

8-2010

Investigating Determinants of Voting for the “Helpfulness” of Online Consumer Reviews: A Text Mining Approach

Wenjing Duan

The George Washington University, wduan@gwu.edu

Qing Cao

Texas Tech University, qing.cao@ttu.edu

Qiwei Gan

Texas Tech University, qing.cao@ttu.edu

Follow this and additional works at: <http://aisel.aisnet.org/amcis2010>

Recommended Citation

Duan, Wenjing; Cao, Qing; and Gan, Qiwei, "Investigating Determinants of Voting for the “Helpfulness” of Online Consumer Reviews: A Text Mining Approach" (2010). *AMCIS 2010 Proceedings*. 497.

<http://aisel.aisnet.org/amcis2010/497>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2010 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Investigating Determinants of Voting for the “Helpfulness” of Online Consumer Reviews: A Text Mining Approach

Wenjing Duan

School of Business
The George Washington University
Email: wduan@gwu.edu

Qing Cao

Rawls College of Business
Texas Tech University
Email: qing.cao@ttu.edu

Qiwei Gan

Rawls College of Business
Texas Tech University
Email: qing.cao@ttu.edu

ABSTRACT

The “helpfulness” feature of online user reviews helps consumers cope with information overloads and facilitates decision making. However, many online user reviews lack sufficient helpfulness votes for other users to evaluate their true helpfulness level. This study empirically examines the impact of the various features, that is, basic, stylistic, and semantic, of online user reviews on the number of helpfulness votes those reviews receive. Text mining techniques are employed to extract semantic characteristics from review texts. Our findings show that the semantic characteristics are more influential than other characteristics in affecting how many helpfulness votes reviews receive. Our findings also suggest that reviews with extreme opinions receive more helpfulness votes than those with mixed or neutral opinions. This paper sheds light on the understanding of online users’ helpfulness voting behavior and the design of a better helpfulness voting mechanism for online user review systems.

Keywords

Online user review, helpfulness vote, text mining, latent semantic analysis

INTRODUCTION

Understanding the role of online user reviews in e-commerce has become an increasingly important subject for both academics and practitioners (Duan, Gu, and Whinston 2008). The great amount of information available on the Web has created information overload among online users (Brynjolfsson and Smith, 2000; Jones, Ravid and Rafaeli, 2004). The information overload spans two dimensions. First, the number and types of products available online have grown exponentially. Online consumers often find they lack the knowledge and time to make the best possible decision out of numerous competing products. Online user review systems, in this sense, provide a venue for consumers to share their opinions and experience on products. Second, even though the information overload created by the availability of a large amount and variety of products online could be mitigated by referring to online user reviews to some extent, the utter volume of available online user reviews create another big obstacle to consumers (Liu, Huang, An, Giannotti, Gunopulos, Turini, Zaniolo, Ramakrishnan and Wu, 2008). It is virtually impossible for consumers to read all the reviews before making purchase decisions. Many websites encourage users to evaluate the “helpfulness” of user reviews by simply asking anyone who read the review to vote on the question “Was this review helpful to you?” More and more websites even display reviews based on the helpfulness voting. This feature allows consumers to quickly find the most helpful reviews and makes the consumer decision-making process more efficient. It is estimated that this simple question “Was this review helpful to you?” brings about \$2.7 billion in additional revenue to Amazon.com (Spool, 2008).

However, the helpfulness voting is not a panacea. A large portion of online user reviews on many popular websites do not receive any helpfulness votes. Since consumers are expected to pay greater attention to the most helpful reviews, less helpful reviews become less attractive to consumers. This may create a vicious cycle in which the more helpful reviews attract more readers and hence receive even more helpfulness votes, while the less helpful reviews are even less likely to receive additional helpfulness votes. To make things worse, many reviews do not receive helpfulness votes merely because they were posted more recently than others, and thus have not obtained enough exposure to the users. Hence, the true helpfulness level

of these reviews cannot be simply evaluated by the number of helpfulness votes. Consumer decision-making facilitated by the helpfulness votes, therefore, can be skewed without considering when the review is posted and what the context is.

In order to take all the user reviews into consideration in evaluating their impact, previous studies employed various approaches to assess or predict the helpfulness level of reviews without any helpfulness votes (Forman, Ghose and Wiesenfeld, 2008; Kim, Pantel, Chklovski and Pennacchiotti, 2006; Liu et al., 2008). The assessment and predictions are usually based on the number of helpfulness votes, and rely heavily on the correct and precise evaluation of other content characteristics of reviews that have received helpfulness votes. However, the definition and measurement of the helpfulness features of reviews are neither clear nor consistent in the existing literature (Forman et al., 2008; Liu et al. 2008).

In this paper, we approach this problem from a different perspective. Instead of predicting a helpfulness level for reviews that have no votes, we investigate the factors that determine the number of helpfulness votes a particular review receives. Our objective is to understand why some reviews receive many helpfulness votes while others receive few or no votes, and to explore what characteristics of online user reviews influence the number of helpfulness votes. To the best of our knowledge, there is no prior study addressing this most basic yet critical question. A better understanding of what drives the helpfulness voting would help e-commerce websites improve the design of user review systems to encourage more helpfulness votes on online user reviews and facilitate users to derive more informed decisions.

Using the reviews collected from a well-known website, we empirically examined the effects of basic, stylistic, and semantic characteristics of online user reviews on the number of helpfulness votes reviews have received. Our research expands the text mining research literature by employing text mining methodology in quantifying large amount of text data such as online user reviews. As an alternative to previous arduous content analysis method (Kassarjian, 1977; Kolbe and Burnett, 1991), text mining techniques have gained more and more attention from academic researchers. We use the text mining methodology to extract semantic characteristics from the review texts, and then compared and contrasted the effects of the basic, stylistic and semantic characteristics on the number of helpfulness votes the reviews received using ordinal logistic regression models. Our findings show that semantic characteristics are more influential than the other characteristics in affecting how many helpfulness votes the reviews receive. This finding demonstrates the importance of employing more viable text mining techniques in uncovering the information content and exploring the influence of online user reviews. Our findings also suggest that reviews with extreme opinions receive more helpfulness votes than those with mixed or neutral opinions. This finding also provides major implications to marketing practitioners in that “extremeness” of opinions could be used to attract customers’ attention.

RESEARCH METHODS

Data Collection

Data for this research were collected from CNET Download.com (CNETD: <http://www.download.com>), which is a leading and representative online platform of the software market. CNETD is a library of more than 50,000 free or free-to-try software programs for Windows, Mac, mobile devices, and Webware. CNETD offers a widely accepted user feedback system for online users to share their opinions and experiences. The user review system includes detailed comments and an overall evaluation indicated by a five-star user rating system.

We collected the entire history of review data up to May 2009 for software programs in one of the largest groups of Windows software (Enterprise Computing), which includes a wide variety of various categories. For each software program, we collected all the reviews posted, with each record consisting of: reviewer’s user ID, post time, title, pros, cons, summary, number of helpfulness votes, and total number of votes. Our sample consists of 87 software programs, which belong to 28 unique categories that are large enough to provide a diversified coverage of various software programs. The total number of user reviews for the 87 software programs are 3,460.¹ Table 1 presents the distribution of the number of reviews received by the software programs.

Table 2 shows the distribution of reviews by the number of helpfulness votes. Note that one category contains reviews that received seven or more helpfulness votes. This is because as the number of helpfulness votes for a review increases, the number of such reviews decreases sharply. In analyzing the number of votes that reviews received, we combined the number of reviews that received seven votes and above into one group representing the reviews with “many” helpfulness votes, in order to ensure enough observations.

¹ The detailed description of the number of software programs and reviews in each category is available upon request.

Number of Reviews (n)	Number of Software	Percentage (%)
n ≤ 10	23	26.4
10 < n ≤ 30	42	48.3
30 < n ≤ 100	16	18.4
n > 100	6	6.9
Total	87	100.0

Table 1. Number of Software Programs by Number of Reviews

Number of helpfulness votes	Number of reviews	Percentage (%)
0	1785	51.59
1	700	20.23
2	363	10.49
3	215	6.21
4	93	2.69
5	57	1.65
6	42	1.21
≥7	205	5.93
Total	3460	100.00

Table 2. Reviews by Number of Helpfulness Votes

User Reviews at CNETD

We categorize the available information from the reviews into three types of characteristics, namely, basic, stylistic, and semantic characteristics, which have been extensively used in previous studies (Forman et al., 2008; Kim et al., 2006; Liu et al., 2008). In our study, we aim to examine whether those three types of characteristics have any impact on the number of helpfulness votes that reviews receive.

Basic Characteristics of Review

The first type of information is what we can directly observe from a review, including (1) whether the reviewer wrote about “pros,” (2) whether the reviewer wrote about “cons”, (3) whether the reviewer wrote anything in the summary, (4) how many days since the posting date, and (5) the “extremeness” level of the review, which can be roughly estimated as the absolute value of the difference between the reviewers’ rating and the average of all user ratings.

Stylistic Characteristics of Review

The second type of information is the stylistic characteristics of a review, which represent key features of reviewers’ writing style that cannot be easily derived by simply browsing the review texts. Table 3 lists the characteristics we have examined.

Semantic Characteristics of Review

The third type of information is semantic characteristics, which are related to the substance of the review. Examining the exact meaning of text is extremely difficult and often subjective. As such, we turn to a more practical way to parse the meaning of reviews with the help of Latent Semantic Analysis (LSA). LSA is a widely used statistical approach to analyze relationships between a set of documents and terms in these documents. It can produce a set of meaningful patterns related to the documents and terms (Deerwester, Dumais, Furnas, Landauer and Harshman, 1990). Different from previous studies, rather than trying to identify what semantic characteristic(s) causes viewers to vote on the helpfulness of reviews, we

examine the more fundamental question of whether semantic characteristics as a whole have any impact on the number of helpfulness votes that reviews receive.

Description
Number of words
Number of sentences
Average characters per word
Average words per sentence
Number of words in pros
Number of words in cons
Number of words in summary
Number of words in title
Number of 1- 9 letter words, respectively
Number of 10 or more-letter words

Table 3. Stylistic Characteristics

Text Mining Methodology

The LSA-based text mining methodology employed in this paper is shown in Figure 1. Details of each step are explained in the following sections.

Text Preprocessing

In this step, we calculate the number of words, number of sentences, words per sentence, number of characters per word, and number of words of different length for the title, pros, cons, and summary of each review. This preprocessing provides the stylistic characteristics of the review. In order to obtain an overall view of the review, we combined the title, pros, cons and summary of each review into one text block.

Parsing

In the parsing step, we used SAS® Enterprise Text Miner to perform stemming, part-of-speech tagging and term identification. The purpose of stemming is to treat words with different tense as one term. Part-of-speech tagging identifies the part-of-speech of each term and classifies the term as a noun, verb, adjective, and etc. Term identification is used to regard a word in the text as one term after stemming and part-of-speech tagging. Homonyms, words spelled the same but belonging to different parts-of-speech, are counted as multiple terms; multiple words with the same root, however, are counted as only one term. The purpose of term identification is to construct a so-called “term-by-frequency matrix” with each row referring to each review text and each column representing each term. However, one of the problems of this matrix is that when there are too many reviews and too many terms, dimensions of the matrix become extremely large, which makes it extremely difficult to conduct computations. In our data set, 16,168 terms were identified in 3460 reviews, resulting in a 3460×16168 matrix. Hence, another procedure is required in LSA text mining procedure to reduce the number of terms.

Term Reduction

One of the objectives of LSA is to discriminate one text from another in a semantic sense. In our study, first, we try to discriminate reviews with many helpfulness votes from those with none or very few. Since relatively meaningless words such as “a, an, the” in the reviews are not useful in discrimination, we compiled a list of the meaningless words (generally called “stop words”) and eliminated them from the Term-by-Frequency Matrix, which reduced the number of columns. Second, we deal with synonyms in the text. We compiled a synonyms dictionary to treat synonyms as one single term. As a result, many synonyms are consolidated into single terms in our analysis, resulting in fewer columns in the matrix.

However, only using term frequency cannot discriminate reviews effectively. One term that appears very frequently in one review may also appear so in other reviews. The most commonly used terms may appear frequently in almost all review texts, and thus are not useful in distinguishing one review from others. Conversely, the less frequent and unique terms are

more useful. In order to solve this problem, the term frequencies are transformed by TF-IDF (Term Frequency-Inverse Document Frequency, here, document refers to review) weighting (Salton, Wong and Yang, 1975). TF-IDF weighting is used to place less weight on more frequent terms and more weight on less frequent terms.²

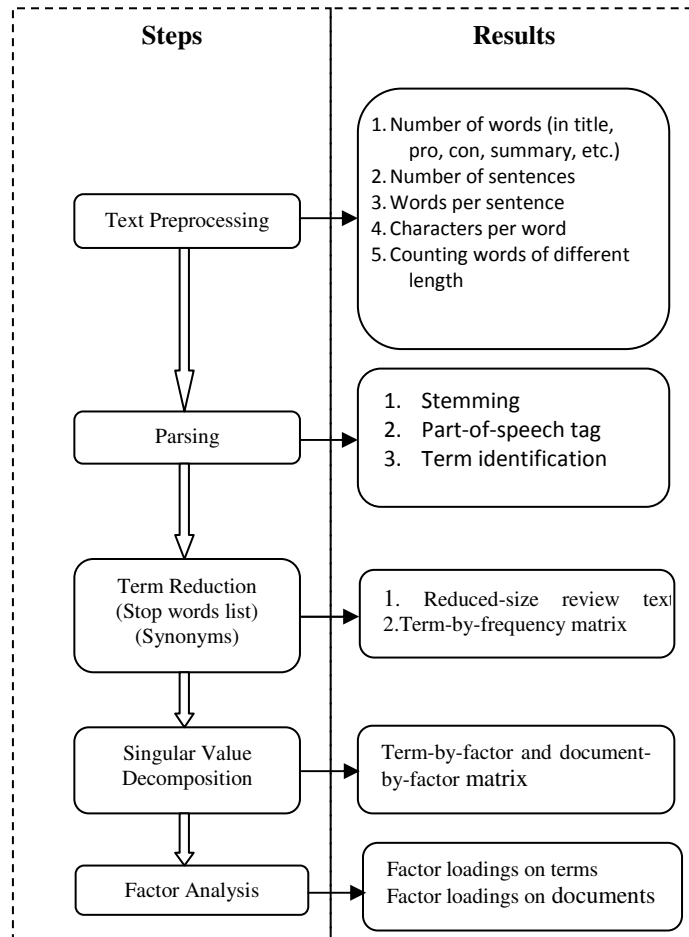


Figure 1. Text Mining Methodology

Singular Value Decomposition (SVD)

Although the term reduction process reduced a large number of terms, there were still too many terms remaining (3,457 terms in our case). Singular value decomposition (SVD) was then implemented to reduce the dimensionality of the transformed term-by-frequency matrix. With SVD, a matrix can be decomposed into the product of three matrices. “One matrix describes the original row entities as vectors of derived orthogonal factor values, another describes the original column entities in the same way, and the third is a diagonal matrix containing scaling values such that when the three matrices are multiplied, the original matrix is reconstructed” (Landauer et al., p.26, 1998).³

Factor Analysis

Following the SVD process, we perform the factor analysis, where terms are grouped into factors and are given appropriate loadings. In this study, each SVD factor represents a summarization of words in reviews with similar properties in a higher dimension, which is distinct from the other SVD factors. The number of factors represents the number of semantic characteristics from the reviews. As the purpose of this study is to examine whether semantic characteristics as a whole affect

² Please refer to Salton, Wong and Yang (1975) and Han and Kamber (2006) for more details on TF-IDF method.

³ More information regarding SVD can be found in (Landauer et al., 1998).

the number of helpfulness votes, we tried several numbers of factors, namely, 50, 100, 150, and 200. For demonstration purposes, we report the 100-factor solution in this paper. The 100 SVD factor loadings for each review are what we mean by “semantic characteristics” of reviews in our research context. To put it differently, each review now has 100 new variables to describe its semantic characteristics, and these 100 variables will be used in the subsequent analyses.

Empirical Models

Using ordinal logistic regression (OLR) models, we investigate relationships between three types of characteristics of online reviews and the number of helpfulness votes that those reviews receive. The dependent variable in our study is not whether or not a review receives helpfulness vote(s) as in the binomial logistic regression, but a “helpfulness rank” based on the number of votes a review receives. An OLR model is particularly suitable for our study because we are interested in not only whether a review receives at least one vote or not, but also what review characteristics lead to more helpfulness votes. Equations 1 and 2 depict the basics of our OLR approach.

$$g(\text{Pr}(Y \leq i|x)) = \alpha_i + \beta'x, \quad i = 1, 2, \dots, k \tag{1}$$

Where Y is the dependent variable, the ranks are denoted by 1, 2 ... k, $\alpha_1, \alpha_2, \dots, \alpha_k$ are k intercept parameters, β is the vector of slope parameters and x is the vector of independent variables. The function g(μ) is called the link function that allows the response (μ) to be linearly related to the independent variables. The log-odds scale has been used as the link function as in the form of Equation 2 (Cox and Snell, 1989; Walker and Duncan, 1967).

$$\begin{aligned} g_i(\text{Pr}(Y \leq i|x)) &= \ln \frac{P_r(Y \leq i|x)}{P_r(Y > i|x)} = \ln \frac{P_r(Y \leq i|x)}{1 - P_r(Y \leq i|x)} \\ &= \ln \frac{\phi_1(x) + \phi_2(x) + \dots + \phi_i(x)}{1 - (\phi_1(x) + \phi_2(x) + \dots + \phi_i(x))} = \alpha_i + \beta'x, \tag{2} \\ & \quad i = 1, 2, \dots, k \end{aligned}$$

Where $\phi_i(x)$ is the probability of being in class i given x.

In our context, Y is the number of votes on helpfulness. “0” to “6” denotes “0” to “6” number of votes respectively while “7” denotes “7 or more” votes.

In order to examine what and how different characteristics of reviews influence the number of votes they receive, we construct five OLR models with various combinations of the three types of characteristics. We explore whether the effect of various single characteristics remains unchanged and whether the addition of semantic characteristics has any additional effect on the performance of the model. Table 4 presents the five model descriptions.

Model	Description of Explanatory Variables (x)	# of independent variables
Model 1	Use only basic characteristics of the reviews (Basic)	5
Model 2	Use only stylistic characteristics of the reviews (Style)	19
Model 3	Use only semantic characteristics of the reviews (SVD factor loadings)	100
Model 4	Use basic characteristics + stylistic characteristics	24
Model 5	Use basic characteristics + stylistic characteristics + SVD factor loadings	124

Table 4. Model Descriptions

There are a large number of independent variables in the aforementioned five models. In particular, there are 100 SVD factor variables in models 3 and 5, which have much more variables than models 1, 2 and 4. Adding more variables in the model may increase the fit of the model, however, it may also decrease the model's predictive power. In order to make the models more parsimonious and comparable, we employ the stepwise variable selection method (Hocking, 1976). Thus, the final model should have a reasonable fit with fewer variables. For instance, we have only 18 independent variables in model 5 after the stepwise selection.

Model Comparison Criteria

To compare the models, we choose three widely used fit indices: misclassification rate, Akaike’s Information criterion (AIC), and lift ratio. Misclassification rate is often used to see how inaccurate the classification is. The larger the misclassification rate, the less accurate is the classification, and the poorer performance of the model is. Akaike’s Information criterion (AIC) is a measure of goodness-of-fit proposed by Akaike (1974). It describes the tradeoff between bias and variance in model construction, that is, between complexity and precision of the model. The smaller the AIC, the better the model is. The lift ratio is a widely used model accuracy measure in data mining literature. The lift ratio of a subset of the population is the ratio of the predicted response rate for that subset to the predicted response rate for the population. As a rule of thumb, the larger the lift ratio, the better the performance of the model will be.

RESULTS AND DISCUSSION

Table 5 summarizes the empirical results of comparing the five OLR models. Model 5, which combines all three characteristics of reviews, has the lowest misclassification rate and AIC index with the highest lift ratio; therefore it has the best performance among all models. The results indicate that integrating semantic characteristics into the model along with basic and stylistic characteristics significantly enhances the performance of the model. Such a finding suggests that semantic characteristics play a very important role in influencing the number of helpfulness votes a review receives. Results demonstrated in Table 5 also indicate that stylistic characteristics are the least critical criterion to encourage helpfulness votes from users.

Model	Misclassification Rate	Akaike's Information Criterion (AIC)	Lift Ratio
Model 1: Basic characteristics (Basic)	0.48301	9863.15	5.92
Model 2: Style characteristics (Style)	0.48589	9920.73	4.32
Model 3: SVD factors (SVD)	0.48560	9990.31	4.64
Model 4: Basic+Style	0.48243	9792.87	5.76
Model 5: Basic+Style+SVD	0.47753	9599.26	7.36

Table 5. Fit Statistics for Model Comparison

As discussed previously, model 5, which contains all three types of characteristics, has the best fit. In order to examine exactly what characteristics have the key impact on the number of helpfulness votes, we investigate model 5 further to include the parameter estimates, standard error, Wald test statistics for the parameters and corresponding P values.⁴ After applying stepwise OLR, 18 variables are included in model 5, including one control variable (whether the software is free or free to try), three basic characteristic variables (days since posting, difference between reviewer rating and average rating, and whether summary has content or not), three stylistic characteristic variable (number of 4-letter words, number of words in cons, and number of words in title), and eleven semantic variables.

First, we find that the P values for the estimates of control, basic, and stylistic variables are all smaller than 0.01, indicating they are statistically significant at 99% confidence level. The P values for six of eleven semantic characteristics are also smaller than 0.01, demonstrating statistical significance at 99% confidence level. The P values for the other five semantic variables are smaller than 0.05, illustrating statistical significance at 95% confidence level.

Second, it is also interesting to find that the estimate for “number of words in cons” is positive, indicating that the more words in the cons part of the review, the more helpfulness votes the review is likely to receive. Online purchase is a risky venture for most consumers; hence they tend to pay more attention to the negative part of reviews on the product as shown in "cons" part. As a result, more words in "cons" part of the review may encourage more people to read it and then vote on it. This finding is consistent with widely known "negativity bias" effect in psychology, which states that there is a generally bias in humans to give greater weight on negative entities (Rozin and Royzman, 2001).

⁴ The detailed estimates of the parameters and p values of the 18 variables are available upon request.

Third, it is also worth noting that the estimate for “the difference between reviewer rating and the average rating” is positive, which indicates that the larger the difference between reviewers’ rating and the average rating on the review, the more votes the review is likely to receive. This difference represents the extremeness of the review. A review that is drastically different from the average reviews (more extreme) is more likely to stand out and attract significantly more attention from the users.

Fourth, it is unusual for us to find that the estimate for “days since posting” is negative, indicating that the longer the review has been posted, the fewer votes it is likely to receive. This could be due to the ranking mechanism of the website. CNETD ranks the most recent review first by default. The older reviews are listed near the end of the list, hence reducing their chances of being spotted by viewers. Providing more ranking options or changing the default review listing for viewers might increase the exposure of more reviews to users.

Finally, the estimates for some of the SVD variables are positive while others are negative, indicating that certain words have positive impact on encouraging helpfulness votes while other words have negative impact. In our study, we do not attempt to explore the details of SVD factors, and instead examine semantic characteristics as a whole with an emphasis on their effect on the number of helpfulness votes that reviews receive. The empirical results show that semantic characteristics do have a significant impact on the number of helpfulness votes, with certain words encouraging more votes while other words discouraging votes.

CONCLUSIONS

In this paper, we examine a previously ignored yet important research question concerning the online user reviews: Why do some reviews not receive any votes on their helpfulness, while other reviews receive many votes? We address this question by investigating the impact of various characteristics of online user reviews on the number of helpfulness votes that reviews receive. We categorize characteristics of online reviews into three types, namely, basic, stylistic and semantic. Text mining techniques and ordinal logistic regression models are employed to investigate more than 3,400 online reviews of 87 different software programs from CNET Download.com. A number of practical and scholarly implications can be derived from this study.

This study is complementary to the previous studies on the helpfulness of online user reviews. Helpfulness votes on user reviews help online users locate the most potentially helpful reviews more efficiently and effectively for their decisions. Previous studies made efforts to predict the helpfulness of reviews without any helpfulness votes for users to make informed decisions. This study approaches from a new perspective to facilitate users' decision making. Rather than predicting the helpfulness, this study examines the factors influencing the number of helpfulness votes reviews receive. These perspectives are both important in helping online users get quality information efficiently.

This study also has significant implications for website designers in that it can guide them in designing helpfulness voting mechanisms that may garner more helpfulness votes. Our findings indicate that some design features encourage voting while others curtail viewers’ voting intention. For example, semantic characteristics have the most impact on the number of helpfulness votes that reviews received. Websites could provide incentives to encourage reviewers to write more meaningful comments. In addition, websites could provide more ranking options (e.g. based on the extremeness of opinion) to rank the reviews instead of ranking the most recent reviews first.

Our findings may have important implications for behavioral researchers by providing a new perspective on the online user voting behavior. In addition to confirm the widely known "negative bias" effect, our results show that the reviews with the most extreme opinions have a higher probability of getting more votes, suggesting that people tend to pay more attention to those extreme opinions. The effect of “extreme opinions” on other’s attention has not been fully investigated in previous research and would be an interesting area for future research.

REFERENCES

1. Akaike, H. (1974) A new look at the statistical model Identification, *IEEE Transactions on Automatic Control*, 19, 6, 716-723.
2. Brynjolfsson, E. and Smith, M. D. (2000) Frictionless commerce? A comparison of Internet and conventional retailers, *Management Science*, 46, 4, 563-585.
3. Cox, D. R. and Snell, E. J. (1989) *Analysis of binary Data*, Chapman & Hall, London,
4. Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. and Harshman, R. (1990) Indexing by latent semantic analysis, *Journal of the American Society for Information Science*, 41, 6, 391-407.

5. Duan, W., Gu, B. and Whinston, A.B. (2008) Do online reviews matter? - An empirical investigation of panel data, *Decision Support Systems*, 45, 4, 1007-1016.
6. Forman, C., Ghose, A. and Wiesenfeld, B. (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets, *Information Systems Research*, 19, 3, 291-313.
7. Hocking, R. R. (1976) The analysis and selection of variables in linear regression, *Biometrics*, 32, 1-49.
8. Jones, Q., Ravid, G. and Rafaeli, S. (2004) Information overload and the message dynamics of online interaction spaces: A theoretical Model and Empirical Exploration, *Information Systems Research*, 15, 2, 194-210.
9. Kassirjian, H. H. (1977) Content analysis in consumer research, *Journal of Consumer Research*, 4, 1, 8.
10. Kim, S. M., Pantel, P., Chklovski, T. and Pennacchiotti, M. (2006) Automatically assessing review helpfulness, *Proceedings of Conference on Empirical Methods in Natural Language Processing*, Morristown, NJ, USA, 423-430.
11. Kolbe, R. H. and Burnett, M. S. (1991) Content-analysis research: An examination of applications with directives for improving research reliability and objectivity, *Journal of Consumer Research*, 243-250.
12. Landauer, T. K., Foltz, P. W. and Laham, D. (1998) An introduction to latent semantic analysis, *Discourse Processes*, 25, 259-284.
13. Liu, Y., Huang, X., An, A., Yu, X., Giannotti, F., Gunopulos, D., Turini, F., Zaniolo, C., Ramakrishnan, N. and Wu, X. (2008) Modeling and predicting the helpfulness of online reviews, *Proceedings of Eighth IEEE International Conference on Data Mining*, Piscataway, NJ, USA, 443-452.
14. Rozin, P. and Royzman, E. B. (2001) Negativity bias, negativity dominance, and contagion, *Personality and Social Psychology Review*, 5, 4, 296-320.
15. Salton, G., Wong, A. and Yang, C. S. (1975) A vector space model for automatic indexing, *Communications of the ACM*, 18, 11, 620.
16. Spool, J. (2009) The magic behind Amazon's 2.7 billion dollar question, available online at <http://www.uie.com/articles/magicbehindamazon/>.
17. Walker, S. H. and Duncan, D. B. (1967) Estimation of the probability of an event as a function of several independent variables, *Biometrika*, 54, 1, 167-179.
18. Wei, C. P., Yang, C. C. and Lin, C. M. (2008) A latent semantic indexing-based approach to multilingual document clustering, *Decision Support Systems*, 45, 3, 606-620.