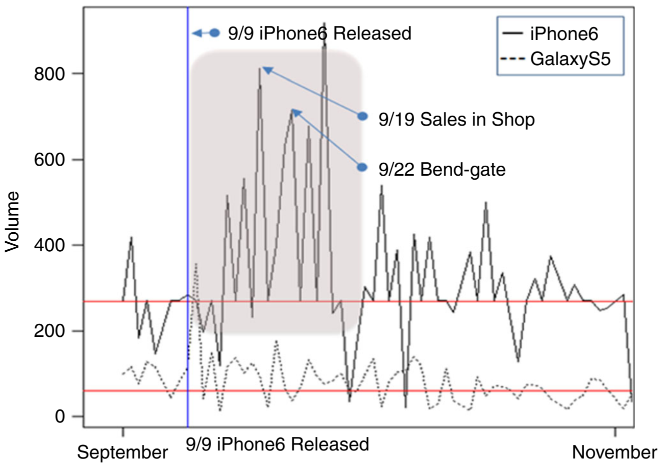

Figure 4. A comparison of tweet volumes

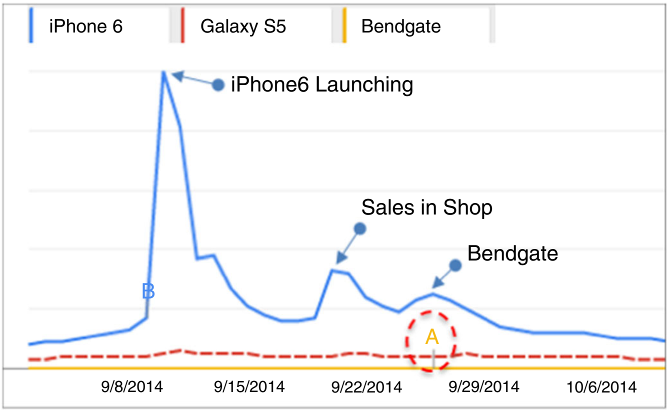

Figure 5.
Search trends in Google.com

and 25. Although conducting an in-depth comparison of these events is not the focus of this paper, it is a useful insight suggesting that information disseminated through social media spreads one or two days ahead of other media, such as web search engine trends, traditional news media, and stock markets.

Intention disparity

To validate our second research question which involve identifying the gap in purchase intention between the two rival companies, we used a *t*-test to assess the daily intention score for the two smartphones. Purchase intention was classified as either true (1) or false (0) and we measured the daily intention score, referred as the ratio between purchase intention tweets. Table IX shows that the rivals received significantly different intention scores (p-value < 0.10), although the mean for GalaxyS5 ($m = 0.275$) is slightly higher than iPhone6 ($m = 0.248$). One possible reason for this result is that consumers generally hesitate before purchasing innovative expensive product. Our tweet data clustered around the iPhone6 launch, with many customers confused by its size and Bendgate. Eventually, end users increasingly started liking the new iPhone6 after the design faults were corrected.

In addition, we conducted a χ^2-test with smartphone features as sub-categories. The battery, camera, and OS are linked with significant gaps in intention (see Table X). The GalaxyS5 was more highly rated than the iPhone6 in the battery and camera categories; however, iPhone had a higher OS rating. Indeed, the iPhone series was known to have a short battery life, while the GalaxyS had a strong record. The GalaxyS5's waterproof technology also made a favorable impression (e.g. one tweet said, "so excited how my Galaxys5 can film underwater").

Sentiment disparity

Our third research question involved consumer sentiments, and was assessed through a daily sentiment score ranging between extremely negative (0) and extremely positive (1). We found that, there was no significant sentiment gap between the two smartphones (see Table XI). Both smartphones attracted positive sentiment (iPhone6, $m = 0.591$; GalaxyS5, $m = 0.605$).

Samsung's GalaxyS and Apple's iPhone are the leading brands in the US smartphone market, as well as in the global market. Thus we can assume that, in both cases,

Figure 6.
Stock price movements of Apple Inc. in October 2014

the combination of product quality and "brand loyalty," created positive sentiment. For this reason, the sentiment flow of each of the two smartphones maintained a constant movement in the positive arena. An exception to this pattern was iPhone6, which dropped into negative arena for a short time after the "Bendgate" accident, but sentiments soon returned to normal levels. Ultimately, this event had no influence on iPhone6 sales.

The social opinion gap and sales performance

Here, we attempted to use social media analytics to explain the gap in market share between the rival companies. We estimated the level of shipments of two products in the USA by market research reports, articles, and official announcements. The iPhone6 seemed to have sold about 10.34 million units and the GalaxyS5 2.56 million during the fourth quarter of 2014 in the US market (CNet, 2015; IDC, 2014; KGI, 2015; Localytics, 2014; *Strategy Analytics*, 2015). We thus invented a new indicator by multiplying the mean volume of daily tweets by the purchasing intention score. Finally, we compared our market research with the indicator (see Table XII).

Table IX.
Results of the intention t-test

Intention	Mean	SD	t	p-Value
iPhone6	0.248	0.056	1.962	0.052*
GalaxyS5	0.275	0.117		

Note: *Significant at 10 percent level

Table X.
Intention comparison in relation to specific smartphone features

No.	Feature	iPhone6	GalaxyS5	χ^2	p-Value
1	Battery	55 (21.7%)	60 (33.5%)	7.58	0.006*
2	Camera	108 (14.4%)	68 (26.7%)	9.616	0.002*
3	Price	33 (18.4%)	8 (21.6%)	0.186	0.664
4	Display	141 (19.6%)	40 (20.9%)	0.161	0.688
5	OS	190 (18.5%)	63 (14.7%)	3.068	0.080**

Notes: *,**Significant at 1 and 5 percent level, respectively

Table XI.
Results of the tweet sentiment t-test

Sentiment	Mean	SD	t	p-Value
iPhone6	0.591	0.040	1.644	0.102
GalaxyS5	0.605	0.066		

Table XII.
Opinion mining gap in social media

4Q, 2014	Global	USA	Volume(a)	Intention(b)	$a \times b$
iPhone6(i)	43.09 M	10.34 M	268	0.25	66.46
GalaxyS5(g)	40.0 M	2.56 M	61	0.28	16.78
x-times(i/g)	10.77	4.04	4.39	0.90	3.96

Interestingly, the result shows that the social indicator (3.96) is very similar to the shipment gap (4.04). Although their similarity is based on estimation from the two case studies, it can provide competitive intelligence for market insights.

Discussion and implications

This study employed a multiple case study approach for competitive intelligence in social media and suggested a hybrid opinion mining method combining lexicon-based sentiment analysis and machine learning classification. Our results show the feasibility of the proposed method and thus provides not only the theoretical reference but practical guide to researchers, marketers, and business users as well.

We believe that researchers and practitioners can consider hybrid opinion mining method for carrying out social media analytics to discover market sensing and gaining business insights via business intelligence. The procedure for the proposed method describes how to collect social media data, remove garbage data, conduct sentiment analysis, categorize product features, and classify purchase intention. Furthermore, our results of statistical analysis provide empirical evidence that competitive intelligence from social media explain the consumer behavior, opinions, and competing environment and it can even predict sales performance gap between competitors (see Table XIII).

Business analysts and practitioners who wish to mine competitive intelligence in social media can apply our method and process not just within the smartphone industry but in other industries as well. Findings from this analysis prove that a company can understand the consumers' opinions not only toward its own product but also toward products from competitors' through social media analytics. If a company builds the social media monitoring system for market sensing and competitive intelligence, it can investigate consumers' feedback toward its products and services in real time. In addition, the firm could predict its performance in terms of market sales and estimate the gap of its products or services with competing products or services.

	iPhone6	GalaxyS5	Note
Launched	September 9, 2014	April 11, 2014	
Estimated sales in 4Q	10.34 million units in the USA	2.56 million units in the USA	SA, IDC, KGI, CNET, Localtyics
Market status	Apple, No. 1 in the USA with 41.6% market share	Samsung, No. 2 in the USA with 29.7% market share	December 2014, comScore Report
Volume of tweets	From 142,489 tweets to 18,283 (12.83%) after pre-processing	87,459 tweets to 5,743 (6.57%) after pre-processing	From August to December 2014
	iPhone6 ($m = 268.87$) has a significantly larger volume of tweets than the GalaxyS5 ($m = 61.10$)		t-test, $p < 0.01$
Purchasing intention	Expression: 4,479 None: 13,804	Expression: 1,535 None: 4,208	SVM classification
	The GalaxyS5 ($m = 0.275$) ranks higher than the iPhone6 ($m = 0.248$) in the total set		t-test, $p < 0.10$
	Significant gaps on battery*, camera*, and OS** features		χ^2, $p < 0.01$*, < 0.10**
Sentiment analysis	Positive: 5,413 (29.61%) Neutral: 10,592 (57.93%) Negative: 2,279 (12.47%)	Positive: 1,919 (33.41%) Neutral: 3,172 (55.13%) Negative: 652 (11.35%)	Lexicon-based approach
	No significant gap		t-test, $p = 0.102$
Comparing indicators	The sales volume gap in the US market = 4.04 times The opinion mining gap in tweets = 3.96 times		iPhone6/GalaxyS5

Table XIII. Summary of analysis results

Thus decision makers can compensate for the weaknesses in their own products by contrasting them with rival products/services and adjusting the overall market strategy much more quickly in comparison with traditional ways.

Conclusion and future research

Since people develop relationships through social media and shares sentiments, complaints, opinions of all kinds, social media networks such as Twitter and Facebook have become an important channel for market sensing. This study aims to mine competitive intelligence in social media to find the market insight by comparing consumer opinions and sales performance by analyzing public social media data of a business and its competitors. As the multiple case study approach, we compared two competing smartphones, the iPhone6 and the GalaxyS5, through opinion mining followed by statistical validation of the results. The result shows that social media data contain competitive intelligence. The volume of tweets revealed a significant gap between the market leader and one follower. The purchase intention results also reflected this gap, but to a lesser extent. In addition, the social opinion gap between the two smartphones was also similar to their sales performance gap.

We believe that this work can easily be extended to include various social media platforms and services, including Facebook, LinkedIn, Instagram, and blogospheres. Each media organization has its own characteristics and users can also be distinguished through features including age, gender, area, interests, and preferences. In addition, future studies should focus not only major products but also minor competitors because competitive intelligence could be used to quickly identify new entrants and emerging stars in the market. For example, Xiaomi, a smartphone manufacturer in China, has a meteoric rise over the last couple of years. Lastly, we expect future studies to focus on refining social media data in various different ways because social media typically contains not only formal communications (news, columns, and critiques) but also informal chats (including small talk, jokes, and slang). In particular, informal social media conversations have coined new terms and even emoticons. To extract real consumer opinion and sentiment from the enormous quantity of social media data, we should improve data processing techniques, sentiment lexicon, classification methods, and empirical research.

References

Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013), "New avenues in opinion mining and sentiment analysis", *IEEE Intelligent Systems*, Vol. 28 No. 2, pp. 15-21. doi: 10.1109/MIS.2013.30.

Chau, M. and Xu, J. (2012), "Business intelligence in blogs: understanding consumer interactions and communities", *MIS Quarterly*, Vol. 36 No. 4, pp. 1189-1216.

Chen, H. (2010), "Business and market intelligence 2.0, Part 2", *IEEE Intelligent Systems*, Vol. 25 No. 2, pp. 2-5. doi: 10.1109/MIS.2010.43.

Chen, H. and Zimbra, D. (2010), "AI and opinion mining", *IEEE Intelligent Systems*, Vol. 25 No. 3, pp. 74-76. doi: 10.1109/MIS.2010.75.

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

CNet (2015), "China likely to top US for Apple iPhone sales", *CNet*, available at: www.cnet.com/news/china-likely-to-top-us-for-apple-iphone-sales/ (accessed January 26, 2015).

ComScore (2015), "ComScore reports December 2014 US smartphone subscriber market share", available at: www.comscore.com/Insights/Market-Rankings/comScore-Reports-December-2014-US-Smartphone-Subscriber-Market-Share (accessed February 9, 2015).

Evangelos, K., Efthimios, T. and Konstantinos, T. (2014), "Understanding the predictive power of social media", *Internet Research*, Vol. 23 No. 5, pp. 544-559, available at: www.emeraldinsight.com/doi/full/10.1108/IntR-06-2012-0114

Fan, W. and Gordon, M.D. (2014), "The power of social media analytics", *Communications of the ACM*, Vol. 57 No. 6, pp. 74-81. doi: 10.1145/2602574.

He, W., Zha, S. and Li, L. (2013), "Social media competitive analysis and text mining: a case study in the pizza industry", *International Journal of Information Management*, Vol. 33 No. 3, pp. 464-472. doi: 10.1016/j.ijinfomgt.2013.01.001.

He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A novel social media competitive analytics framework with sentiment benchmarks", *Information & Management*, Vol. 52 No. 7, pp. 801-812. doi: 10.1016/j.im.2015.04.006.

Hu, M. and Liu, B. (2004), *Mining Opinion Features in Customer Reviews*, available at: www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon

Hua, T., Lu, C.-T., Ramakrishnan, N., Chen, F., Arredondo, J., Mares, D. and Summers, K. (2013), "Analyzing civil unrest through social media", *Computer*, Vol. 46 No. 12, pp. 80-84. doi: 10.1109/MC.2013.442.

Hung, C. and Lin, H.-K. (2013), "Using objective words in SentiWordNet to improve word-of-mouth sentiment classification", *IEEE Intelligent Systems*, Vol. 28 No. 2, pp. 47-54. doi: 10.1109/MIS.2013.1.

IDC (2014), "GalaxyS5 sales predict", *Money Today*, available at: http://news.mt.co.kr/mtview.php?no=2014080110401494101 (accessed August 1, 2014).

Kahaner, L. (1998), *Competitive Intelligence: How to Gather, Analyze, and Use Information to Move Your Business to the Top*, Touchstone, New York, NY, available at: www.amazon.com/Competitive-Intelligence-Analyze-Information-Business/dp/0684844044 (accessed February 11, 1998).

KGI (2015), "KGI: iPhone sales forecast at 73M for Q4", *ZDNet*, available at: http://appleinsider.com/articles/15/01/24/kgi-iphone-sales-forecast-at-73m-for-q4-ahead-of-apple-watch-debut-in-march-12-macbook-air-in-q1 (accessed January 27, 2015).

Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S. (2011), "Social media? Get serious! Understanding the functional building blocks of social media", *Business Horizons*, Vol. 54 No. 3, pp. 241-251. doi: 10.1016/j.bushor.2011.01.005.

Kim, E., Sung, Y. and Kang, H. (2014), "Brand followers' retweeting behavior on Twitter: how brand relationships influence brand electronic word-of-mouth", *Computers in Human Behavior*, Vol. 37, pp. 18-25. doi: 10.1016/j.chb.2014.04.020.

Kim, Y. and Jeong, S.R. (2015), "Opinion-mining methodology for social media analytics", *KSII Transactions on Internet and Information Systems*, Vol. 9 No. 1, pp. 391-406.

Kim, Y., Kwon, D.Y. and Jeong, S.R. (2015), "Comparing machine learning classifiers for movie WOM opinion mining", *KSII Transactions on Internet and Information Systems*, Vol. 9 No. 8, pp. 3178-3190. doi: 10.3837/tiis.2015.08.025.

Komando, K. (2013), "10 smartphone must-have features", *USA Today*, available at: www.usatoday.com/story/tech/columnist/komando/2013/12/13/smartphone-battery-processing-display-camera/3921399/ (accessed December 13, 2015).

Li, Z. and Li, C. (2014), "Twitter as a social actor: how consumers evaluate brands differently on Twitter based on relationship norms", *Computers in Human Behavior*, Vol. 39, pp. 187-196. doi: 10.1016/j.chb.2014.07.016.

Lim, J. and Kim, J. (2014), "An empirical comparison of machine learning models for classifying emotions in Korean Twitter", *Journal of Korea Multimedia Society*, Vol. 17 No. 2, pp. 232-239. doi: 10.9717/kmms.2014.17.2.232.

Liu, B. (2010), "Sentiment analysis: a mutlifaceted problem", *IEEE Intelligent Systems*, Vol. 25 No. 3, pp. 76-80, available at: http://doi.org/10.1109/MIS.2010.75

Liu, B., Hu, M. and Cheng, J. (2005), "Opinion observer: analyzing and comparing opinions on the web", *Proceedings of the 14th International Conference on World Wide Web – WWW 05*, ACM Press, New York, NY, p. 342. doi: 10.1145/1060745.1060797.

Liu, Y., Chen, Y., Lusch, R.F., Chen, H., Zimbra, D. and Zeng, S. (2010), "User-generated content on social media: predicting market success with online word-on-mouth", *IEEE Intelligent Systems*, Vol. 25 No. 1, pp. 8-12. doi: 10.1109/MIS.2010.75.

Localytics (2014), "Samsung Galaxy S5 grabs nearly 1% of all android smartphones after first week", available at: http://info.localytics.com/blog/samsung-galaxy-s5-grabs-nearly-1-percent-of-android-market

Lu, X., Ba, S., Huang, L. and Feng, Y. (2013), "Promotional marketing or word-of-mouth? Evidence from online restaurant reviews", *Information Systems Research*, Vol. 24 No. 3, pp. 596-612, available at: http://pubsonline.informs.org/doi/abs/10.1287/isre.1120.0454

Lusch, R.F., Liu, Y. and Chen, Y. (2010), "The phase transition of markets and organizations: the new intelligence and entrenreneurial frontier", *IEEE Intelligent Systems*, Vol. 25 No. 1, pp. 5-8. doi: 10.1109/MIS.2010.75.

McCarthy, J., Rowley, J., Ashworth, C.J. and Pioch, E. (2014), "Managing brand presence through social media: the case of UK football clubs", *Internet Research*, Vol. 24 No. 2, pp. 181-204. doi: 10.1108/IntR-08-2012-0154.

Mangold, W.G. and Faulds, D.J. (2009), "Social media: the new hybrid element of the promotion mix", *Business Horizons*, Vol. 52 No. 4, pp. 357-365. doi: 10.1016/j.bushor.2009.03.002.

Ortigosa, A., Martín, J.M. and Carro, R.M. (2014), "Sentiment analysis in Facebook and its application to e-learning", *Computers in Human Behavior*, Vol. 31, pp. 527-541. doi: 10.1016/j.chb.2013.05.024.

Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in inforamation Retrieval*, Vol. 2 Nos 1-2, pp. 1-135. doi: 10.1561/1500000011.

Pang, B., Lee, L. and Vaithyanathan, S. (2002), "Thumbs up? Sentiment classification using machine learning techniques", *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 79-86. doi: 10.3115/1118693.1118704.

Riaga, O. (2014), "A critical look at iPhone 6 in relation to 10 key smartphone features", available at: www.kachwanya.com/iphone-6-features/#sthash.5ADh8Ipf.dpuf (accessed September 12, 2014).

Rui, H., Liu, Y. and Whinston, A. (2013), "Whose and what chatter matters? The effect of tweets on movie sales", *Decision Support Systems*, Vol. 55 No. 4, pp. 863-870. doi: 10.1016/j.dss.2012.12.022.

Shen, J., Brdiczka, O. and Ruan, Y. (2013), "A comparison study of user behavior on Facebook and Gmail", *Computers in Human Behavior*, Vol. 29 No. 6, pp. 2650-2655. doi: 10.1016/j.chb.2013.06.043.

Sonnier, G.P., McAlister, L. and Rutz, O.J. (2011), "A dynamic model of the effect of online communications on firm sales", *Marketing Science*, Vol. 30 No. 4, pp. 702-716. doi: 10.1287/mksc.1110.0642.

Strategy Analytics (2015), "Apple becomes world's largest smartphone vendor in Q4 2014", *Strategy Analytics*, available at: http://blogs.strategyanalytics.com/WSS/post/2015/01/29/Apple-Becomes-Worlds-Largest-Smartphone-Vendor-in-Q4-2014.aspx (accessed January 29, 2015).

Wu, Y., Wei, F., Liu, S., Au, N., Cui, W., Zhou, H. and Qu, H. (2010), "OpinionSeer: interactive visualization of hotel customer feedback", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 16 No. 6, pp. 1109-1118. doi: 10.1109/TVCG.2010.183.

Ye, Q., Law, R., Gu, B. and Chen, W. (2011), "The influence of user-generated content on traveler behavior: an empirical investigation on the effects of e-word-of-mouth to hotel online bookings", *Computers in Human Behavior*, Vol. 27 No. 2, pp. 634-639. doi: 10.1016/j.chb.2010.04.014.

Yin, R. (2002), *Case Study Research: Design and Methods*, 3rd ed., Vol. 18, SAGE Publications, Thousand Oaks, CA, available at: www.ncbi.nlm.nih.gov/pubmed/7727570 (accessed December 24, 2002).

Zeng, D., Chen, H., Lusch, R. and Li, S.H. (2010), "Social media analytics and intelligence", *IEEE Intelligent Systems*, Vol. 25 No. 6, pp. 13-16. doi: 10.1109/MIS.2010.151.

Zhang, Z., Ye, Q., Law, R. and Li, Y. (2010), "The impact of e-word-of-mouth on the online popularity of restaurants: a comparison of consumer reviews and editor reviews", *International Journal of Hospitality Management*, Vol. 29 No. 4, pp. 694-700. doi: 10.1016/j.ijhm.2010.02.002.

Corresponding author
Professor Seung Ryul Jeong can be contacted at: srjeong@kookmin.ac.kr

This article has been cited by:

1. ChangHsin Hsin, Hsin Hsin Chang, WongKit Hong, Kit Hong Wong, ChuTsun Wei, Tsun Wei Chu. Online advertorial attributions on consumer responses: materialism as a moderator. *Online Information Review*, ahead of print. [Abstract] [Full Text] [PDF]
2. Anant Saxena, K. R. Chaturvedi, Sapna Rakesh. 2018. Analysing Customers Reactions on Social Media Promotional Campaigns. *Paradigm* **22**:1, 80-99. [Crossref]
3. Vásquez RojasClaudia, Claudia Vásquez Rojas, Roldán ReyesEduardo, Eduardo Roldán Reyes, Aguirre y HernándezFernando, Fernando Aguirre y Hernández, Cortés RoblesGuillermo, Guillermo Cortés Robles. 2018. Integration of a text mining approach in the strategic planning process of small and medium-sized enterprises. *Industrial Management & Data Systems* **118**:4, 745-764. [Abstract] [Full Text] [PDF]
4. JebleShirish, Shirish Jeble, DubeyRameshwar, Rameshwar Dubey, ChildeStephen J., Stephen J. Childe, PapadopoulosThanos, Thanos Papadopoulos, RoubaudDavid, David Roubaud, PrakashAnand, Anand Prakash. 2018. Impact of big data and predictive analytics capability on supply chain sustainability. *The International Journal of Logistics Management* **29**:2, 513-538. [Abstract] [Full Text] [PDF]
5. Song Gao, Ou Tang, Hongwei Wang, Pei Yin. 2018. Identifying competitors through comparative relation mining of online reviews in the restaurant industry. *International Journal of Hospitality Management* **71**, 19-32. [Crossref]
6. ThelwallMike, Mike Thelwall. 2018. Gender bias in sentiment analysis. *Online Information Review* **42**:1, 45-57. [Abstract] [Full Text] [PDF]
7. Veera Boonjing, Daranee Pimchangthong. Data Mining for Positive Customer Reaction to Advertising in Social Media 83-95. [Crossref]
8. Wu He, Xin Tian, Andy Hung, Vasudeva Akula, Weidong Zhang. 2017. Measuring and comparing service quality metrics through social media analytics: a case study. *Information Systems and e-Business Management* **55**. . [Crossref]
9. WangHongwei, Hongwei Wang, GaoSong, Song Gao, YinPei, Pei Yin, LiuJames Nga-Kwok, James Nga-Kwok Liu. 2017. Competitiveness analysis through comparative relation mining. *Industrial Management & Data Systems* **117**:4, 672-687. [Abstract] [Full Text] [PDF]