Contextual polarity and influence mining in online social networks

Hassan Alzahrani[1*] , Subrata Acharya[1], Philippe Duverger[2] and Nam P. Nguyen[1]

*Correspondence:
halzah1@students.towson.edu
[1] Department of Computer &
Information Sciences, Towson
University, Towson, MD
21252, USA
Full list of author information
is available at the end of the
article

Abstract

Crowdsourcing is an emerging tool for collaboration and innovation platforms. Recently, crowdsourcing platforms have become a vital tool for firms to generate new ideas, especially large firms such as Dell, Microsoft, and Starbucks, Crowdsourcing provides firms with multiple advantages, notably, rapid solutions, cost savings, and a variety of novel ideas that represent the diversity inherent within a crowd. The literature on crowdsourcing is limited to empirical evidence of the advantage of crowdsourcing for businesses as an innovation strategy. In this study, Starbucks' crowdsourcing platform, Ideas Starbucks, is examined, with three objectives: first, to determine crowdsourcing participants' perception of the company by crowdsourcing participants when generating ideas on the platform. The second objective is to map users into a community structure to identify those more likely to produce ideas; the most promising users are grouped into the communities more likely to generate the best ideas. The third is to study the relationship between the users' ideas' sentiment scores and the frequency of discussions among crowdsourcing users. The results indicate that sentiment and emotion scores can be used to visualize the social interaction narrative over time. They also suggest that the fast greedy algorithm is the one best suited for community structure with a modularity on agreeable ideas of 0.53 and 8 significant communities using sentiment scores as edge weights. For disagreeable ideas, the modularity is 0.47 with 8 significant communities without edge weights. There is also a statistically significant quadratic relationship between the sentiments scores and the number of conversations between users.

Keywords: Social networks, Sentiment analysis, Generalized least squares method, Starbucks, Community detection

Introduction

To succeed with products requires an understanding of customers' needs [1–3]. Hence, it is essential to understand how and why knowledge of customer needs can maximize a firm's profit. This knowledge would allow firms to effectively manage and optimize the outcomes of customer involvement in a crowdsourcing platform. For example, on April 1, 2010, a customer posted an idea in the "innovation community" category on the Starbucks crowdsourcing platform, Ideas Starbucks, about payment with a Starbucks mobile card. Later, in 2011, Starbucks implemented this idea and started to provide mobile payment options.

Another example is the backlit keyboard, an idea posted on February 21, 2007, by a customer in the accessories category on Dell's crowdsourcing platform. Now, backlit keyboards are commonly available on most of their laptop computers. Starbucks and Dell believe that being close to consumers can be a profitable source of knowledge, critical for new product and service development, that cannot be attained from other sources [4–6]. On the other hand, the shutdown of several crowdsourcing platforms is correlated with a low implementation rate of ideas and also seems to be consistent with the arguments against the crowdsourcing [7].

Problem description

Crowdsourcing platforms have emerged as a method of outsourcing work to consumers not directly affiliated with the firm [8]. These platforms possess the capacity to overcome limitations in traditional business research, such as small participant sample sizes and narrow participant demographic backgrounds [9]. While the benefits of crowdsourcing platforms are salient to many businesses and academics, the research on methodological approaches to extract robust data from these platforms is somewhat polarized between qualitative and quantitative researchers. There is a salient paucity in business research on crowdsourcing that presents the benefits of multiple methodological approaches to effectively extract this data to benefit a firm's product differentiation strategy.

While generalized linear mixed models and social network analysis have become common methodological approaches in crowdsourcing research, some scholars fail to consider other methods, such as text mining and sentiment analysis, as robust methodological approaches. Both text mining and sentiment analysis are rooted in the disciplines of sociology, anthropology, and psychology. These methodological approaches are manifestations of theories of affective stance and appraisal, which focus on emotions shaping cognition [10]. The main objective of this study is to examine and explicate how companies can benefit from crowdsourcing platforms by using a multitude of empirical methods, such as text mining, sentiment analysis, social network analysis, and generalized linear mixed models, to generate new product ideas.

Research objectives

The research presented here has three primary objectives. The first is to examine the users' perception of the company. Viable ideas stemming from crowdsourcing initiatives are typically influenced from a user's experience or perception of the firm. When a user's experience or perception of a firm is positive, ideas tend to be more constructive and feasibly implementable [7]. The second objective is to identify which communities of users generate the best ideas and which communities generate the worst ideas. Crowdsourcing platforms allow users to either promote, through "likes", or demote, through "dislikes", the ideas of their peers [7]. Therefore, communities of users with the largest number of likes are deemed to have the best ideas, while those with the largest number of dislikes are deemed communities that generate the worst ideas. Poorly performing communities can be isolated so that greater attention is paid to better ideas, accelerating the evaluation. The third objective is to extrapolate the sentiment of discussions between two different groups of crowdsourcing users based on the frequency of their conversations. The frequency of word usage is a common method to characterize the type and

degree of sentiment of users. Similarly, the frequency of interactions among users can help researchers determine both type and degree of sentiment among crowdsourcing platform users [11].

These objectives were addressed with three empirical frameworks. In the first, sentiment analysis was used to calculate and categorize sentiment and emotion scores for successful and unsuccessful ideas to show that the platform's users had an accurate impression of the company. In the second framework, the sentiments and emotions were used to construct the communities using social network analysis. The communities were linked by users identified as idea launchers. The third framework was constructed to show the relationship between sentiment and the number of messages exchanged by users. Agreeable ideas are the top 5% while disagreeable ideas are the lowest 5%. The relationship was positive for agreeable ideas and slightly less so for disagreeable ideas.

Literature review

Crowdsourcing

The concept of harvesting ideas from the crowd has been around for centuries, including the longitude prize, offered by the British government in 1714 to anyone who could solve the practical problem of measuring longitude at sea [12]. Both the physical and digital worlds are connected, and data are readily available to firms beyond their immediate workforce. Now, firms possess the ability to reach out to the masses for ideas that can be commercialized [13]. Web 2.0 technologies have radically changed the way people communicate on the Internet. Crowdsourcing is increasingly performed via the Internet, enabling a plurality of contributors from around the globe to harvest ideas [14].

With these technologies, several terms were invented, among them, crowdsourcing, which was used for the first time by an anonymous user on an Internet forum 10 years ago [15]. The term was popularized by journalist Jeff Howe in 2006, in his article published in the online magazine Wired [16]. Howe [17] defined crowdsourcing as an act of outsourcing a task to a large and undefined group of people, coordinating the knowledge of the group with those who need it. Howe wondered if many solutions to our problems were already present in the wisdom of the crowd, just waiting to be uncovered. Later, Estellés-Arolaset al. [18] created a more comprehensive definition that conceptualized crowdsourcing as a participative online activity, in which individuals or firms propose a voluntary task to a group of individuals with varying knowledge bases, demographic heterogeneity, and size. The resulting process is mutually beneficial as users receive satisfaction from a given need, such as economic, social recognition, self-esteem, or skill development. The crowdsourcer obtains work, money, knowledge, experience, and other advantages from the crowd.

Hosseini et al. [19] describe the taxonomy of features that characterize crowdsourcing into its four constituent parts, which are the crowd, the crowdsourcer, the crowdsourced task, and the crowdsourcing platform. The crowd consists of the people that participate in crowdsourcing activities. Crowds are characterized by diversity, anonymity, size, undefinedness, and suitability. A *crowdsourcer* is an individual, an institution, or a firm that seeks the inherent power in crowds to complete a specific task. Often, incentive provisions, open calls, ethicality provisions, and privacy provisions are put into place by crowdsourcers to create parameters of conduct to elicit useful ideas. The *crowdsourced*

task is the outsourced activity that is provided by the crowdsourcer. Often, it is in the form of a problem, an innovative model, a data collection issue, or a form of fundraising. A *crowdsourced task* requires the expertise, experience, ideas, knowledge, skills, technologies, or money from the crowd. The features of a crowdsourced task are modularity, complexity, solvability, automation, and user-drivenness. Finally, the *crowdsourcing platform* is where the actual crowdsourcing task takes place. The platform is usually a website or an online venue where crowd-related interactions take place.

The literature describes many forms of crowdsourcing [20–22]. A general taxonomy of the modern forms of crowdsourcing was proposed by Doan et al. [23], distinguishing between explicit and implicit crowdsourcing forms. In the explicit form, companies ask for contributions directly (e.g., Ideas Starbucks, IdeaStorm, Amazon Mechanical Turk). In the implicit form, companies embed tasks to motivate users to participate. In the taxonomy created by Doan et al. [23], users' implicit participation can be categorized as standalone or can be piggybacked onto another platform. In standalone implicit crowdsourcing, companies use the input provided by the users to solve a problem that is related to the issue of the platform (e.g., the ESP game, the Peekaboom game, and reCAPTCHA). In the piggyback form, companies gather and retrieve information using third-party websites, such as search engines (e.g., Google, Yahoo, and Bing!). The content generated by users is used, for example, for product recommendation, spelling correction, and keyword generation [15, 23]. There are other taxonomies offered for crowdsourcing: internal and external [15]; micro-tasks, macro-tasks, simple projects, and complex projects [24]; and integrative and selective tasks, routine tasks, complex tasks, and creative tasks [25].

Howe [17] proposed another taxonomy of crowdsourcing forms, distinguishing collective intelligence, crowd creation, crowd voting, and crowdfunding. Howe [17] conceptualizes *collective intelligence* as a group of individuals collaborating to create synergy. Ultimately, this synergy creates something more significant than its constituent parts. *Crowd creation* is the most common form of crowdsourcing and emphasizes solving a particular dilemma through satisfactory solutions, in the form of tangible products, to a specific problem. The most significant output from the crowd creation process is an end product of either intellectual or physical form that holds values to others [14, 17]. *Crowd voting* is considered to be among the most popular forms of crowdsourcing that have the highest participation rate among all forms of crowdsourcing. It involves leveraging the community's judgment to elicit ratings of products and services. *Crowdfunding* is a form of alternative finance that consists of the funding of a project or business venture by raising monetary contributions from a large number of people.

Since the adoption of Web 2.0, the majority of multinational companies have established a crowdsourcing platform for their business to obtain customers' opinions and ideas, to facilitate communication both with and between the customers, and to enhance the loyalty of customers, and increase product recognition [26].

Sentiment analysis

Opinions are critical influencers of behavior and central to most human activity. Beliefs, perceptions of reality, and decisions that we make are conditioned upon how others perceive and evaluate the world around them. An integral part of the decision-making

process is seeking out the opinion of others. Not only is this true of individuals, but also organizations [27]. Sentiment analysis, or opinion mining, is a social science methodology that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward products, services, and organizations [27, 28].

Sentiment analysis began earlier in 2000 [27, 29, 30], but it was not until [31] used the term that it became widespread. Meanwhile, Dave et al. [32] introduced the term "opinion mining" for the same activity. As a methodological tool, sentiment analysis, or opinion mining, is a machine-learning approach that extracts opinions, sentiments, appraisals, attitudes, and emotions toward entities (e.g., products, services, and topics) and their attributes (e.g., picture quality, battery life, quality of service, and support) from a text [30]. This methodology is commonly seen as a subarea of natural language preprocessing because it contains lexical semantics, co-reference resolution, word sense disambiguation, discourse analysis, information extraction, and semantic analysis [27]. Sentiment analysis is used in a vast number of domains, including marketing, finance, health, tourism, politics, and social science.

In addition to the business applications, sentiment analysis has research applications. For example, [33] and [34] show that positive sentiment is a better predictor of movie success than simple buzz; Liu et al. [35] use sentiment to predict box-office revenue; [36] conclude that negative emotion significantly affects innovation activities in the brand community (i.e., on a crowdsourcing platform); Lee et al. [37] extract idea content characteristics, such as subjectivity and negativity, using Starbucks crowdsourcing, which indicates that the subjectivity and negativity of ideas have a positive impact on user agreement and organization adoption. Pestian et al. [38] predict the suicide of patients based on anonymized clinical tests and annotated suicide notes based on the assignment of emotions to suicide notes. Zhang et al. [39] identify positive and negative public moods on Twitter and use them to predict the movement of stock market indices. The examples of sentiment analysis are prevalent in several business and non-business-related literature, indicating the prevalence of the technique among academics from various fields.

Sentiment can be classified as linguistic-based, psychology-based, and consumer research-based [40]. The latter conceptualization is selected for simplicity [27, 40], classifying sentiment as rational or emotional. Rational sentiment consists of rational reasoning, tangible beliefs, and utilitarian attitudes [27, 40] and can be classified as positive, negative, or neutral. Emotional sentiments are non-tangible and originate in people's psychological states of mind [27, 40]. While there are many classification systems for emotions, we used the one created by [41], who proposed eight evolutionary-created emotions: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. This classification system was chosen for its parsimony and seminal place in the literature stream. Several contemporary taxonomies of sentiment analysis maintain roots in [41].

There are three different levels on which sentiment analysis can be performed depending on the study requirements: at the document level, in which each document (e.g., product review, idea, comment) is considered as basic unit of information. Entities and aspects inside the document are not studied, and sentiments expressed about them are not determined. However, document-level sentiment analysis is less meaningful because the author's opinion may be positive about some entities and negative about others, for

example, *"Jane has used this camera for a few months. She said that she loved it. However, my experience has not been great with the camera. The pictures are all quite dark."* [42]. At the sentence level, each sentence is treated as a short document. However, sentence level has two classes: subjective (i.e., an opinionated sentence) and objective (i.e., a not-opinionated sentence). At the feature level, also called *aspect level*, each piece of text is identified as a feature of some product and is based on the idea that an opinion consists of a sentiment and a target. In our study we used sentence-level sentiment analysis because it allows for the removal of objective sentences that are assumed to imply or express no opinion or sentiment. Sentiment scoring is a process to calculate sentiment and polarity by matching words back to a dictionary of words flagged as "positive", "negative", or "neutral".

Social network analysis

The second objective establishes the discovery of community structure in the network of users proposing ideas. This process should be part of decision-making by stakeholders of any company. Crowdsourcing platforms allow interaction among users who are proposing ideas. These interactions can be investigated as a network using graph theory. The graphs generated by the analysis are interactive, showing the network structure and the links that interconnect the structure. Social network analysis is used to study relationships among individuals, families, households, villages, and firms. In social networks, the actors can be modeled by a network structure consisting of vertices and edges. Vertices represent actors (known as nodes), and the edges (known as ties) represent the relationships between the vertices. The strength of a connection indicates how strong a relationship is [43].

In social network analysis, a social network is conceptualized as a graph when the relationships have no direction and as a digraph in the presence of direction. In a digraph, the edges are presented as arcs. The main goal of social network analysis is to detect and interpret patterns of social ties among actors, with community structure being an important property. In the classification of nodes into groups, the within-groups connections are dense, but between-groups connections are sparse [44]. The community structure in social network analysis is closely related to clustering and graph partitioning concepts. The identification of the optimal number of communities is a difficult task that depends on the algorithms used.

A network-based perspective is frequently used to analyze the key users in crowdsourcing [45, 46]. For example, Martínez-Torres [47] showed that social network analysis can be used in crowdsourcing to identify innovative users, defined as those users who post ideas that are potentially applicable to the organization [47]. Arenas-Marquez et al. [48] used social network analysis to identify influencers who can have a significant impact on the decision-making of other users based on the participation features of word-of-mouth crowdsourcing platforms [49]. Toral et al. [50] used social network analysis to identify the users who play the role of middleman among other users on a crowdsourcing platform [25]. Basole [51] stated that social network analysis can be used in the mobile ecosystem of crowdsourcing to help decision-makers to: (1) visualize the complexity of interfirm relationships and interactions among current and emerging mobile segment; (2) estimate how convergence influences ecosystem structure, and

(3) evaluate the firm's position relative to its competitors. These results can then be applied to improve innovation strategies or business models.

Social network analysis is used in different disciplines, such as computational biology, where researchers study systems of interacting genes, proteins, chemical compounds, or organisms [52]. Researchers in the field of finance have used social networks to analyze the interplay among world banks as part of the global economy [53]. In marketing, researchers often assess the extent to which product adoption is induced as a type of contagion [54]. Engineering scholars can utilize social network analysis to establish best practice designs to deploy networks of sensing devices [55]. The field of neuroscience uses social network analysis to explore voltage dynamic patterns in the brain associated with epileptic seizures [56]. Political science researchers use social network analysis to examine the evolution of voting practices when groups are faced with varying internal and external forces [57]. Finally, public health scholars study the spread of infectious diseases in populations and formulate plans of action to address those infections by employing social network analysis [58, 59].

Generalized least squares

The third objective examines the relationship between the ideas' sentiment scores and the discussion frequency of crowdsourcing users, for which we constructed a linear regression model. Sentiment scores calculate the text polarity sentiment at the sentence as measured by Hu et al. [60]. The parameters of the linear regression can be estimated using different approaches. The ordinary least squares method is widely used for the optimal linear unbiased estimation of the parameters of a linear regression. The validity of the inferences of a linear regression model via the ordinary least squares estimator depends on four assumptions for the residuals: zero conditional mean, independence, constant variance, and normality. Testing the assumptions is vital, and a violation can result in biased estimates of the relationships, confidence intervals, and significance tests [61]. The normality assumption is required for confidence intervals in small samples. For large samples, this assumption can be accepted because of the Central Limit Theorem. If the conditional mean of residuals is non-zero, the relationship between the outcome and the predictors is nonlinear and the regression coefficient may be biased [61]. The assumptions of independence and constant variance are known as homoscedasticity. If the assumption of homoscedasticity is violated, the estimates of standard errors, confidence intervals, and significance tests may be biased [61]. Nonlinearity is eliminated with the logarithmic transformation of variables. However, this linearization of the variables cannot eliminate the presence of assumption violations in ordinary least squares. In this study, any possible presence of non-independence of residuals and non-constant variance of residuals is eliminated by using the generalized least squares estimator [61, 62] for the linear regression. The generalized least squares method has the properties of being unbiased (the difference between the estimated value and real value of the parameter converges to zero), consistent (convergence to the real value of the parameter), and asymptotically normal (the probability distribution converges to the normal distribution) [61, 62]. Generalized least squares is the best linear unbiased estimator for the parameters. The parameters obtained with the generalized least squares approach fit a linear regression model [61].

When a database is too large for the available computing power, it is necessary to use the statistical technique of bootstrapping [63, 64], in which a set of several random datasets resampled from the original data are adjusted with generalized least squares to find appropriate parameters. The final model estimated is calculated as the average of all simulations with its associated standard error and the statistical significance of the parameters [64]. With the existence of an exorbitant amount of unlabeled natural language data and the lack of labeled data, bootstrapping has become a common practice among scholars [65]. We chose this method because of its foundations and evidence-based efficacy in computational linguistics and business scholarship.

Data collection and description

My Starbucks Idea is an online social platform (mystarbucksidea.force.com) that serves as a crowdsource, where users propose, comment, and vote on ideas of other users. Each proposed idea on this site contains information about the author, a score based on the number of votes received, and the comments of users. Each user earns points corresponding to the idea's rating. Highly rated ideas are reviewed and may be implemented by Starbucks.

The database we use was collected from the "My Starbucks Idea" website in December 2016, offering data from 2008 to 2016 containing 17 attributes. We employed web-scraping techniques, such as R, Rvest, and XML2 packages to harvest the data. Table 1 presents the description of each attribute with the number of valid and missing values. The total number of records is 340,811. There are 134,192 unique ideas with 340,811 associated comments. All the attributes related to ideas also have 134,192 unique entries, while the attributes of comments contain 340,811 records. Users are participative and let others comment on their ideas. On average, each idea receives 2.5 comments. Users are participative and let others comment on

Table 1 Variables in the My Starbucks Idea database

Variable	Description	Unique	Missing
id.idea	Unique identifier of each idea	134192	0
index.comment	Unique identifier of comments by idea	340811	0
title.idea	Title of the idea	134192	0
category	Categories of the ideas	340811	0
sub.category	Sub-categories of the ideas	340811	0
author.idea	Username of authors of ideas	134192	0
author.comment	Username of authors of comments	340811	0
text.idea	Text content of the idea	134192	0
text.comment	Text content of the comment	340811	0
date.time.idea	Date and time of publication of the idea	134192	0
date.time.comment	Date and time of publication of the comment	340811	0
state	State of residence of the author of the idea	46164	294347
city	City of residence of the author of the idea	46164	294347
member.since	Date of membership of the author of the idea	340811	0
idea.action	Actions of the author of the comment of the idea	134192	0
score.idea	Number of people who like minus dislike the idea	134192	0
points.earned	Points eared by the author of ideas	340811	0

Table 2 Ideas and comments by categories and subcategories

Categories	Sub-categories	Frequency of comments	Frequency of ideas
Product	Food	23,974	10,648
	Frappuccino Beverages	16,429	6173
	Merchandise Music	20,027	9846
	New Technology	19,828	6116
	Other Product	22,662	10,070
	Starbucks Card	33,181	11,606
	Tea Other Drinks	22,947	10,347
	Coffee Espresso Drinks	33,079	11,541
Experience	Atmosphere Locations	23,473	12,134
	Ordering Payment Pick Up	27,872	10,495
	Other Experience	29,634	10,209
Involvement	Building Community	19,488	6844
	Other Involvement	16,297	6261
	Outside USA	4212	2362
	Social Responsibility	27,708	9540
Total	340811	134,192	

Table 3 Idea/comment productivity in the analyzed five percent of agreeable and disagreeable ideas

	No. Users	Ideas		Comments		
		Total	Per user	Total	Per user	Average wd/idea
Agreeable ideas	6376	6710	1.05	51,232	8.04	34.15
Disagreeable ideas	5310	6710	1.26	30,975	5.83	38.27
	Likes / dislikes per idea					
	Min.	Average	Max.			
Agreeable ideas	190	1369	376,870			
Disagreeable ideas	− 710	− 197	−150			

their ideas. Users often do not prefer to disclose their geographic locations. Only 14% of users provided information on their city and state of residence. The rest of the database is complete without missing values. Categories group the ideas. There are three categories: Product, Experience, and Involvement. Each of these categories contains at least three subcategories. In Table 2, Product and its subcategory, Food, have the highest number of ideas and comments. In contrast, Involvement and its Outside USA subcategory have the lowest number of ideas and comments.

We selected only the top 5% of agreeable ideas and the lowest 5% of disagreeable ideas. The main reason behind this approach is to verify if the same authors are active in both groups. We find that there are differences between the two groups in terms of the quantity and quality of ideas (Table 3).
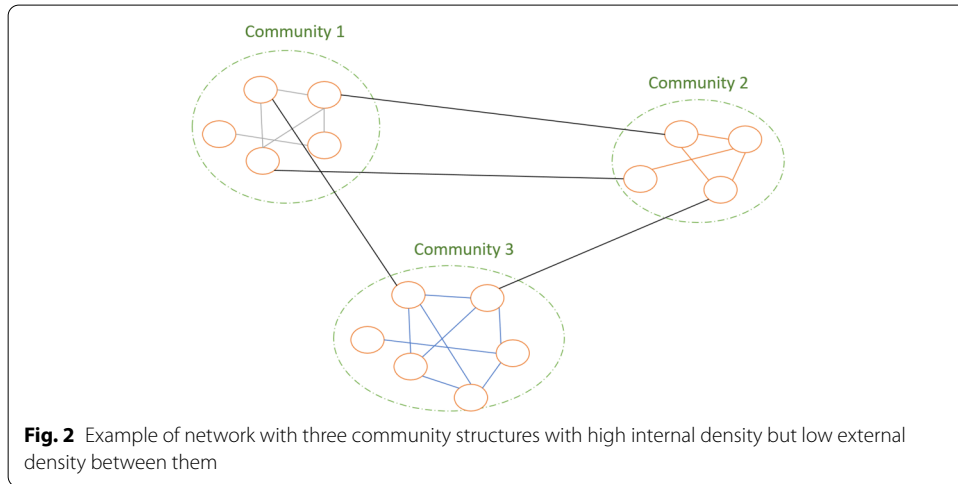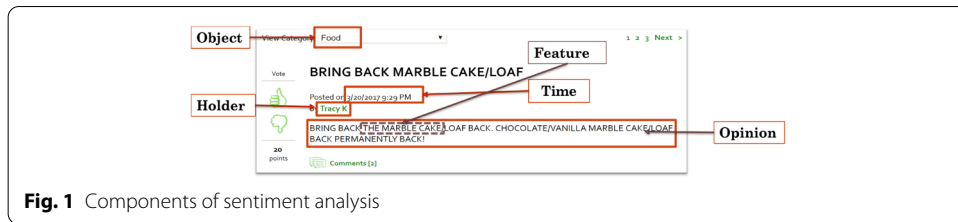
Methodology

Experimental setup

Each objective of this study requires a special design. For the first objective, determining the perception of the company by crowdsourcing participants when generating ideas on the platform, the database needs to be analyzed with special statistical tools. For the second objective, organizing users with community structure to locate users more likely to produce ideas, the statistical algorithm needs to be able to classify information in a way that meets this objective. The third objective, studying the relationship between the ideas' sentiment scores and the discussion frequency of crowdsourcing users, can be investigated once the classification of information is completed. The study is then split in two stages: classification and causation. The classification with community detection and popularity-sentiment approaches is performed first, then the causation part is classified using the generalized least squares technique.

The following sections present the description of the approaches used in this study. Section "Popularity sentiments" discusses popularity sentiments. A contribution of this study here is the utilization of three dictionary methods (natural language preprocessing, an augmented dictionary method, and a sentiment dictionary) to infer sentiments and emotions using their direction. Section "Community detection" presents the community-detection process. A contribution of this study in this section is the implementation of eight algorithms of modularity to obtain the best community-detection approach. Section "Generalized least squares" presents the generalized least squares approach. Here, a contribution of this study is the utilization of a non-parametric approach in the calculation of the causality relationship between scores and discussion frequencies.

Popularity sentiments

The first objective requires the mining of opinions since online platforms that allow users to express open-ended product reviews, ideas, and comments contain a huge amount of text. Some text may contain unknown words and abbreviations. This is where text mining analysis is be used to discover the patterns, connections, and trends [50]. Sentiment analysis [27, 27, 28, 61, 66] is used to evaluate opinions, sentiments, evaluations, appraisals, attitudes, and emotions. In this research, sentiment analysis or opinion mining is used. A preprocessing text technique is required for information retrieval, information extraction, and computational linguistics research that transforms unstructured, original-format content into structured, formatted information [27, 50, 61]. The technique used in this study is natural language preprocessing [67]. Natural language preprocessing is composed of different tasks, including tokenization, part-of-speech tagging, syntactical parsing, and shallow parsing [27, 50]. The orientation of an opinion (sentiment orientation, polarity of opinion, semantic orientation, and orientation score) is determined by dictionary methods. Those include the augmented dictionary method [68] for positive, negative, or neutral opinions and the Hu Liu sentiment dictionary for tagging polarized words in an opinion. The main advantages of a dictionary approach are to speed up processing of large datasets and

Fig. 1 Components of sentiment analysis

Fig. 2 Example of network with three community structures with high internal density but low external density between them

to increase processing accuracy [60]. The lookup dictionary method combined with the National Research Council Canada (NRC) sentiment dictionary [69] is used to calculate frequency of emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust) in an opinion.

Sentiment analysis is the computational study of opinions, sentiments, and emotions expressed in text. The formula to calculate the polarity of a sentence is:

$$\left(o_j, f_{jk}, oo_{ijkl}, h_i, t_l\right).$$

Object o_j is the entity target. In our case, it will be the Starbucks products. f_{jk} is a feature representing the components and attributes of the object o_j. In our case, it will be, for example, the marble cake. Sometimes, the object can be itself seen as a feature. The opinion holder is the person h_i (or organization) that expresses the opinion. The polarity measure oo_{ijkl} can be a positive, negative, or neutral value. The i, j, k, and l values are the indexes. The polarity measure in the example in Fig. 1 is based on sentence level (two sentences) and the dictionary we get of polarity measure $= 0$ (neutral)

Community detection

The second objective of this study requires the identification of structures in idea proposals. The structures of many phenomena can be represented as networks [49, 70–72]. A typical network has two components: the nodes (vertices, actors) and the collection of relationships between the nodes. Nodes can be regrouped together in clusters; community detection refers to regrouping similar nodes of the network together. Figure 2

illustrates the property of community detection, where the network is split into sets of nodes with high internal density but low external density. The inner density of nodes is defined as follows:

$$\delta_{\text{internal}} = \frac{\text{\# internal edges of} C_x}{n_{C_x}(n_{C_x} - 1)/2},$$

and the external density as:

$$\delta_{\text{external}} = \frac{\text{\# inter cluster edges of } C_x}{n_{C_x}(n_{C_x} - 1)/2},$$

where C_x is the community x and n_{C_x} is the number of nodes in the community x.

The identification of the optimal number of communities is an open problem. Table 4 presents the list of eight most common algorithms for community identification [73], the performance of which was compared in this study. However, every algorithm has its own limitations depending on the network topology. Thus, a quantitative criterion is needed to assess the quality of community structures. The modularity measure, as defined by Newman et al. [44], is the function used to assess the groups of nodes in the network that interact more with each other than with the rest of the network [37, 73]. The Newman and Girvan modularity measure can be written as:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j),$$

where for two vertices, i and j, the A_{ij} is the adjacency matrix that indicates if pairs of nodes are adjacent or not, k_i and k_j are the degrees of the vertices i and j, and $m = \sum_i k_i / 2$ is the total number of edges in the network. The δ-function is defined as:

$$\delta(C_i, C_j) = \begin{cases} 1, & \text{if } C_i = C_j \\ 0, & \text{Otherwise} \end{cases}$$

The possible values of the modularity measure are between 0 and 1. Values closest to 1 indicate better community-structure quality. The modularity increases if the size of the graph increases as well as if the number of well-separated communities increases [73]. The statistical significance of a community is calculated with the Wilcoxon rank-sum

Table 4 Popular community-detection algorithms

Algorithm
Multi-level (Louvain)
Fast Greedy
Info Map
Edge Betweenness
Leading Eigenvector
Walk Trap
Label Propagation
Spin Glass

test [74]. This statistical test allows verification that the internal degree of density of nodes in the community is higher than the external degree.

The multilevel (Louvain) algorithm for finding a community structure is based on the multilevel optimization of modularity measure and a hierarchical clustering approach [75]. The fast greedy algorithm is based on the locally optimal choice at each stage (greedy optimization) and a hierarchical approach [76]. This algorithm is fast and well adapted for networks with large numbers of vertices and edges. The info map algorithm is based on the probability flow of a random walk trajectory [77] and minimizing the map equation [78] over possible network partitions. The edge betweenness algorithm is based on the number of the shortest paths that go through an edge in a network and a hierarchical approach [44]. This algorithm is very slow and is not recommended for very large networks. The leading eigenvector approach is based on the optimization of the modularity measure combined with a divisive approach [79]. This algorithm is not appropriate for degenerated networks. The walk trap algorithm is based on short random walks with the same principle as the fast greedy algorithm, but it is slower than the latter [80]. The label propagation algorithm [81] is a fast algorithm based on the propagation of a small subset of a priori labeled data through unlabeled points in the network. The spin glass approach is derived from statistical mechanics and based on the Potts model [82, 83].

Generalized least squares

The third objective of the study established that positive sentiment between two different users increases with the number of discussions. This relationship is explained with the generalized least squares model, a linear model well adapted for variables that are not normally distributed. That is, the use of non-parametric methods. The linear model with the generalized least squares estimator instead of ordinary least squares is preferable even if the independent variables are numerical [61, 62]. The dependent variable is the total sentiment score in a discussion between two users. The independent variable is the number of discussions exchanged between two different users and is transformed with the logarithmic function.

Proposed approach

Given the large size of the collected data, 5% (randomly resampled 11 times) of the ideas and their comments with the highest and lowest scores were analyzed separately. Table 3 summarizes the frequencies of the 5% (randomly resampled 11 times) of agreeable and disagreeable ideas per user and the average likes or dislikes per idea.

The polarity sentiment scores are calculated with the augmented dictionary method. The Hu Liu lexicon was extracted for each idea and annotated to determine whether the content was positive, negative, or neutral. Subjective information about the emotions was extracted with the lookup dictionary method. The NRC lexicon was used to obtain the frequency of anger, anticipation, disgust, fear, joy, sadness, surprise, and trust emotions.

Figures 3 and 4 show the trajectory plots to understand and visualize how these sentiments and emotions were activated across the ideas and comments, and how the narrative is structured over time. The x-axis is time, the y-axis represents the
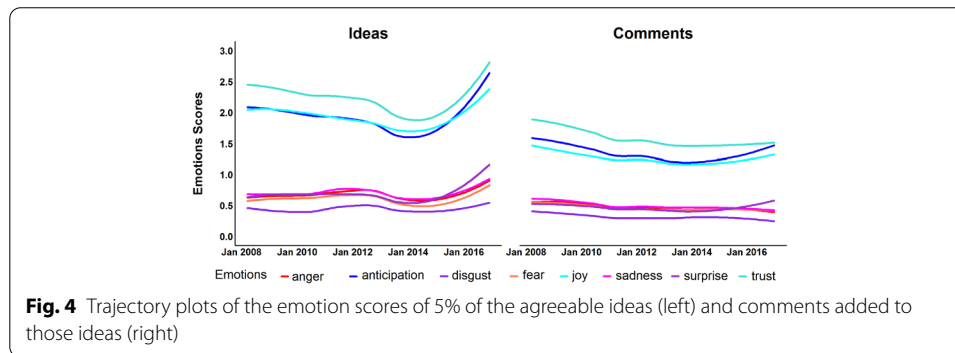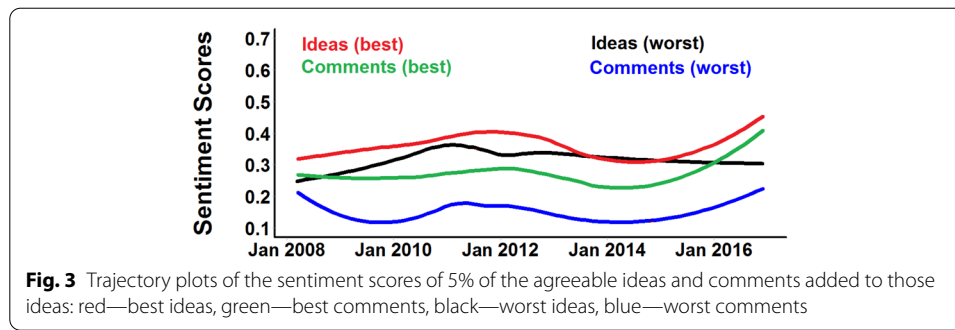
Fig. 3 Trajectory plots of the sentiment scores of 5% of the agreeable ideas and comments added to those ideas: red—best ideas, green—best comments, black—worst ideas, blue—worst comments

Fig. 4 Trajectory plots of the emotion scores of 5% of the agreeable ideas (left) and comments added to those ideas (right)

Table 5 Attributes with their description for social network analysis

Attribute	Description
from	Usernames who posted comments
to	Usernames who posted ideas
frequency	Number of times two users discussed
sentiment	Sentiment scores (total) in discussions between to users

sentiment scores or emotions frequency. A Fourier transformation and low-pass filtering were used to reveal simple shapes and remove the noise from the trajectory plots.

For the social network analysis, the frequencies of discussions exchanged between two users were calculated with the constraint of pairing users with more than two discussions. This calculation was made separately for agreeable and disagreeable ideas. The resulting datasets contained only four attributes, as described in Table 5. For the sentiment analysis, the package igraph in R was used to build the undirected network graphs and to identify the community structures. The algorithms employed are listed in Table 3. Multiple edges were combined and isolated nodes (i.e., nodes with centrality degree equal to zero) were removed. The nodes were aligned in an esthetic manner using Fruchterman–Reingold's layout based on force-directed graph drawing algorithms [84]. The adequate community structure was selected using the modularity

measure and the Wilcoxon rank-sum test [74]. For the best modularity measure, two network graphs were created: one with sentiment scores as the edge weights and one without edge weights.
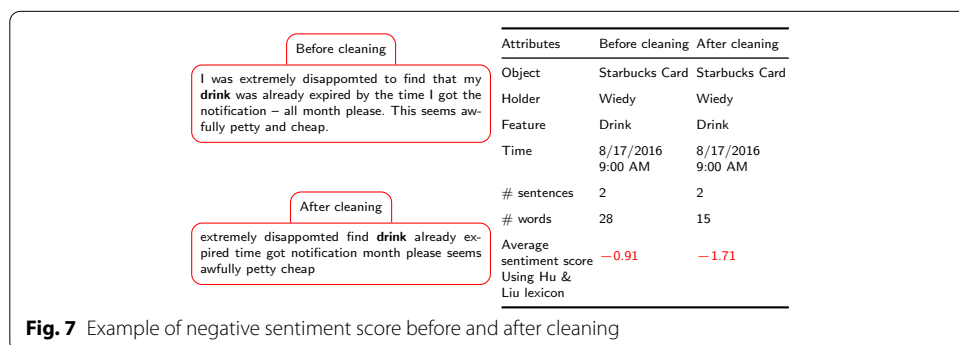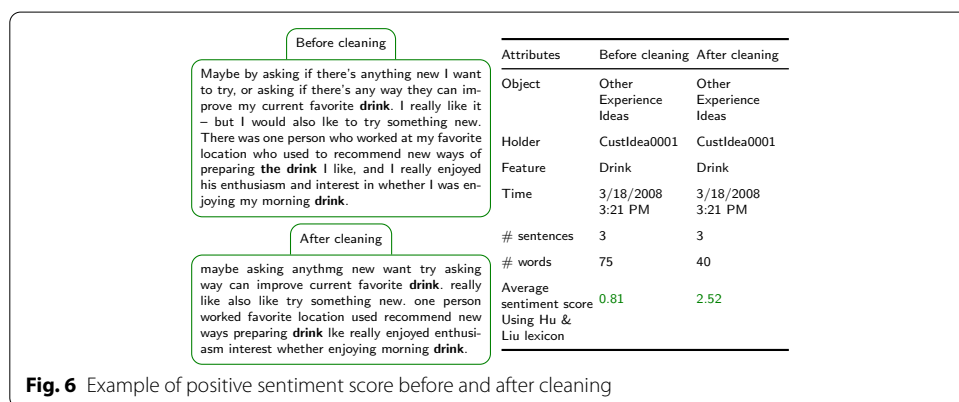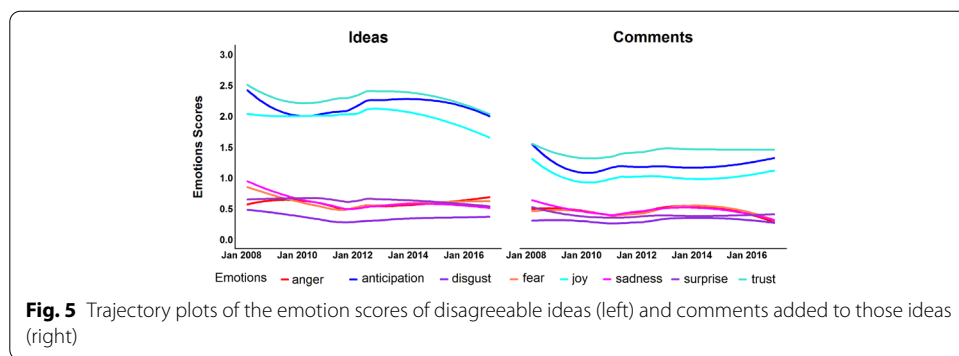
To investigate the relationship between sentiment scores and discussion frequency, the generalized least squares technique was complemented with the bootstrap technique. A set of 22 simulations were constructed with 5% of the sample (randomly resampled 11 times). Eleven simulations for the best ideas and 11 simulations for the worst ideas were conducted. The records selected were randomly chosen. For each one of the simulations, the parameters of the linear regression were calculated. The final parameters of the linear regression were calculated as the average value of each coefficient. The standard error and statistical significance were calculated from the 11 simulations.

Results

Sentiment analysis results

Sentiment analysis provides the insights for the first objective, which is to find out if crowdsourcing users have an accurate impression of the company. Having an understanding and knowledge of the emotions of users when they post ideas and comments, the company can better assess the benefit of crowdsourcing for its operations.

The trajectory plots in Fig. 3 show that the narratives of the agreeable ideas and disagreeable ideas with their comments are structured. Comments on the best ideas (green) follow the shape of the best ideas (red): the sentiment of the best ideas is correlated with that of their comments. Meanwhile, comments on the worst ideas (blue) do not follow the shape of the worst ideas (black): the worst ideas' sentiment is not correlated with their comments' sentiment. The sentiment scores of the best ideas are always above those of the worst ideas (except for the year 2014). This signals a positive involvement of users with the company. The users feel comfortable with the company. There is a parallelism of best comments with best ideas that reinforces the good perception of the company. The comments on the worst ideas present a curvature opposite of those ideas. Users are not in the mood to attack or criticize the worst ideas. They give a low score to the worst ideas, yet their sentiment scores do not signal a negative attitude. The curves corresponding to best ideas and best comments never cross paths. The situation is similar for worst ideas and worst comments. Ideas and comments are the main effects and have no interactions between their scores. This means that comments can explain both the best and worst ideas. There was a positive sentiment all along the time period between 2008 and 2016. Both the agreeable and disagreeable ideas began with the lowest sentiment scores and slightly moved up until they reached their maximum between 2010 and 2011. The positive sentiment of the language dropped down slightly until it reached a local minimum for the agreeable ideas in 2014 to 2015 and for the disagreeable ideas in 2012. After 2015, the positive language used in the agreeable ideas increased until it reached the maximum in 2016. The shape of the agreeable ideas was sinusoidal, and that of the disagreeable ideas was a parabola. The shapes of sentiment of the language used in the comments were approximately linear without trend and with some negligible movements, except in 2015 when the positive language used in the comments added to the agreeable ideas increased significantly until it reached its maximum in 2016.

Fig. 5 Trajectory plots of the emotion scores of disagreeable ideas (left) and comments added to those ideas (right)

Fig. 6 Example of positive sentiment score before and after cleaning

Fig. 7 Example of negative sentiment score before and after cleaning

Figures 4 and 5 show that the trust, joy, and anticipation emotions dominate the contents of the language used in both the agreeable and disagreeable ideas and their comments. However, there were no significant movements in the trajectory plots of the emotions, except in 2015. The trust, joy, and anticipation emotions increased significantly in the agreeable ideas after 2015 until they reached their maximum in 2016. Posted ideas and comments indicated that users were mainly in a good mood. The scores of ideas posted in a good mood are always above those of the comments. Here are two examples of positive and negative sentiment scores (Figs. 6 and 7). In both cases, the score went up after the cleaning the text. The score was calculated with the Hu and Liu lexicon.

For the users who are in a bad mood, ideas and comments are in the same range for both the best and worst ideas or comments. The curves of positive and negative never cross paths: there are no interactions between good and bad moods. Within good moods and bad moods there are interactions throughout the study. In particular, trust and joy never interact, but anticipation interacts with both joy and trust. Users feel trust or joy towards the company, but anticipation is always present. The users have a positive perception of the company and they provide the company with good input. The number of users with bad moods and negative emotions is smaller than that of those with positive emotions.

Network analysis results

The network analysis gives an analytic graphical solution to the second objective. Among all the participants, it is possible to regroup users into communities that are more likely to generate the best ideas. At the same time, the links between and within communities can be found. There are three sets of results in the network analysis used to identify the communities of users with the best ideas. The identification of communities needs to be optimal (Tables 6 and 7 for modularity). For the best set of communities, the community structure is calculated and presented in Tables 8 and 9. The graphical interactive representation of the network is presented in Figs. 8 and 9.

Table 6 shows the number of communities, the modularity measures, and the number of significant communities detected by each algorithm for the agreeable ideas. The fast greedy and multilevel algorithms have the highest values of the modularity measure with and without weights, but when the sentiment scores are not used as edge weights, the

Table 6 (Agreeable ideas) Modularity measures, detected communities, and significant communities at level 0.05 (based on the Wilcoxon rank-sum test)

Community detection algorithm	Modularity	Number of communities	Significant communities
With weight			
Multi-level	0.54	9	6
Fast Greedy	0.53	9	8
Info Map	0.50	34	6
Edge Betweenness	0.48	12	9
Leading Eigenvector	0.48	13	9
Walk Trap	0.47	31	6
Label Propagation	0.42	30	2
Spin Glass	0.21	12	9
Without weight			
Multi-level	0.54	9	6
Fast Greedy	0.53	11	9
Info Map	0.50	33	7
Edge Betweenness	0.53	9	7
Leading Eigenvector	0.51	10	7
Walk Trap	0.47	35	7
Label Propagation	0.45	35	4
Spin Glass	0.21	11	7

Each community structure is detected by one of the eight algorithms, with and without sentiment scores as weight

Table 7 (Disagreeable ideas) Modularity measures, detected communities, and significant communities at level 0.05 (based on the Wilcoxon rank-sum test)

Community detection algorithm	Modularity	Number of communities	Significant communities
With weight			
Fast Greedy	0.44	11	10
Multi-level	0.44	9	9
Edge Betweenness	0.41	15	9
Leading Eigenvector	0.41	15	9
Info Map	0.41	34	8
Walk Trap	0.39	23	7
Spin Glass	0.25	15	8
Label Propagation	0.14	9	4
Without weight			
Fast Greedy	0.47	8	8
Multi-level	0.44	9	9
Edge Betweenness	0.45	20	8
Leading Eigenvector	0.43	20	8
Info Map	0.42	35	7
Walk Trap	0.38	32	6
Spin Glass	0.25	11	8
Label Propagation	0.05	9	4

Each community structure is detected by one of the eight algorithms, with and without sentiment scores as weight

edge betweenness algorithm has the same modularity measure as the fast greedy algorithm. The Wilcoxon rank-sum tests show that the community structure based on the fast greedy algorithm has the highest number of significant communities. The top communities are significantly dense internally and significantly sparse externally, and one community is not significant by this measure (the fourth community in Table 8).

Figures 8 and 9 show interactions among users in the community. The groups depicted in these figures can be detected and isolated. The graphs present the distance between communities, the spread of communities, and the location of communities. Each node has the interaction between communities and within groups. In the case of agreeable ideas, Fig. 8 shows the undirected network in which the nodes represent the users who posted ideas and/or comments, and there is an edge from user A to user B if B frequently had discussions with A. The widths of the edges indicate the frequency of discussions between users A and B, and the colors of the nodes indicate the communities found by the fast greedy algorithm with sentiment scores as edge weights. The network has a total of 254 users and 545 edges, where the positive, negative, and neutral sentiments are colored in blue, red, and grey, respectively. Most edges (88.4%) are blue, which means that the language used in discussions between users is positive. Table 8 shows that the 4th and 8th communities contain users who posted the most valuable ideas, with 7724.6 and 6890 likes per idea on average, respectively. The users of these communities are not active on the My Starbucks Idea platform, with an average of 62.3 and 113 points earned per user on average, 4.4 and 3.1 connections on average, and only 40 and 18 posted ideas,

Table 8 (Agreeable ideas) User characteristics of each community based on the fast greedy algorithm with sentiment scores as edge weight

Communities		No. users posted ideas	No. users not posted ideas	No. of ideas	Average number of			
					Ideas likes	Points earned by user	Member since (days)	Relations by users
	1	29	4	67	2862.4	260.8	1678.3	4.6
	2	35	2	189	2043.4	826.7	2263.5	5.6
	3	42	2	215	1474.8	3726.8	1285.1	4.9
	4	13	1	40	7724.6	62.3	1927.7	4.4
	5	53	2	387	3412.5	2589.9	2636.2	7.1
	6	19	0	50	4551.6	156.4	2219.8	3.8
	7	34	2	481	3881.8	11882.9	2086.3	7.0
	8	6	2	18	6980.0	113.0	3186.5	3.1
	9	6	2	13	3831.7	246.8	2609.0	2.8

respectively. The 7th community contained the most active users with 481 posted ideas (13.4 ideas per user on average), 11,882.9 points earned per user on average, and 7.0 connections per user on average. The 1st and 3rd communities contain new members with an average of 1678.3 days and 1285.1 days per user, respectively, but the users of the 1st community are less active than the 3rd, with two ideas per user on average, 260.8 points earned per user on average, and 4.6 connections per user on average. The 9th community has the lowest number of posted ideas and connections per user on average. The 2nd and 6th contain occasional users, and the 5th community contains the second most active users.

In the case of the disagreeable ideas, Table 7 shows that the fast greedy algorithm without sentiment scores as edge weights has the highest value of modularity measure, and that all communities are significantly dense internally and significantly sparse externally.

In the case of the disagreeable ideas, Fig. 9 shows that the language used in discussions between users is positive (81.5%). The graph contains 274 users and 647 connections. In Table 9, the 2nd and 5th communities had the least valuable ideas with dislikes per idea, on average, of 3612.7, and 3416.4, respectively, but the users of both have different characteristics; the users of the 2nd community have the highest number of connections per user and the oldest accounts with 2763 days per user on average. Meanwhile, the users of the 5th community have higher points earned per user on average (8415.8 points). The 3rd community contains the users who have the highest points earned per user on average (34,732.6 points) and the highest number of posted ideas per user on average (8.1 ideas). The 7th and 8th communities contain the newest users with less than one idea per user on average, with the users of the 7th community being the oldest of the two (1751 days on average) and having more

Table 9 (Disagreeable ideas) User characteristics of each community based on the fast greedy algorithm without sentiment scores as edge weight

| Communities | | No. users posted ideas | No. users not posted ideas | No. of ideas | Average number of | | | |
					Ideas likes	Points earned by user	Member since (days)	Relations by users
👥	1	10	21	69	1001.0	389.8	2337.8	4.2
👥	2	30	41	407	3612.7	5028.6	2763.0	7.8
👥	3	14	34	390	1970.7	34732.6	2259.3	5.4
👥	4	8	23	23	1190.0	126.8	2142.5	4.3
👥	5	14	24	348	3416.4	8415.8	2409.1	5.9
👥	6	9	34	46	1392.2	393.7	2589.4	6.0
👥	7	1	7	1	640.0	1.0	1751.0	3.0
👥	8	1	3	1	200.0	2.0	776.0	2.8

dislikes per idea on average (640). The users from the 1st, 4th, and 6th communities are occasional users with different characteristics.
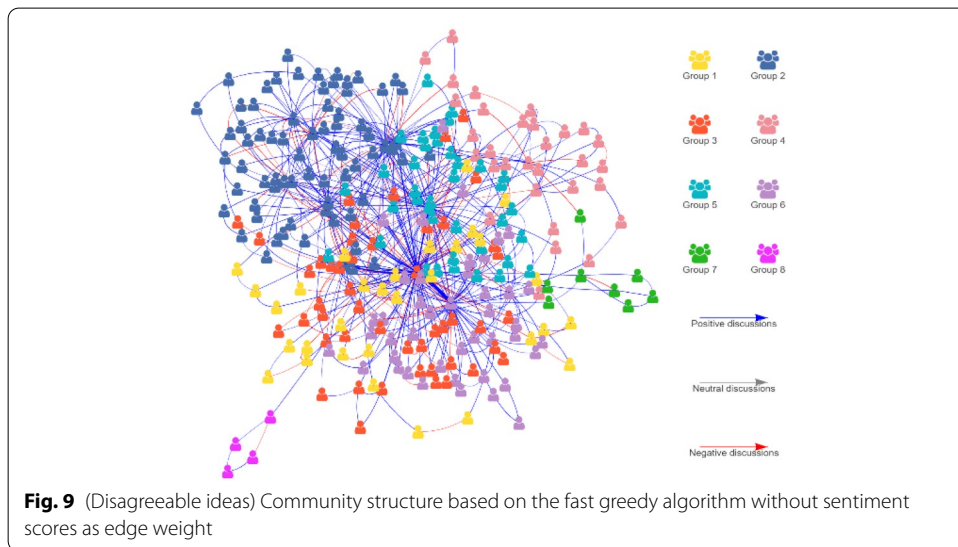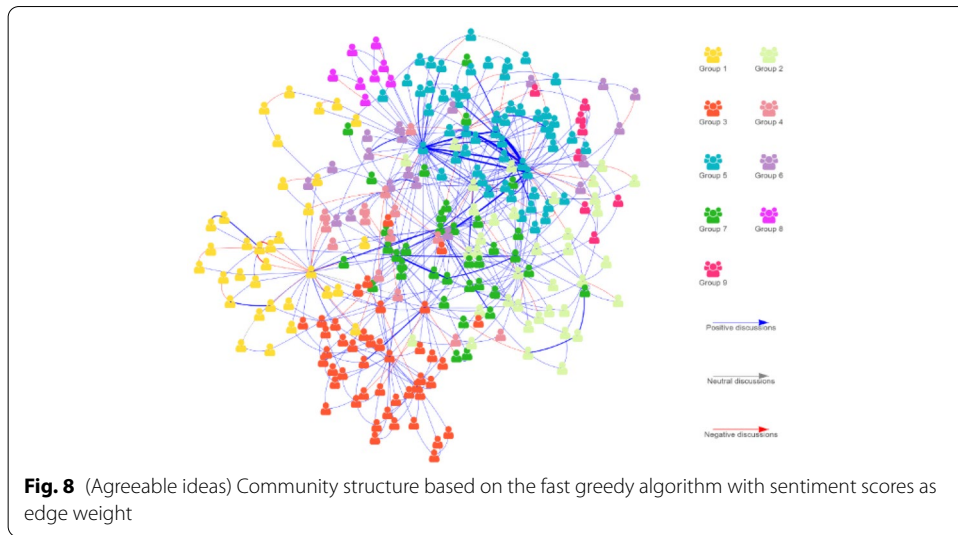
Table 8 should be read together with Fig. 8. A stakeholder sees that community 5 (blue) generates 387 ideas and community 7 (green) generates 481 ideas. These are the most prominent communities with the best ideas. Figure 8 shows the stakeholder that those two communities interact with each other and with the other communities. For the stakeholder, it becomes clear that the users inside community 5 and 7 are the most valuable ones. Focusing attention on the ideas proposed by communities 5 and 7 minimizes analysis time and improves operational efficiency for the company.

In the same way, Table 9 shows that communities 2, 3, and 5 produce the largest number of the worst ideas. The company can save time by bypassing all those ideas and the users who are more likely to continue generating bad ideas.

Linear models

A linear model was constructed to satisfy objective three, which is to determine the relationship between ideas'sentiment scores and the discussion frequency of crowdsourcing users. Tables 10 and 11 show the coefficients with standard errors in parentheses for each simulation. The final column contains the coefficients and standard errors of all the simulations. The standard error of the final result is greater than any of each simulation. However, the coefficients remain statistically significant. The standard error also remains in the same order of magnitude. Table 10 presents the coefficients and their standard errors for agreeable ideas. Table 11 shows the coefficients and standard errors for disagreeable ideas.

Figures 10 and 11 show a quadratic relation between the sentiment of the language used by users and the number of discussions exchanged in the case of agreeable/disagreeable ideas. This association is statistically significant, as seen in Table 12. This result confirms the third hypothesis. The generalized least squares equations can be written as:

Fig. 8 (Agreeable ideas) Community structure based on the fast greedy algorithm with sentiment scores as edge weight

Fig. 9 (Disagreeable ideas) Community structure based on the fast greedy algorithm without sentiment scores as edge weight

$$\text{Agreeable} = 3.05 - 3.54 \log(\text{frequency})$$
$$+ 1.29 \log(\text{frequency})^2$$
$$\text{Disagreeable} = 2.76 - 2.98 \log(\text{frequency})$$
$$+ 0.98 \log(\text{frequency})^2$$

For agreeable ideas, with a frequency, the score will be $3.05 - 3.54 \cdot 3 + 1.29 \cdot 9 = 4.04$. That is, the score of the idea will increase by 4.04 points. For disagreeable ideas, $2.76 - 2.98 \cdot 3 + 0.98 \cdot 9 = 2.64$. The perception of the disagreeable idea will increase by 2.64 points, which remains below the value of agreeable ideas. Given that this is a quadratic relationship, the higher the frequency, the bigger the impact on the ideas. For agreeable ideas, the relationship has a positive impact on the score. For disagreeable ideas, there is a positive impact, smaller than agreeable ideas, that reinforces a bad perception and a bad evaluation of the score.

Table 10 Generalized least squares model for the bootstrap method with 11 simulations of 11 randomly selected samples for agreeable ideas

	Agreeable ideas					
	1	2	3	4	5	6
Intercept	4.557	3.310	4.319	2.043	3.991	3.949
	(0.767)	(0.796)	(0.790)	(0.804)	(0.786)	(0.814)
log(frequency)	− 5.145	− 3.89	− 5.101	−2.574	− 4.55	− 4.401
	(0.789)	(0.814)	(0.802)	(0.821)	(0.794)	(0.828)
log (frequency)^2	1.696	1.401	1.752	1.081	1.544	1.498
	(0.194)	(0.199)	(0.195)	(0.200)	(0.192)	(0.201)
	Agreeable ideas					Final
	7	8	9	10	11	
Intercept	3.873	2.846	1.928	0.730	1.991	3.049
	(0.771)	(0.801)	(0.823)	(0.788)	(0.810)	(1.230)
log(frequency)	− 4.225	− 3.261	− 2.320	− 1.080	− 2.346	− 3.536
	(0.779)	(0.814)	(0.841)	(0.799)	(0.821)	(1.317)
log (frequency)^2	1.406	1.205	0.970	0.676	0.974	1.291
	(0.187)	(0.198)	(0.204)	(0.193)	(0.199)	(0.338)

Standard errors in parentheses

Table 11 Generalized least squares model for the bootstrap method with 11 simulations of 11 randomly selected samples for disagreeable ideas

	Disagreeable ideas					
	1	2	3	4	5	6
Intercept	3.921	1.742	2.182	1.825	1.970	4.413
	(0.473)	(0.655)	(0.581)	(0.617)	(0.561)	(0.602)
log(frequency)	− 4.174	− 1.875	− 2.360	− 2.084	− 2.208	− 4.554
	(0.413)	(0.57)	(0.501)	(0.529)	(0.489)	(0.524)
log (frequency)^2	1.270	0.692	0.842	0.804	0.803	1.319
	(0.083)	(0.11)	(0.096)	(0.102)	(0.095)	(0.101)
	Disagreeable ideas					Final
	7	8	9	10	11	
Intercept	2.423	2.860	2.438	2.983	3.614	2.761
	(0.611)	(0.583)	(0.629)	(0.633)	(0.568)	(0.892)
log(frequency)	− 2.778	− 3.024	− 2.777	− 3.276	− 3.700	− 2.983
	(0.531)	(0.509)	(0.544)	(0.55)	(0.492)	(0.872)
log (frequency)^2	0.976	0.961	0.963	1.050	1.114	0.981
	(0.103)	(0.098)	(0.105)	(0.106)	(0.095)	(0.197)

Standard errors in parentheses

Fig. 10 Agreeable ideas

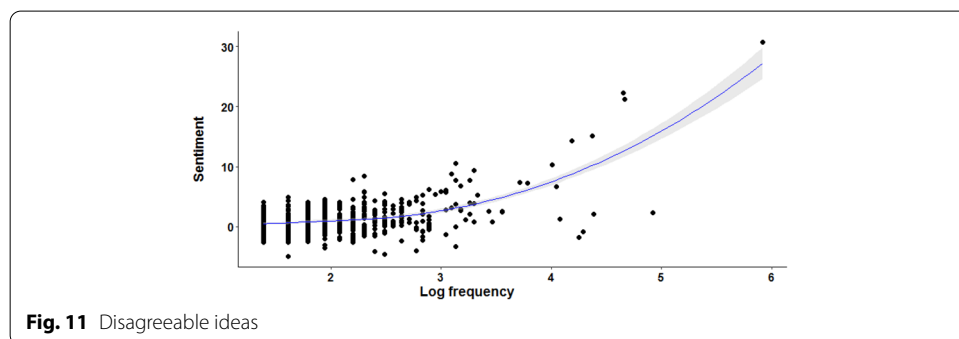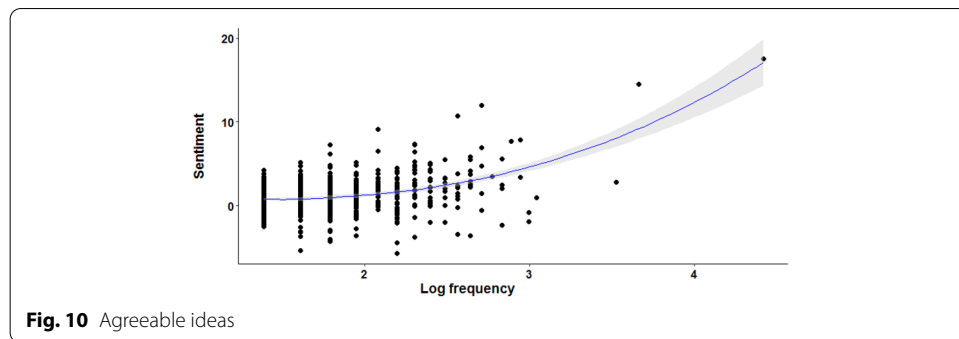**Fig. 11** Disagreeable ideas

Table 12 Generalized least squares estimated model with sentiment scores as outcome for agreeable and disagreeable ideas

	Agreeable ideas	Disagreeable ideas
Intercept	3.049***	2.761***
	(1.230)	(0.892)
log(frequency)	− 3.536***	− 2.983*
	(1.317)	(0.872)
log (frequency)^2	1.291**	0.981**
	(0.338)	(0.197)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Standard errors in parentheses

Conclusion

Firms recognize the intrinsic value in the sentiments and emotions of users in response to prospective ideas, and their comments can be valuable for innovation at the initial ideation stage. Firms' increased use of crowdsourcing is a clear indication that they recognize the value of cooperation. While the use of crowdsourcing continues to grow, there is an evident paucity in the development of appropriate empirical methods to analyze that data effectively. The research and results presented here contribute to the advancement of both crowdsourcing scholarship and empirical methods used to examine that phenomenon by illustrating that a plurality of methods yields a greater depth of understanding of the phenomenon.

While some scholars chose one empirical methodology paradigm to analyze data from crowdsourcing, we decided to employ a few different methods to achieve a more

in-depth insight and benefit from a multitude of perspectives. Each technique used in this study added another meaningful dimension to the study of crowdsourcing. Sentiment analysis, a psychology-based analysis tool, yielded rich data that firms can use to commercialize users' ideas through a systematic analysis of language. Social network analysis provided valuable insight into the dynamics among and between users that can aid firms in determining appropriate parameters for both the crowdsourcing platform and the subsequent analysis of data. Finally, the ordinary least squares method provided for a robust statistical analysis that further solidified the validity of the results of the study.

Social network analysis allows the interaction of users to be depicted in an interactive graph with users clustered into communities. The graphs provide a visual tool that aids researchers to aptly identify influential communities that are at the center of discussions on the platform. Communities that are isolated or have fewer interactions can be filtered to target only the group of users who suggest valuable ideas. This finding can help firms implement suggestions for idea innovations and service improvements more efficiently, adding value to new product generation and innovation. Finally, we provide empirical evidence that indicates that positive language in a discussion between two individuals increases with the frequency of the conversations exchanged. This finding suggests that, over time, more frequent interactions can influence individuals' evaluation of ideas within the community.

Like all empirical research studies, there were some limitations. When evaluating idea polarity, only one layer of the wheel of emotions from [41] was used to simplify word polarization. Although this approach is commonly used to achieve parsimonious and generalizable results, additional layers may have resulted in more specific results. Also, the database used in this study is vast, so we had to employ bootstrapping to obtain accurate samples to avoid the use of supercomputers. Some scholars criticize bootstrapping as a time-consuming process that makes assumptions, such as independent samples, that could skew results. However, this commonly used approach is necessary when the use of supercomputers is prohibitive.

Authors' contributions
All authors contributed to this work. All authors read and approved the final manuscript.

Availability of data and materials
The data and materials are available upon requests.

Declarations

Competing interests
There are no competing interests in this manuscript.

Author details
[1]Department of Computer & Information Sciences, Towson University, Towson, MD 21252, USA. [2]Department of Marketing, Towson University, Towson, MD 21252, USA.

References

1. Cooper, R.G., Kleinschmidt, E.J.: Success factors in product innovation. Ind. Market. Manag. **16**(3), 215–223 (1987)
2. Narver, J.C., Slater, S.F., MacLachlan, D.L.: Responsive and proactive market orientation and new-product success. J. Prod. Innov. Manag. **21**(5), 334–347 (2004)
3. Sigala, M.: Social networks and customer involvement in new service development (nsd) the case of www.mystar-bucksidea.com. Int. J. Contemp. Hosp. Manag. **24**(7), 966–990 (2012)
4. Duverger, P.: Using dissatisfied customers as a source for innovative service ideas. J. Hosp. Tour. Res. **36**(4), 537–563 (2012)
5. Lee, H., Han, J., Suh, Y.: Gift or threat? an examination of voice of the customer: the case of mystarbucksidea.com. Electr. Commerce. Res. Appl. **13**(3), 205–219 (2014)
6. Prahalad, C.K., Ramaswamy, V.: The future of competition: co-creating unique value with customers. Harvard Business Press, Boston (2004)
7. Huang, Y., Singh, P.V., Srinivasan, K.: Crowdsourcing "blockbuster" ideas: a dynamic structural model of ideation. In: ICIS (2011)
8. Brew, A., Greene, D., Cunningham, P.: Using crowdsourcing and active learning to track sentiment in online media. In: 19th European Conference on Artificial Intelligence, pp. 145–150 (2010)
9. Borgo, R., Micallef, L., Bach, B., McGee, F., Lee, B.: Information visualization evaluation using crowdsourcing. In: Computer Graphics Forum, vol. 37, pp. 573–595 (2018). Wiley Online Library
10. Rambocas, M., Gama, J., et al.: Marketing research: The role of sentiment analysis. Fep working papers 489, Universidade do Porto, Faculdade de Economia do Porto (2013)
11. Razzaq, A., Asim, M., Ali, Z., Qadri, S., Mumtaz, I., Khan, D.M., Niaz, Q.: Text sentiment analysis using frequency-based vigorous features. China Commun. **16**(12), 145–153 (2019)
12. Burton, M.D., Nicholas, T.: Prizes, patents and the search for longitude. Explor. Econ. Hist. **64**, 21–36 (2017)
13. Vukovic, M.: Crowdsourcing for enterprises. In: IEEE Congress on Services, pp. 686–692 (2009). IEEE
14. Davis, J.G.: From crowdsourcing to crowdservicing. IEEE Internet Comput. **15**(3), 92–94 (2011)
15. Dimitrova, S.G.: Implementation of crowdsourcing into business and innovation strategies: A case study at bombardier transportation, Germany. PhD thesis, École Polytechnique de Montréal (2013). Doctoral dissertation
16. Howe, J.: The rise of crowdsourcing. Wired Mag. **14**(6), 1–4 (2006)
17. Howe, J.: Crowdsourcing: how the power of the crowd is driving the future of business. Random House, New York (2008)
18. Estellés-Arolas, E., González-Ladrón-de-Guevara, F.: Towards an integrated crowdsourcing definition. J. Inf. Sci. **38**(2), 189–200 (2012)
19. Hosseini, M., Phalp, K., Taylor, J., Ali, R.: On the configuration of crowdsourcing projects. Int. J. Inf. Syst. Model. Des. **6**(3), 27–45 (2015)
20. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R.C., Schader, M.: Managing the crowd: Towards a taxonomy of crowdsourcing processes. In: AMCIS (2011)
21. Saxton, G.D., Oh, O., Kishore, R.: Rules of crowdsourcing: models, issues, and systems of control. Inf. Syst. Manag. **30**(1), 2–20 (2013)
22. Rouse, A.C.: A preliminary taxonomy of crowdsourcing. ACIS 2010 Proceedings **76** (2010)
23. Doan, A., Ramakrishnan, R., Halevy, A.Y.: Crowdsourcing systems on the world-wide web. Commun. ACM. **54**(4), 86–96 (2011)
24. Frei, B.: Paid crowdsourcing: Current state & progress toward mainstream business use. Produced by smartsheet.com (2009)
25. Schenk, E., Guittard, C., et al.: Crowdsourcing: What can be outsourced to the crowd, and why? In: Workshop on Open Source Innovation, Strasbourg, France, vol. 72, p. 3 (2009). Citeseer
26. Lee, H., Jeong, S., Suh, Y.: The influence of negative emotions in an online brand community on customer innovation activities. In: System Sciences (HICSS), 2014 47th Hawaii International Conference on System Sciences, pp. 1854–1863 (2014). IEEE
27. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**(1), 1–167 (2012)
28. Bakshi, R.K., Kaur, N., Kaur, R., Kaur, G.: Opinion mining and sentiment analysis. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 452–455 (2016). IEEE
29. Ceccato, V.A., Snickars, F.: Adapting gis technology to the needs of local planning. Environ. Plan. B. **27**(6), 923–937 (2000)
30. Piryani, R., Madhavi, D., Singh, V.K.: Analytical mapping of opinion mining and sentiment analysis research during 2000–2015. Inf. Proces. Manag. **53**(1), 122–150 (2017)
31. Nasukawa, T., Yi, J.: Sentiment analysis: Capturing favorability using natural language processing. In: Proceedings of the 2nd International Conference on Knowledge Capture. K-CAP'03, pp. 70–77. Association for Computing Machinery, New York, NY, USA (2003)
32. Dave, K., Lawrence, S., Pennock, D.M.: Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In: Proceedings of the 12th International Conference on World Wide Web, pp. 519–528 (2003)
33. Mishne, G., Glance, N.S., et al.: Predicting movie sales from blogger sentiment. In: AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, pp. 155–158 (2006)
34. Sadikov, E., Parameswaran, A., Venetis, P.: Blogs as predictors of movie success. In: Proceedings of the Third International Conference on Weblogs and Social Media (ICWSM-2009) (2009)
35. Liu, Y., Huang, X., An, A., Yu, X.: Arsa: a sentiment-aware model for predicting sales performance using blogs. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 607–614 (2007)
36. Fournier, S., Lee, L.: Getting brand communities right. Harv. Bus. Rev. **87**(4), 105–111 (2009)

37. Lee, H., Seo, S.: What determines an agreeable and adoptable idea? a study of user ideas on mystarbucksidea.com. In: System Sciences (HICSS), 2013 46th Hawaii International Conference on System Sciences, pp. 3207–3217 (2013). IEEE

38. Pestian, J., Nasrallah, H., Matykiewicz, P., Bennett, A., Leenaars, A.: Suicide note classification using natural language processing: a content analysis. Biomed. Inf. Insights. **2010**(3), 19–28 (2010)

39. Zhang, X., Fuehres, H., Gloor, P.A.: Predicting stock market indicators through twitter – "i hope it is not as bad as i fear", pp. 1–8 (2010)

40. Chaudhuri, A.: Emotion and reason in consumer behavior. Routledge, England (2006)

41. Plutchik, R., Kellerman, H.: Theories of emotion. Academic Press, New York (1980)

42. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends. Inf. Retr. **2**(1–2), 1–135 (2008)

43. Scott, J.: Social network analysis. Sociology. **22**(1), 109–127 (1988)

44. Newman, M.E., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E. **69**(2), 026113 (2004)

45. Faggiani, A., Gregori, E., Lenzini, L., Luconi, V., Vecchio, A.: Smartphone-based crowdsourcing for network monitoring: opportunities, challenges, and a case study. IEEE Commun. Mag. **52**(1), 106–113 (2014)

46. Sha, Z., Chaudhari, A.M., Panchal, J.H.: Modeling participation behaviors in design crowdsourcing using a bipartite network-based approach. J. Comput. Inf. Sc. Eng.**19**(3),(2019)

47. Martínez-Torres, M.R.: Analysis of open innovation communities from the perspective of social network analysis. Technol Anal Strateg Manag **26**(4), 435–451 (2014)

48. Arenas-Marquez, F.J., Martínez-Torres, M.R., Toral, S.: Electronic word-of-mouth communities from the perspective of social network analysis. Technol. Anal. Strateg. Manag. **26**(8), 927–942 (2014)

49. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Modern. Phys. **74**(1), 47–97 (2002)

50. Toral, S..L., Martínez-Torres, M..d.R., Barrero, F.: Analysis of virtual communities supporting oss projects using social network analysis. Inf. Softw. Technol. **52**(3), 296–303 (2010)

51. Basole, R.C.: Visualization of interfirm relations in a converging mobile ecosystem. J. Inf. Technol. **24**(2), 144–159 (2009)

52. Mehler, A.: In search of a bridge between network analysis in computational linguistics and computational biology-a conceptual note. In: International Conference on Bioinformatics & Computational Biology, pp. 496–502 (2006). Citeseer

53. Allen, F., Babus, A.: Networks in finance. In: Kleindorfer, P.R., Wind, Y.J.R., Gunther, R.E. (eds.) The Network Challenge: Strategy, Profit, and Risk in an Interlinked World, pp. 367–382. Prentice Hall Professional, Hoboken, New Jersey, USA (2009)

54. Fazeli, A., Jadbabaie, A.: Game theoretic analysis of a strategic model of competitive contagion and product adoption in social networks. In: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), pp. 74–79 (2012). IEEE

55. Hodge, V.J., O'Keefe, S., Weeks, M., Moulds, A.: Wireless sensor networks for condition monitoring in the railway industry: a survey. IEEE Trans. Intell. Transp. Syst. **16**(3), 1088–1106 (2014)

56. Bassett, D.S., Sporns, O.: Network neuroscience. Nat. Neurosci. **20**(3), 353–364 (2017)

57. Sokhey, A.E., McClurg, S.D.: Social networks and correct voting. J. Politics. **74**(3), 751–764 (2012)

58. Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V.: Episimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks. In: SC'08: Proceedings of the 2008 ACM/IEEE Conference on Supercomputing, pp. 1–12 (2008). IEEE

59. Kolaczyk, E.D., Csárdi, G.: Statistical analysis of network data with R, vol. 65. Springer, New York (2014)

60. Hu, M., Liu, B.: Mining opinion features in customer reviews. In: Proceedings of the 19th National Conference on Artifical Intelligence. AAAI'04, pp. 755–760. AAAI Press, San Jose, California, USA (2004)

61. Williams, M.N., Grajales, C.A.G., Kurkiewicz, D.: Assumptions of multiple regression: correcting two misconceptions. Pract. Assess. Res. Eval. **18**(1), 11 (2013)

62. Kariya, T., Kurata, H.: Generalized least squares. Wiley, Hoboken (2004)

63. Efron, B.: Bootstrap methods: another look at the jackknife. Ann. Stat. **7**(1), 1–26 (1979)

64. James, G., Witten, D., Hastie, T., Tibshirani, R.: An introduction to statistical learning: with applications in R. Springer, New York (2017)

65. Abney, S.: Bootstrapping. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp. 360–367 (2002)

66. Maas, A., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C.: Learning word vectors for sentiment analysis. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 142–150 (2011)

67. Jusoh, S., Al-Fawareh, H.M.: Natural language interface for online sales systems. In: 2007 International Conference on Intelligent and Advanced Systems, pp. 224–228 (2007). IEEE

68. Ding, X., Liu, B., Yu, P.S.: A holistic lexicon-based approach to opinion mining. In: Proceedings of the 2008 International Conference on Web Search and Data Mining, pp. 231–240 (2008)

69. Mohammad, S.M., Turney, P.D.: Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In: Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, June 2010, LA, California, pp. 26–34 (2010). Association for Computational Linguistics

70. Dorogovtsev, S.N., Mendes, J.F.: Evolution of networks: from biological nets to the internet and WWW. Oxford University Press, Oxford (2013)

71. Newman, M.E.: The structure and function of complex networks. SIAM Rev. **45**(2), 167–256 (2003)

72. Watts, D.J.: Networks, dynamics, and the small-world phenomenon. Am. J. Sociol. **105**(2), 493–527 (1999)

73. Fortunato, S.: Community detection in graphs. Phys. Rep. **486**(3–5), 75–174 (2010)

74. Hollander, M., Wolfe, D.A., Chicken, E.: Nonparametric statistical methods. Wiley, Hoboken (2014)

75. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. **2008**(10), 10008 (2008)

76. Clauset, A., Newman, M.E., Moore, C.: Finding community structure in very large networks. Phys. Rev. E. **70**(6), 066111 (2004)
77. Good, B.H., De Montjoye, Y.-A., Clauset, A.: Performance of modularity maximization in practical contexts. Phys. Rev. E. **81**(4), 046106 (2010)
78. Surowiecki, J.: The wisdom of crowds. Anchor, Hamburg (2005)
79. Newman, M.E.: Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E. **74**(3), 036104 (2006)
80. Reichardt, J.: Stefan bornholdt: statistical mechanics of community detection. Phys. Rev. E. **74**(1), 1–14 (2006)
81. Raghavan, U.N., Albert, R., Kumara, S.: Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev. E. **76**(3), 036106 (2007)
82. Potts, R.B.: Some generalized order-disorder transformations. In: Mathematical Proceedings of the Cambridge Philosophical Society, vol. 48, pp. 106–109 (1952). Cambridge University Press
83. Joy, P., Kumar, P.A., Date, S.: The relationship between field-cooled and zero-field-cooled susceptibilities of some ordered magnetic systems. J. Phys. **10**(48), 11049–11054 (1998)
84. Fruchterman, T.M., Reingold, E.M.: Graph drawing by force-directed placement. Software. **21**(11), 1129–1164 (1991)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.