



On the relationship between number of votes and sentiment in crowdsourcing ideas and comments for innovation: A case study of Canada's digital compass



Daniel E. O'Leary

University of Southern California, Marshall School of Business, Los Angeles, CA 90089, United States

ARTICLE INFO

Article history:

Received 18 December 2015

Received in revised form 20 May 2016

Accepted 21 May 2016

Available online 31 May 2016

Keywords:

Crowdsourcing

Innovation

Number of votes

Sentiment analysis

ABSTRACT

Recently, PWC facilitated an innovation crowdsourcing effort entitled “Canada’s Digital Compass” aimed at generating creative directions for Canada to pursue in its future. As part of this effort, over 70 ideas were submitted and over five hundred votes on those ideas were generated. This paper analyzes data developed as part of that effort, investigating a number of relationships between different variables elicited from the analysis and the number of votes for those ideas. The analysis confirms some previous results, provides a structure for interpreting those results and generates some new results. In particular, this paper generates sentiment measures for both the innovation ideas and for the comments. Using a decision tree approach we find that the number of comments and the extent to which the sentiment in those comments is positive are statistically significantly related to the number of votes. In addition, using regression analysis, this paper finds that the number of votes is statistically significantly related to the interaction between sentiment measures for ideas and comments. Finally, a rationale for these results, based on game theory is proposed and investigated.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Recently, there has been substantial interest in gathering the “wisdom of the crowd” (e.g., [19]). One application that has generated substantial interest is in the area of innovation (e.g., [5]). In particular, idea gathering and rating systems provide the opportunity to solicit information from the crowd about innovations. Such systems allow users to suggest ideas, vote on ideas and make comments on ideas. In practice, voting platform systems, such as Chaordix¹ and others have been developed and implemented in different environments to help organizations facilitate idea generation and enhancement. Such voting systems generally are used to gather ideas and information about ideas rather than facilitate any co-creation of ideas by creating virtual teams or groups.

Typically, the information that is gathered is immediately available to the crowd. Ideas are formulated and placed on line virtually immediately after their submission. Participant comments are made available to the other participants in real time and information about who made the comments is broadly available to all. User votes are then accumulated for the different ideas. As a result, as part of their use, these systems typically generate information about votes, ideas and comments. The number of votes for an idea provides a measure of the relative support for the innovation, the relative quality or the innovativeness of the idea, by the crowd, as compared to the other ideas. Accordingly, these

systems provide decision support information to management and can help management focus on the better ideas. Thus, there is interest in understanding what factors relate to the number of votes that an innovation idea generates and how these variables are related.

Some characteristics and relationships between the number of votes, and idea and comment information have been analyzed by previous researchers. Gangi and Wasko [5] as part of an innovative analysis of using crowdsourcing for innovation using Dell’s IdeaStorm found the strongest correlation (0.74) between “total number of votes” was with “total number of comments.” In particular, that relation was positive and statistically significant at better than .01.

However, that analysis also raised a number of additional questions. First, the analysis was a case study analyzing one set of innovations. As a result, there is interest in determining if that relationship holds in other settings. Second, it is not clear a priori why that relationship would hold in crowdsourcing votes on innovations. Why would voters tend to comment on those innovation ideas that they vote for? For example, it would seem that ideas with few votes could easily generate a substantial number of comments detailing how and why the ideas are not effective and why people did not or should not vote for them. In that case there would be a negative relationship between the number of comments and the number of votes. Accordingly, the empirical result suggests that voters largely limit their comments to those ideas that they also voted for. As a result there is interest in finding a theory or theories that might help explain that result. Third, although that analysis suggests that a model aimed at estimating the total number of votes

E-mail address: oleary@usc.edu.

¹ <http://www.chaordix.com/>.

for innovations, likely would include the total number of comments, it is not clear what other variables might be related to the number of votes. For example, does information about the idea count, is it related to the number of votes? Fourth, previously researchers in innovations have focused primarily on the quantitative variables that count the number of something, such as “number of comments,” and others. Accordingly, previous researchers have ignored information that relates to issues such as the sentiment of the ideas and the comments on the innovations. As a result, it is unclear how idea or comment sentiment might relate to the number of votes for innovations.

Accordingly, the purpose of this paper is to extend the research of Gangi and Wasko [5] to investigate those four issues associated with using the crowd to vote on innovations. Although Gangi and Wasko [5] did find a relationship between the number of comments and the number of votes, they did not propose a theory, discuss idea analytics or analyze the qualitative variables. Accordingly, this paper investigates those issues within a crowdsourcing idea voting system. The system used was Canada's Digital Compass project which was moderated by PWC. This resulted in the following findings.

1.1. Findings of this paper

These questions lead to a number of findings. First, this paper finds a similar result to the previous literature for the correlation between the number of comments and the number of votes, with a positive and statistically significant result. Second, this paper investigates the relationship between number of votes and comments, in particular, those ideas that get comments (compared to those that do not). Voting systems provide a number of capabilities that have led to changes in transparency in information about content and contributors with some interesting implications. As a result of this transparency, the game theoretic notion of “tit-for-tat” is suggested as a potential rationale for why more popular ideas get comments and less popular ideas do not receive as many comments. Tit-for-tat is a notion that equivalent actions will be given in return. Thus if you give a negative comment on my idea, then I would give a negative comment on your idea, or conversely. This paper suggests that because of the transparency of such systems, crowdsourced voters tend to comment mostly on the high vote getting ideas and provide largely positive comments. Third, this paper finds that information about the idea is related to the number of votes. In particular, the number of words in the idea is related to the number of votes. Fourth, although such crowd voting systems provide many different structured variables to examine (e.g., number of comments, number of words, etc.) such variables do not account for the unstructured information resulting from such systems, e.g., the sentiment content of the ideas or the comments. As a result, this paper uses sentiment analysis in order to provide some structured analysis on unstructured text data as part of the analysis of the number of votes that an idea gets.

1.2. This paper

As a means of structuring the analysis of these findings, this paper proceeds as follows. Section 1 has reviewed and motivated the paper. Section 2 provides a brief review of some background concepts, such as crowdsourcing and a review of some of the previous literature in crowdsourcing and the number of votes. Section 3 provides some background information on the Digital Compass project. Section 4 summarizes the data and the variables used in this analysis. In particular, this section provides structured analytics to analyze unstructured text. Section 5 summarizes the expected relationships between the data variables and the number of votes that an idea gets. Section 6 investigates some of the findings from the summary data, the correlation matrix and the regression analyses between idea and comment variables as used to estimate the number of votes. Section 7 briefly summarizes the paper, reviews its implications, contributions and investigates some extensions.

2. Background and selected previous research in innovations and crowdsourcing

The notion of crowdsourcing was examined by Howe [10] and others, who distinguished between crowdsourcing and outsourcing. In particular, Howe [10] called crowdsourcing “... the process by which the power of the many can be leveraged to accomplish feats that were once the province of a specialized few.” While quoting a Vice President of Innovation at Procter & Gamble, Howe [10] notes “Outsourcing is when I hire someone to perform a service and they do it and that's the end of the relationship. That's not much different from the way employment has worked throughout the ages. We're talking about bringing people in from outside and involving them in this broadly creative, collaborative process. That's a whole new paradigm (crowdsourcing).”

Starting at least with IBM's Innovation Jams in 2007 [2,8] there has been interest in using crowdsourcing systems to facilitate innovation. As an example, Bjelland and Wood [2] focused on documentation of the process and outcomes. O'Leary [15] investigated a case focusing on Accenture's use of crowdsourcing to develop and rate innovations. As another example, Hossain and Islam [9] investigated Dell's IdeaStorm, focusing on the extent of implementation of proposed ideas.

However, there seems to have been limited empirical analysis of the relationships between key variables and the number of votes, associated with these crowdsourcing innovation system efforts. In particular, our scope of analysis of the previous literature is limited to the task of using crowdsourcing software for innovation and the resulting relationships between different characteristics and the number of votes an idea gets. Baily and Horvitz [1] investigated a number of variables related to innovation management systems, but provided limited analysis of relationships between sets of variables. Gangi and Wasko [5] found a statistically significant correlation between number of comments and number of votes, when they examined the broader innovation process. Similarly, Fuller et al. [4] also found that the number of comments was statistically related to the number of votes. Although other innovation researchers have investigated the correlation between different variables their analyses have not considered emerging variables, such as sentiment or the relationship between sentiment-based characteristics of ideas and comments.

Further, previous researchers have not fully examined the potential effects of transparency on the use of such crowdsourcing systems. Instead, previous research with transparency and voting systems has largely been concerned with government transparency [13]. However, in this paper we investigate the potential impact on innovation systems of requiring comments and ideas to be attributed to particular people, rather than allowing anonymity. Specifically, including identities of the participants on the ideas and the comments removes asymmetries of information.

2.1. Use of sentiment analysis in other settings: blogs, product reviews and on-line forums

There are a number of differences with the use of sentiment in crowdsourcing for innovation, in contrast to blogs, product reviews and other settings. In crowdsourcing for innovation, there is both an idea and the comments on the idea. As a result, there is a question as to the impact of sentiment from the idea statement, sentiment from the comments and potential interaction effects. In addition, it is not clear that settings designed for gathering innovations and innovation information from the crowd is directly comparable to other settings. In particular, different settings provide unique context capabilities that likely can be leveraged.

Although this paper focuses on using crowdsourcing for innovation, sentiment analysis has been used in the analysis of a range of different problem settings. Li and Wu [12] investigated sentiment in the analysis of on-line forums. O'Leary [14] investigated the use of sentiment analysis in blogs. Saleham and Kim [18] used sentiment analysis in online

Download English Version:

<https://daneshyari.com/en/article/552370>

Download Persian Version:

<https://daneshyari.com/article/552370>

[Daneshyari.com](https://daneshyari.com)