1-1-2009

Assessing Public Opinions Through Web 2.0: A Case Study on Wal-Mart

David Zimbra  
*University of Arizona, zimbra@arizona.edu*

Tianjun Fu  
*University of Arizona, futj@email.arizona.edu*

Xin Li  
*City University of Hong Kong, xinli.is@acm.org*

Follow this and additional works at: [http://aisel.aisnet.org/icis2009](http://aisel.aisnet.org/icis2009)

Recommended Citation  

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2009 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
ASSESSING PUBLIC OPINIONS THROUGH WEB 2.0: A CASE STUDY ON WAL-MART

Research-in-Progress

David Zimbra
Department of MIS
University of Arizona
Tucson, AZ 85721
zimbra@email.arizona.edu

Tianjun Fu
Department of MIS
University of Arizona
Tucson, AZ, 85721
futj@email.arizona.edu

Xin Li
Department of Information Systems
City University of Hong Kong
Kowloon, Hong Kong, 220
xinli.is@acm.org

Abstract

The recent advancement of Web 2.0 enables people to exchange their opinions on a variety of topics. Among these discussions, the opinions of employees, customers, and investors are of great interest to companies. Insight into such perspectives can help managers make better decisions on business policies and strategy. However, assessing online opinions is a nontrivial task. The high volume of messages, casual writing style, and the significant amount of noise require the application of sophisticated text mining techniques to digest the data. Previous research has successfully applied sentiment analysis to assess online opinions on specific items and topics. In this research, we propose the integration of topic analysis with sentiment analysis methods to assess the public opinions expressed in forums with diverse topics of discussion. Using a Wal-Mart-related Web forum as an example, we found that combining the two types of analysis can provide us with improved ability to assess public opinions on a company. Through further analysis on one cluster of discussions, several abnormal traffic and sentiment patterns were identified related to Wal-Mart events. The case study validates the propose framework as an IT artifact to assess online public opinion on companies of interest. Our research promotes further efforts to combine topic and sentiment analysis techniques in online research supporting business decision making.

Keywords: Web 2.0, topic analysis, sentiment analysis, user-generated content
ASSESSING PUBLIC OPINIONS THROUGH WEB 2.0: A CASE STUDY ON WAL-MART

Introduction

The recent advancement of Web 2.0 provides several new channels, such as Web forums and Web blogs, for individuals to communicate. These interactive social media enable different constituencies to publish opinions on a wide variety of topics, including many of concern to businesses. Compared to traditional news media that typically focus on the opinions of authoritative sources, Web 2.0 may accumulate opinions from participants including employees, customers, and investors in a company. These discussions could provide managers insight into perspectives on companies’ internal operations, public relationships, and even evidence for external issues affecting their business. Such information could assist managers in making better decisions on their policies and business strategy. Additionally, it may also be possible for companies to utilize Web 2.0 media to shape public opinions. Previous research has shown that consumer behavior is increasingly influenced by peer opinions (Surowiecki 2005), while the influence of traditional media is being reduced (Gillin 2007). Thus, assessing online public opinion is critical for managers to take advantage of this dynamic networked world.

However, assessing online opinions is a nontrivial task. The high volume of messages, casual writing style, and significant amount of noise requires the application of sophisticated text mining techniques to digest the data. In this research, we propose an integrated framework that applies automatic topic and sentiment extraction methods to assess the public opinions on companies of interest. Based on the proposed framework, we conduct a case study on a Web forum related to Wal-Mart and analyze the communication traffic patterns, discussion sentiment characteristics, and opinion leader activities. The case study clearly shows the value of Web 2.0 as a relevant source of knowledge for companies, and highlights the need to combine multiple text mining techniques to provide a clear understanding of public opinions.

The remainder of the paper is organized as follows. Section 2 reviews the background research on assessing and analyzing public opinions for business purposes. Section 3 summarizes our research questions. In Section 4, we present our methodology for topic-sentiment based public opinion mining. In Section 5 we present the preliminary analysis of a case study on a Wal-Mart-related forum, and discuss our findings. We conclude the paper with the planned extensions of this research and expected contributions in Section 6.

Related Work

Assessing public opinion (such as the opinions of customers) has been of interest to business managers for years, because of their implications on decision making. Traditional methods for assessing public opinion include collecting comment cards, setting up toll-free telephone numbers for feedback, and distributing questionnaires. At the end of the 1990s, the World Wide Web prompted the use of email as a feedback channel, which significantly improved the ease and effectiveness of collecting public opinions (Sampson 1998). The emergence of Web 2.0 significantly changed the nature of consumer feedback. Virtual communities are established online to discuss issues related to the company. In such an environment, participants may not only reveal their own opinions but also be influenced by those of other members (Gillin 2007).

Online discussions in Web 2.0 are on a wide variety of topics, a factor making them difficult to analyze. In previous research, a major stream has focused on understanding online opinions on specific items or products. For example, Liu et al. proposed a system to visualize online product reviews to aid in purchasing decisions (Liu et al. 2005). Furthermore, Glance et al. developed a prototype that can analyze customer comments on products in Web blogs and forums (Glance et al. 2005). Wenger also studied Web blogs for the purpose of assessing tourists’ comments on Austrian tourism destinations (Wenger 2008). Researchers have also studied public opinion of companies, while focusing on specific aspects of the company. Motivated by applications in financial management, several studies have analyzed the Web forums that discuss stock and other financial issues. Previous research has found that forum discussion volumes are correlated with the abnormal returns related to companies’ stock price (Tumarkin et al. 2001; Wysocki 1998). Bagnoli found the “whisper forecasts” in forum discussions to be more accurate than First Call analyst forecasts (Bagnoli et al. 1999). Several follow up studies using heuristics and machine learning methods to
extract investors’ opinions from stock discussion forums found a statistically significant relationship with the abnormal returns of stocks (Antweiler et al. 2004; Das et al. 2005; Das et al. 2007; Gu et al. 2006).

The previous studies on product reviews and stock price prediction based on user generated content belong to a wider field of research: sentiment analysis. Sentiment analysis is described as the process of determining the level of subjectivity and polarity (i.e., direction and intensity of expressions of opinion) in text. Applications include business problems, such as analysis of product reviews (Dave et al. 2003; Pang et al. 2002) and customer feedback (Gamon 2004), and in knowledge management, such as enhanced information retrieval (Godbole et al. 2007; Mishne 2006). Approaches to sentiment analysis can be divided into those that rely upon semantic resources such as sentiment or affect lexicons (Mishne 2006), and those that try to learn patterns directly from tagged text using machine learning techniques (Dave et al. 2003; Pang et al. 2002). A lexical approach to sentiment analysis is often favored by researchers because it does not require the training of a learning component to identify sentiment, which often necessitates the laborious annotation of a sample of the data. However, learning approaches are capable of determining aspects of context, and detecting the subtle and indirect expressions of sentiment (Pang et al. 2002) that may be mistaken or omitted by a strictly lexical approach.

Sentiments analysis only provides us with a part of the story of online opinions. Online discussions are on a wide variety of topics. Sentiment analysis could provide more valuable information to managers if conducted while focusing on specific topics of interest. Topic identification is a traditional research question in text mining, which is related to text classification, categorization, summarization, and automatic document indexing. The extraction of topics can be based on the linguistics statistics of a corpus, including term frequency, term location, term co-occurrence, etc. Latent Semantic Indexing (LSI) is an example of such an approach (Deerwester et al. 1990). Topic extraction can also be conducted based on external knowledge such as dictionaries, syntactic and semantic parsers, and is often applied in conjunction with clustering algorithms. For example, Clifton et al. applied mutual information-based methods with clustering algorithms on news articles for topic identification (Clifton et al. 2004).

While previous studies have shown that rich and valuable sentiment information is embedded in online discussions, the existing analyses are limited to discussion on specific items and topics. However, there is limited research that takes advantage of both topic and sentiment analysis techniques in assessing public opinion on a company through Web 2.0, which contains a diverse set of topics and mixture of sentiments.

Research Questions

In light of the limitations of previous research, we propose to examine IT artifacts that can be used for uncovering topics and sentiments from online discussion to support managerial decision making. In this research, we pose the following questions.

1) How can topic and sentiment analyses be combined to better assess online public opinions of a specific company?
2) What can the uncovered online opinions tell us about various aspects of a company?

Methodology

In this research, we propose to combine topic analysis and sentiment analysis to assess the facets of discussion and public opinion of a company using Web 2.0. Our research framework consists of four steps: data collection, topic extraction, sentiment identification, and opinion analysis, as shown in Figure 1.

1) Data Collection

In the data collection step, we aim to collect Web 2.0 public discussions pertaining to specific companies. To identify potential sources of information, it is usually necessary to utilize Web search engines and Web forum/blog platforms. Online discussions are collected by Web crawlers and then parsed into relational databases for analysis.

2) Topic Extraction

In our framework, we propose a clustering approach together with Natural Language Processing techniques for topic extraction. We first perform thread-level clustering based upon the discussion content similarity using the maximum entropy model. For each identified cluster of threads, we apply a mutual information (MI)-based noun phrase
extractor, Arizona Noun Phraser (Tolle et al. 2000), to extract major terms that can represent the topics discussed in the threads comprising the cluster.

3) Sentiment Identification

In parallel to the topic extraction, we identify the sentiments expressed in the online discussions. We present a hybrid model to sentiment analysis that leverages both lexical and learning approaches in a combined fashion. The model couples the OpinionFinder system for subjectivity analysis (Wilson et al. 2005a) with the SentiWordNet sentiment lexicon (Esuli et al. 2006).

After preprocessing steps including sentence splitting, tokenization, and part of speech tagging, the OpinionFinder system uses a Naïve Bayes classifier to distinguish subjective and objective sentences (Riloff et al. 2003; Wiebe et al. 2005). Then, a rule-based classifier is applied to identify direct subjective expressions (Choi et al. 2006). Finally, OpinionFinder uses two boosting-based classifiers to classify whether these expressions are positive or negative (Wilson et al. 2005b). However, the system provides limited information on the intensity of the expression polarity. Therefore, we incorporate the SentiWordNet sentiment lexicon, which provides knowledge on the intensity and polarity of terms, to enhance the interpretation of the OpinionFinder output. The SentiWordNet lexicon consists of positive, negative, and objectivity scores for each WordNet synset. These scores can be applied more accurately if knowledge on a term’s particular usage and context are provided. Thus, our model considers the part of speech and the polarity classification results of an expression to make more effective use of the lexicon.

In our research, to score the sentiment value of online discussions we focus only on the sentences classified by OpinionFinder as subjective. From the subjective sentences, we extract each term identified as a sentiment expression together with their contextual information, such as part of speech and polarity (positive or negative). We use this part of speech and polarity information to identify the appropriate sentiment value in SentiWordNet for the usage of the term in the context of the message. If a particular term has no entry in SentiWordNet, we assign it a sentiment value corresponding to the average of all entries in the lexicon having that particular part of speech and polarity, to make use of the available information. The values of each sentiment expression are aggregated (summed) to the sentence level, and then averaged to the message level. Therefore we define the sentiment value for a particular message as follows:

\[
sent(m) = \frac{1}{m} \sum_{s \in m} \delta(s) \times \sum_{t \in s} swn(t, pos(t, s), pol(t, s))
\]

Where:

- \(sent(m)\) = sentiment score for message \(m\)
- \(\delta(s) = 1\) if sentence is subjective, \(0\) if sentence is objective
\(|m| = \text{the number of sentences in the message}
\)

\(S = \text{sentences in the message}
\)

\(t = \text{sentiment expression terms in the sentence}
\)

\(\text{pos}(t) = \text{part of speech of } t \text{ in sentence } s
\)

\(\text{pol}(t) = \text{polarity of } t \text{ in sentence } s
\)

\(\text{swn}(t,\text{pos}(t),\text{pol}(t)) = \text{SentiWordNet value for term } t \text{ based on its part of speech and polarity in the sentence}
\)

4) Opinion Analysis

To answer our research questions, we first measure the effect of including topical information in the analysis of online opinions. We utilize the topic and sentiment analyses in combination, and study their relationship with stock price as a proxy for a public opinion index. The selection of the approach is due to two reasons: 1) this topic is of interest to a variety of managers, investors, and researchers; 2) for evaluation purposes, the stock price and related measures are easy to access, and provide a reasonable proxy for public opinion; other data characterizing public opinion on a company may not be readily available. We evaluate a simple time series model to see how the sentiments expressed in the forum relate to the stock price at a weekly level. We compare the predictive power of using only sentiment analysis versus using both sentiment and topic analyses to predict the stock movement using Web 2.0 discussions. As suggested by the literature, we utilize control variables measuring the price volatility and trading volume of the stock (Das et al. 2007). The models take the forms below, with the ‘Sentiment’ model utilizing only the sentiment and control time lagged variables, while the ‘Sentiment + Topic’ model adding the topical cluster variables.

**Sentiment Model**

\[ sp_t = \alpha_0 + \alpha_1 \text{sent}_{t-1} + \alpha_2 \text{range}_{t-1} + \alpha_3 \text{volume}_{t-1} \]

**Sentiment + Topic Model:**

\[ sp_t = \alpha_0 + \alpha_1 \text{sent}_{t-1} + \alpha_2 \text{range}_{t-1} + \alpha_3 \text{volume}_{t-1} + \sum_{i=1}^{\text{# of clusters}} \beta_i \text{Cluster}_{i,t-1} \]

Where:

\(sp_t = \text{average daily stock price close during week } t\)

\(\text{sent}_{t-1} = \text{aggregated forum sentiment during week } t-1\)

\(\text{range}_{t-1} = \text{daily stock price close volatility during week } t-1\)

\(\text{volume}_{t-1} = \text{average daily stock trading volume during week } t-1\)

\(\text{cluster}_{i,t-1} = \text{number of unique threads in a cluster active during week } t-1\)

In the second part of the opinion analysis, we inspect the major topic groups describing the online discussions and examine the traffic statistics (posting frequency), opinion leaders, and the development of sentiments on the company over time. The purpose of these analyses is to provide an exploratory approach to assessing the major topics and associated sentiments in Web 2.0 discussions.

**Analysis and Results**

**Dataset**

As a preliminary implementation of our research framework we conducted a case study on a Web forum with Wal-Mart-related discussions: the “Wal-Mart Sucks” board of the Wal-Mart blows forum (http://www.walmart-blows.com/forum/viewforum.php?f=3). The dataset consists of 1,354 threads with 19,624 postings by 1,855 authors from November 2003 to November 2008. Discussions in this forum mainly express negative opinions on Wal-Mart.
Figure 2 shows the participant activity in the forum. In general, the number of active authors and postings are proportional to each other. It should be noted that the postings in the WSB dataset experienced a surge in the third quarter of 2006, to nearly twice the usual volume. Following the surge, many participants maintained their active participation in the forum. Further analysis showed that participants in the forum contributed unevenly to the posting volume. The top 1% of authors was responsible for 30% of the postings and the top 10% produced about 70% of the postings.

![Figure 2. Activity on the Wal-Mart Sucks Board](image)

**Experiment Procedure**

After collecting the Web forum data, we applied a clustering algorithm using the maximum entropy model implemented in the Weka package (Witten et al. 2005). We used the standard bag-of-words as our feature representation in this preliminary study. The model parameters were specified to extract 10 clusters with 100 maximum iterations. We also applied the sentiment method detailed in the research design for sentiment evaluation. We then compared the two models in the prediction of the company’s stock price, as a proxy for a public opinion index. Finally, we conduct a case study on one major topical cluster identified in the forum and analyze its traffic/sentiment trends and opinion leaders.

**Predictive Analysis**

To conduct the analysis of the sentiments expressed in the forum and their relation to a common public opinion measure, the daily historical prices of the Walmart common stock (WMT) were collected to serve as the dependent variable in the regression models. Since the number of daily forum postings is insufficient to warrant comparative analysis at this level of time granularity, the forum traffics, sentiments, trading volumes, and stock prices were averaged to a weekly value. The volatility in price was calculated using the daily stock close prices as the difference between the weekly high and low price. The topical clusters were represented using 10 variables, one for each of the clusters. The values for the cluster variables were calculated as the number of unique threads active within each respective topical cluster during the specified week. Consistent with similar studies (Das et al. 2007) all variables were standardized prior to inclusion in the regression model.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha_i$</th>
<th>t-stat</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentiment</td>
<td>0.194**</td>
<td>3.05</td>
<td>0.172</td>
</tr>
<tr>
<td>Sentiment + Topic</td>
<td>0.165*</td>
<td>2.05</td>
<td>0.346</td>
</tr>
</tbody>
</table>

Table 1 compares the prediction of the two models (Sentiment model and Sentiment + Topic model). In the sentiment model, the variable representing the previous week’s sentiment expression is significant at the 0.005 level, with an Adjusted $R^2 = 0.172$. Incorporating the topical cluster variables resulted in an increase in Adjusted $R^2$ to a value of 0.346, while the sentiment variable remains significant at level 0.05. These results indicate that the
inclusion of topical information can enrich the sentiment analysis of a forum, and provide knowledge that more closely aligns with the public perception represented by the stock price.

**Exploratory Analysis**

We further inspect the topics/sentiments in the forum in exploratory analyses that may answer a variety of managerial questions. Figure 3 reports the 4 largest topical clusters with more than 100 forum discussion threads. We extracted the major topics (top 20 noun phrases) from the clusters using the Arizona Noun Phraser. After inspecting the extracted keywords, we divided the keywords into three categories: employee-related, management-related, and customer-related. All four clusters contain similar customer-related keywords. However, cluster 2 represents more management-related discussions, while cluster 4 focuses more on employee-related topics. For our case study, we chose to analyze the threads in cluster 4 due to our interest in Wal-Mart’s internal perspectives given the employee-driven nature of the forum.

<table>
<thead>
<tr>
<th>Employee</th>
<th>Management</th>
<th>Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>home office</td>
<td>store manager</td>
<td>Lee Scott</td>
</tr>
<tr>
<td>district manager</td>
<td>parking lot</td>
<td>low price</td>
</tr>
<tr>
<td>customer service</td>
<td>credit card</td>
<td>gift card</td>
</tr>
<tr>
<td>Cluster 1</td>
<td>Cluster 2</td>
<td>Cluster 3</td>
</tr>
<tr>
<td>minimum wage</td>
<td>black friday</td>
<td>state law</td>
</tr>
<tr>
<td>shopping cart</td>
<td>cart pusher</td>
<td>high school</td>
</tr>
<tr>
<td>holiday pay</td>
<td>cell phone</td>
<td>cart pusher</td>
</tr>
<tr>
<td>sporting good</td>
<td>door greeter</td>
<td>black friday</td>
</tr>
<tr>
<td>3rd shift</td>
<td>break room</td>
<td>break room</td>
</tr>
<tr>
<td>home office</td>
<td>high school</td>
<td>cash office</td>
</tr>
<tr>
<td>small business</td>
<td>open door</td>
<td>door greeter</td>
</tr>
<tr>
<td>Lee Scott</td>
<td>home office</td>
<td>open door</td>
</tr>
<tr>
<td>Sam Walton</td>
<td>district manager</td>
<td>new hire</td>
</tr>
<tr>
<td>assistant manager</td>
<td>Sam Walton</td>
<td>shopping cart</td>
</tr>
<tr>
<td>department manager</td>
<td>Lee Scott</td>
<td>home office</td>
</tr>
<tr>
<td>store manager</td>
<td>department manager</td>
<td>low price</td>
</tr>
<tr>
<td>support manager</td>
<td>assistant manager</td>
<td>self check</td>
</tr>
<tr>
<td>low price</td>
<td>store manager</td>
<td>service desk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>customer service</td>
</tr>
<tr>
<td></td>
<td></td>
<td>parking lot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>service desk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>low price</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gift card</td>
</tr>
<tr>
<td></td>
<td></td>
<td>credit card</td>
</tr>
<tr>
<td></td>
<td></td>
<td>food stamp</td>
</tr>
</tbody>
</table>

**Figure 3. Topics of the Major Clusters**

**Traffic Dynamics and Sentiment Evolution**

Figure 4 shows the traffic dynamics and sentiment evolution of the cluster 4 threads. The discussions comprising this cluster experienced three surges, toward the end of 2005, in the third quarter of 2006, and in the third quarter of 2007. During the first surge, the average sentiment maintains the -0.2 level. This may correspond to negative public
perception of Wal-Mart’s health care issues, since Wal-Mart was facing several related lawsuits at that time. During the second surge, the average sentiment rose to a level of -0.1. This may be related to the launch of “Generic Drug Program” at Wal-Mart, which initiated several discussions on its impact on the retail drug industry. In the third surge, the average sentiment of the cluster shows a clear decreasing trend. These discussions were related to the heavy workload during the holidays and lawsuits on labor issues.

**Opinion Leaders**

Table 2 lists the opinion leaders who had most number of messages in the threads comprising cluster 4. In general, all opinion leaders show negative opinions towards Wal-Mart. However, it is noted that the sentiments for user “NightStalker” and “corporate_lackey” were much lower than the other opinion leaders. Their discussions were mostly about customer services and employee issues. The lower scores of “NightStalker” and “corporate_lackey” were due in part to disrespectful and antagonistic behavior aimed at other forum members. Users “markofkane”, “jiggy”, and “No_Name”, while leaders in terms of their participation in forum discussions, were much more respectful and exhibited fewer personal attacks on other members.

<table>
<thead>
<tr>
<th>Author ID</th>
<th># of Postings</th>
<th># of Threads</th>
<th>Avg. Sentiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>markofkane</td>
<td>257</td>
<td>73</td>
<td>-0.11</td>
</tr>
<tr>
<td>jiggy</td>
<td>253</td>
<td>60</td>
<td>-0.07</td>
</tr>
<tr>
<td>NightStalker</td>
<td>156</td>
<td>47</td>
<td>-0.34</td>
</tr>
<tr>
<td>corporate_lackey</td>
<td>150</td>
<td>60</td>
<td>-0.23</td>
</tr>
<tr>
<td>No_Name</td>
<td>119</td>
<td>20</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

**Conclusion and Future Directions**

In this research, we propose a framework that combines topic analysis and sentiment analysis to assess the online public opinions of companies of interest. Using a Wal-Mart-related Web forum as an example, we found that combining the two types of analyses reveals opinion information having a strong relationship with a common measure of public opinion, a company’s stock price. In addition, the case study on one topical cluster of discussions in the forum enables us to identify several abnormal traffic and sentiment patterns in response to Wal-Mart related events. We also found distinctive types of opinion leaders in the clusters. In short, the case study validates the propose framework as an IT artifact designed to assess online public opinions on companies.

In order to better validate the proposed framework, we are in the process of collecting other datasets and conducting more comprehensive experimentation and study. We have started collecting forum discussions reflecting customer, employee and investor opinions of Wal-Mart. We are in the process of examining the effectiveness of the proposed framework on datasets with different characteristics. We will extend our comparative experiments to different topic extraction and sentiment identification techniques in order to assess the effectiveness of the framework by controlling the techniques used in this IT artifact. We will also extend the current research to other industries and companies in order to control the difference across companies.

The goal of the present work is to contribute to our use of online opinions for business management. First, our proposed framework integrates a set of techniques that can be used to dealing with the diverse topics and mixed sentiments in online discussions. Second, our analysis showed the benefit of including both topic and sentiment analysis in the framework, which should be considered in similar studies in the future. Third, we believe our research will promote the use of user generated content in business decision making, especially transitioning from a manual-based analysis fashion to automatic processing.

**References**

Antweiler, W., and Frank, M.Z. "Is All That Talk Just Noise? The Information Content of Internet Stock Message


