Decision Support Systems 66 (2014) 123–134

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Creating social intelligence for product portfolio design

Yung-Ming Li *, Hsuan-Ming Chen, Jyh-Hwa Liou, Lien-Fa Lin

Institute of Information Management, National Chiao Tung University, Hsinchu 300, Taiwan

A R T I C L E I N F O

Article history:
Received 25 March 2013
Received in revised form 5 May 2014
Accepted 14 June 2014
Available online 17 July 2014

Keywords:
Social media
Social creation
Influential experts
Online reviews
Product portfolio

A B S T R A C T

Increasing numbers of people are using social media to express their personal experiences. Recently, these plentiful user-generated data sources have been utilized promisingly by enterprises for product creation. In this research, we propose a social intelligence mechanism that can extract and consolidate the reviews expressed via social media and derive insights (product feature specification and feature importance) to help enterprises make decisions on developing next-generation products by analyzing the reviewers' knowledge and authority and their opinion sentiment toward the target products. The experimental results obtained using Epinions.com show that the proposed mechanism outperforms other benchmark approaches in market trend prediction and customer acceptance.

1. Introduction

Due to the rapid development of Web 2.0 technologies, the interactions between individuals have changed dramatically. More and more people are using blog, forum, and bulletin board systems to express their personal experiences. Gartner [13] reports that 40% of people usually search for products or services through social media, 77% of consumers read online reviews, and 75% of users trust online reviews more than personal recommendations. Nowadays, even strangers can easily exchange information through the Internet. A Nielsen Global Online Consumer Survey [31] indicates that 70% of people trust the online reviews posted by strangers. Online reviews have become a new and useful source for companies wishing to analyze people's opinions and infer their preferences. Enterprises can aggregate consumers' feedback to derive new strategies for product creation and design.

Social creation is a concept by which enterprises can utilize collective intelligence for new product or service creation. A survey conducted by OTX [32] shows that 71% of recommendations and information provided by consumers are valuable to companies. For example, Fiat, the most famous automotive brand in Brazil, built a website for people across the world to provide their opinions in a project called Fiat Mio. It received more than 10,000 suggestions from over 160 countries regarding the production of new cars. A popular e-commerce website, StyleFactory, displayed product images for customers to "make it" or "drop it" [28].

An enterprise can use a group of people to estimate the preference of the public. For example, enterprises usually construct a focus group or organize customer feedback within the company to understand customer needs for decision making [2]. On the other hand, the fiercely growing Internet provides an open and accessible resource to enable the enterprise to become closer to its customers. The online reviews are usually generated by consumers who have used the product. They explicitly or implicitly express their preferences or expectations regarding the next-generation products. Therefore, companies can analyze the reviews to enhance their products and produce new ones to match the needs of the majority of consumers. The next best-selling products will most likely be generated as a result of these insights from reviews [7]. Enterprises usually construct bulletin board systems or websites to allow customers to express their opinions conveniently. Although this method allows companies to collect customer opinions directly, the data are limited to specific aspects of a company. From a developer's viewpoint, understanding the whole market (including the competitors) is more important. Contrary to the specific channel of feedback collection, social media contain diverse perspectives of reviewers and constitute a superior choice for opinion collection. To achieve its goal, this research uses the concept of social creation, which utilizes social intelligence through the reviews expressed in social media to support product portfolio design. There are three main research questions to address in this research:

(1) How can firms identify the relevant information of quality from a large amount of reviews? In order to understand the market better and make better decisions, this study analyzes the reviewers' preferences and recognizes the people who have more related experience and knowledge regarding the target products.
(2) How can firms identify the influence of opinion expressers on online review websites? Opinion leaders are likely to influence the purchase intention or the preference of consumers for

* Corresponding author.
E-mail addresses: yml@mail.nctu.edu.tw (Y.-M. Li), qk5669@hotmail.com (H.-M. Chen), alioujh@gmail.com (J.-H. Liou), lienfa0704@gmail.com (L.-F. Lin).

http://dx.doi.org/10.1016/j.dss.2014.06.013

specific features. Therefore, the proposed approach aims to identify the opinion leaders, who have a powerful influence. Additionally, to discover the information of high influence produced by opinion leaders, the degree of relationship between reviewers should be measured to help enterprises make better decisions.

(3) How can firms consolidate the opinions from a number of identified reviewers? In online reviews, consumers usually express which features they like or dislike, but they cannot illustrate the scope of product specification that they would really like. Therefore, it will be beneficial to develop a mechanism that can extract and analyze customers' emotions regarding the specific feature and produce the scope of product specification. On the other hand, it would also be helpful if the proposed mechanism could predict the weight of each feature to help the developer set priorities for product development.

This research proposes a social media intelligence mechanism that can derive insights for enterprises to make decisions on next-generation product development. The generated insights include two parts: (1) inferred product feature specification and (2) product feature importance distribution. Specifically, inferred product feature specification analyzes reviews' content and recognizes the sentiments of the opinions to construct the scope of product specification [40]. The derived feature importance distribution shows how important a feature is when customers make purchasing decisions by ranking the importance of features in discussions. According to the insights, developers can evaluate the production cost and the selling price of the product to decide which feature should be added to the next-generation product first, and then formulate the development of product specifications.

The rest of the paper is organized as follows. Section 2 presents the related literature. Section 3 illustrates the system framework of the proposed social intelligence mechanism for product design. Section 4 describes the data collection, processing, and experiments. Section 5 demonstrates and evaluates the experimental results. Section 6 summarizes the research contributions and discusses the research limitations and future work.

2. Related literature

2.1. Product design process

For most of the enterprises, product design can roughly be separated into three main processes: concept definition, idea visualization, and mass-production [17]. Concept definition means to understand the similar products, competitions, sales of product, and the behavior of customer via market research. Idea visualization stands for expressing the concept of design through 2D drawing and 3D mock up. Mass-production aims to produce the product which is popular and competitive. There are plentiful studies in this field. For example, [8] focuses on the issue of concept definition. They construct emotional index and use data mining rule to measure consumers' preference. [14] studies the issue of idea visualization and uses a genetic algorithm to increase efficiency of the product design and use CAD to construct the prototype. This research focuses on concept definition. In the past, enterprises usually construct focus groups or organize the customer feedbacks within the company to understand customer needs for decision making [2]. An enterprise can use a group of people to estimate the preference of the public. Recently, prediction markets (PMs) have been applied in the domains of economics, politics, and sports. Researchers also used the concept to the evaluation of new product concepts, new product ideas and early stage technologies [9].

With booming social networking technologies, social media platforms have empowered Internet users to publish their creations and opinions and spread new content. In the past, messages were only delivered among friends. Nowadays, even strangers can search the information through the Internet easily. It also becomes a new way for company to collect people opinion and infer their preference. It is the concept of social design which means that enterprise aggregate consumers' feedback to devise new strategies, improve management flow, and reduce the risk. Abramovici & Lindner [2] analyze enterprise knowledge database established by enterprise customer feedback. Decker & Trusov [10] use product reviews data to infer consumer preferences.

However, previous researches did not consider interactions between people and their profiles. In this research, we use the concept of social creation (design) which combines the enterprise's design and the opinions in the popular social media to infer customers' preference to aggregate the social intelligence for enterprises to support product portfolio design.

2.2. Social creation

The concept of social creation is one type of co-creation, which is broadly defined as "creation of value by consumers" [41]. In the past, domain knowledge has been defined by and limited to a group of experts. The public only received experts' opinions, like those in Encyclopaedia Britannica. Nowadays, everyone who has domain knowledge can co-edit encyclopedias, like Wikipedia [21,38]. The co-creation mechanism transforms people's role from passive recipients to positive participants. It not only improves the efficiency of knowledge sharing, but can also help people to obtain a broader perspective of knowledge.

In the area of e-commerce, consumers are always the most important concern for enterprises. Any business strategies devised by enterprises should take consumers' perspective into account [3]. Hence, more and more companies are bringing consumers into the task of new product development. For example, IBM's Jam, a time limit online discussion platform, was developed to deal with the major issues of customer feedback. Another example is Polyvore, which uses an e-commerce website to collect information regarding consumers' combination of garments and shoes.

Although the concept of social creation in new product development has been proposed, the practical applications are still insufficient. Currently, most enterprises establish their own platform on which users can interact with each other. However, plenty of product and user preference information can be extracted from external data sources, such as the popularly emerging social media. This research focuses on exploiting the external opinions and reviews from social media and derives the social intelligence for new product design.

2.3. Feature selection

Due to the rapid development of web 2.0 technologies, more and more people use social media to express their personal experiences. Feature extraction methods are used to gather product features from a set of product reviews. Analyzing online reviews can help manufacturers better realize consumer opinions to their products and thereafter enhance their products [22]. Allan et al. [4] propose several techniques to discover relevant concepts and the topics of a query. The features of products mentioned in opinions are automatically identified by feature extraction. The authors of [16] generate a set of frequent features by finding out frequent terms and pruning the feature set by calculating term compactness and redundancy to extract product features. Abrahams et al. [1] developed a model of text classification and component feature selection for automatic product component categorization of online user postings.

The Red Opal system [35] also uses frequent nouns and noun phrases for feature extraction. The authors of [11] apply association rule mining mechanism to discover syntax rules of feature term occurrence, which could find out how frequently a feature term happens in some kind of syntax patterns. On feature-based opining summarization, Hu and Liu [15] propose Part-of Speech (POS) tag sequence rules to extract product attributes, and then the polarities of opinion words on the attributes are judged based on the context information. The POS tags of words, such as adjectives and adverbs, have been good indicators for the subjectivity detection and sentiment polarity classification. Besides, [39] have developed systems for analyzing customer reviews, and mining opinions toward a product or attribute, and visually show the mined information for aiding users' decision making.

In this research, we extract product feature from Amazon.com to get the basic product feature. Then analyzes the reviews and extracts the features which are not contained in the dictionary to compensate the lack of the product feature. We adopted POS tagging approach to identify the part of speech of each word in a sentence. After the POS tagging step, the system uses the frequency nouns to identify the product feature. The meaningless or non-feature words will be filtered.

2.4. Sentiment analysis

Recently, sentiment analysis (SA) has been popularly applied to the analysis of online customer reviews. For example, Reyes and Paolo [33] collect ironic reviews from Amazon and use sentiment analysis and opinion-mining techniques for purchase decision making. Cao et al. [7] also use a large collection of extracted texts to analyze the sentiment polarity of opinions. Li and Wu [24] use sentiment analyzing and text mining to classify texts collected from online forums. Archak et al. [5] propose that textual reviews can be extracted to represent different product features. Therefore, we use consumers' reviews to appraise the public's impression of product features.

Seva et al. [36] point out that in new product development users' emotion is one of the important factors. Designers should consider users' needs from the emotional perspective [34]. For example, Liu et al. [25] collect sentences that describe daily situations to construct ConceptNet, which automatically defines the keywords' sentiment perspective through the sentence categories. Shaikh et al. [37] propose an approach named SenseNet in which word characteristics are extracted from WordNet [27]. This research extracts users' experience of using a product, recognizes the polarity of their opinions, and then calculates their preference for the feature.

2.5. Opinion consensus

Opinion consensus integration is one of the important issues in group decision making. Ness and Hoffman [29] define consensus as a decision that is approved by most of the team members; only a few members disagree. Those members who hold opposing views have the ability to influence the decision. How to aggregate team members' opinions can be roughly divided into two different categories [6]. The first is "mathematically aggregated." This type of decision group has a complete arbitration mechanism. The team members do not need to exchange their opinions [30]. Those opinions are aggregated by voting or giving the team members or the opinions different weights. In another type of consensus, the opinions are revised continually by the team members until they can finally reach an approximate agreement.

This research adopts the first approach. A group decision-making model is developed to aggregate the online social network's opinions using team members' profiles and relationship to determine the expert weights.

3. The system framework

The social intelligence mechanism is developed for enterprises to extract easily the users' opinions expressed in social media and derive popular product features to design new product portfolios. To derive the collective intelligence to support new product creation, an enterprise chooses the targeted product category and then the system extracts the opinions related to the product category from social media. The collected online data are organized into a relationship of reviewers, opinion words, and rating scores. In Fig. 1, we illustrate the life cycle for new product portfolio design with the support of social media intelligence. After an enterprise has released a new product, users express their opinions via social media. Then the enterprise collects the relevant online opinions and derives the social intelligence to create or improve the design of the new-generation product.

The architecture of the proposed social intelligence system is depicted in Fig. 2. The components included in the system are described as follows.

(1) Feature dictionary construction module: With the collected data from the social media, this module constructs a feature dictionary according to the product features extracted. The dictionary is used to analyze the opinions expressed in the reviews.
(2) Knowledge analysis module: The module measures a reviewer's knowledge level by analyzing the content fitness and expertise levels. The knowledge level indicates the quality of the opinions he/she has published.
(3) Authority analysis module: The module analyzes a reviewer's authority by analyzing the relationship structure and review ratings. The authority level indicates the influence of the

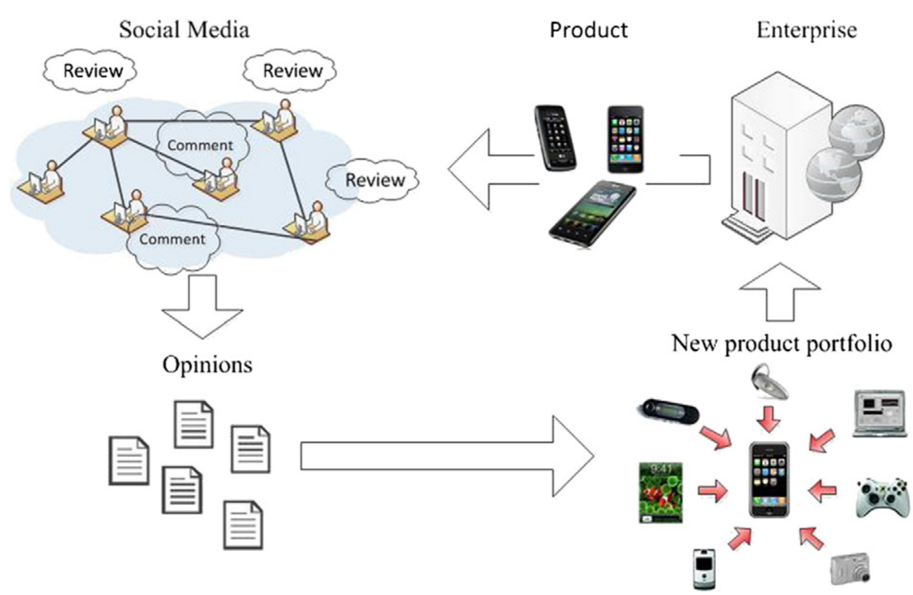

Fig. 1. The life cycle of new product portfolio design with social media intelligence.

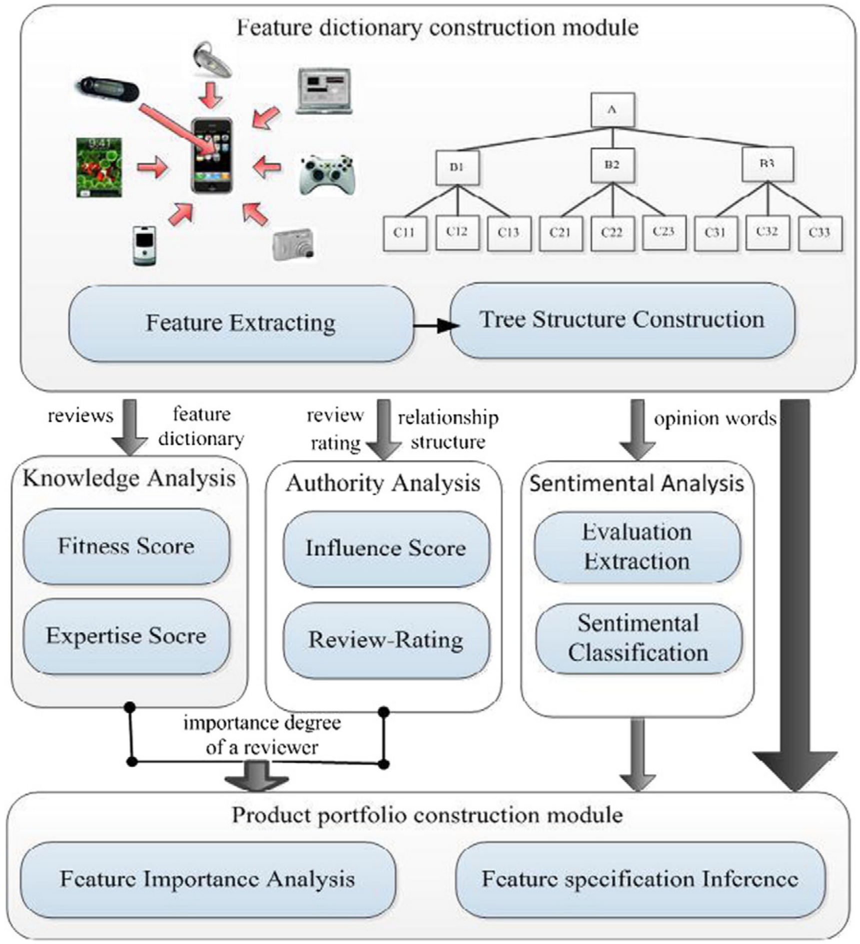

Fig. 2. The framework of the social design mechanism.

opinions he/she has published.

(4) Sentiment analysis module: The module analyzes the opinion words to predict the polarity of a comment (like or dislike).

(5) Product portfolio construction module: Utilizing the scores that are produced by authority analysis, knowledge analysis, and the polarity produced by sentiment analysis, this module aggregates the opinions and the features to create the product portfolio for the developers.

3.1. Feature dictionary construction module

The feature dictionary is composed of several categories of products. Each category contains many product features, which are organized into a tree structure. The tree structure is built by several levels and a lower level of a category is a subset of an upper one. An example of the feature dictionary is shown in Fig. 3.

The following steps are developed to construct a feature dictionary: feature extraction and tree structure construction.

3.1.1. Feature extraction

First of all, this study acquires basic product features from the merchant websites. To compensate for the insufficiency of the product features, we analyze the reviews and extract the features that are not contained in the initial dictionary. In this place, the POS (part-of-speech) tagging technique is used to identify the part of speech of each word in a sentence. The product features are detected based on the terms of nouns and noun phrases. Following the POS tagging step, the system uses the noun frequency [15] to identify the product feature. Meaningless or non-feature words are filtered out. Note that it is common for users to use synonyms to address the same products, so synonyms should be included in the product feature dictionary. For example, customers usually use “app” or “apps” to denote the word “application.” The two words will be added to the dictionary to increase the detection accuracy.

3.1.2. Tree structure construction

The classifications of the product features are acquired from the information on categories shown on the merchant websites. The product features are classified into corresponding categories according to the characteristics of the features identified by the noun frequency method. For example, nowadays, applications can easily be downloaded by smart phones and thus are highly related to the cell phone's operating system. Therefore, “application” should be classified into the category “operating system,” which is under the “cell phone” category.

3.2. Knowledge analysis

The main purpose of this module is to evaluate the quality importance of opinions according to the opinion relevance and the expertise level of an author. We consider two factors – fitness score and expertise score – to measure the knowledge importance of the experiences and opinions of an expresser.

3.2.1. Fitness analysis

People usually make more studies and have more experiences on the topics that interest them, so they are likely to have more knowledge about the topics they profoundly like. It may be easy to observe people's interests through their profiles expressed in social media with social

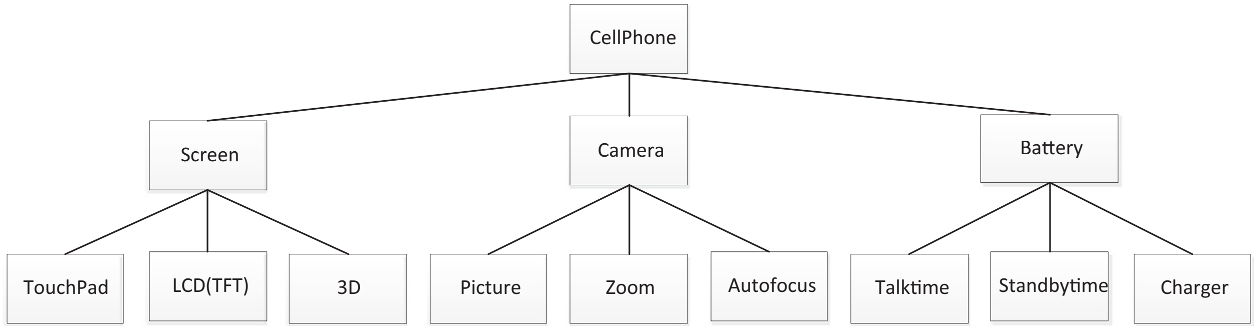

Fig. 3. The tree structure of product features.

networking purposes. However, it is difficult to infer people's interests on most online review websites as they seldom provide complete profiles. Therefore, this research analyzes the historical reviews posted by a reviewer to predict his/her preference. The reviews published by a reviewer are usually located in related categories. Therefore, we can predict the interests and preferences of a reviewer according to the types of reviews he/she has written. Fig. 4 shows an example of a category tree that most online review websites use to classify their reviews for easy retrieval at a later date.

In this research, to measure the fitness degree of a product to a reviewer, we calculate the average distance between the category of the target product and the categories of all the historical reviews published by a specific reviewer in the preference tree. Φ_i denotes the set of reviews written by user u_i and $c(r_j)$ represents the catalog to which a review r_j belongs. Eq. (1) shows the fitness degree of user u_i with target category t:

$$FS(u_i) = \frac{\sum_{r_j \in \Phi_i} dis\left(c\left(r_j \in \Phi_i\right), t\right)}{|\Phi_i|}, \tag{1}$$

where $|\Phi_i|$ indicates the total number of reviews written by user u_i, t stands for the target category, and $dis(c(r_j \in \Phi_i), t)$ represents the distance between the target category t and the category of a review r_j that was written by user u_i. For example, we suppose that the analysis target category is "cell phone", as shown in Fig. 3. If user u_A has written a review in "Charger" category, the preference distance for user u_A is 2. Another user u_B has written a review in "Camera" category, the preference distance for user u_B will be 1. If user u_C has written two reviews in different categories ("Charger" and "Camera"), the preference distance for user u_C will be 1.5 ($\frac{2+1}{2}$). A reviewer has more related knowledge if he/she gets a smaller value of average distance.

3.2.2. Expertise analysis

A person having published relevant reviews on a particular topic more than the average level represents that he/she has a higher knowledge level in the specific domain. A direct way of measuring a user's knowledge level on a certain topic is to count the reviews the reviewer has written in the specific category; however, this may lead to some problems if the relevance of the posted reviews is not evaluated. For example, if a reviewer has written many reviews in a specific category but those reviews are full of meaningless sentences, the accuracy of expert recognition should be problematic. Hence, in addition to counting the quantity of the reviewers a user has written, this proposed system also analyzes the content quality of the reviews. If a review includes many sentences relative to the topic, this review should be considered more important [26]. Therefore, we use the ratio of the sentences containing the product features to identify the expertise degree of the reviews written by a user: denote s as a typical sentence and $S_i = \{s | s \in \Phi_i\}$ as the set of all the sentences expressed in the reviews written by user u_i. W_s represents the set of words included in a sentence s and Θ_t stands for the set of product features included in target category t. The expertise score of a user u_i is formulated as Eq. (2):

$$ES(u_i) = \frac{|\{s | s \in S_i, W_s \cap \Theta_t \neq \emptyset\}|}{|S_i|}. \tag{2}$$

3.3. Authority analysis

The reviewers' authority degree can also be used to measure the credibility importance of the reviews. This module considers two common social factors, influence score and review rating, to measure a reviewer's authority.

3.3.1. Influence analysis

This factor is used to measure the influence of online users and is computed according to their social relationships. The more social relationships a person has built with other people, the more influential power he/she has. In this research, we use the following two network structures to calculate the influence of a reviewer.

3.3.1.1. Explicit link. The techniques of social network analysis can be easily applied to measure the intensity of relationships according to

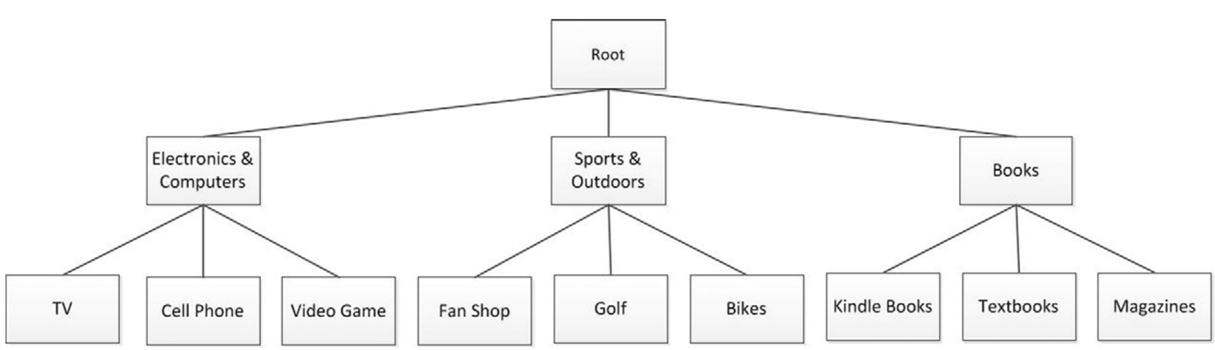

Fig. 4. The tree structure of the preference category.

the explicit links between people on social media (e.g. truster–trustee, follower–followee, or subscriber–subscribee). As the out-degree can easily be manipulated by people who add many friends without their permission, this research adopts the in-degree to prevent this problem.

3.3.1.2. Implicit link. Besides, the relationships could also be constructed by analyzing the activities and interactions among people. For instance, when two people participate in the same activity, a certain relationship between them is built. When a review receives many comments from others, it may indicate that the author prudently writes the review with substantial support or the author's review refers to some key points of the product that attract the readers' interest. Thus, we use the interactions between people including posts and replies to reviews and comments to measure the influential strength of a user. After publishing an article or opinion on the online review website, reviewers might receive some responses.

This research combines explicit and implicit links to construct the relation network. Denote SN as the set of users contained in the social network of reviewers. We use the degree centrality [20] to calculate the influence score as in Eq. (3).

$$IS(u_i) = \frac{\sum_{u_j \in SN} \left(Ea_{ij} + Ia_{ij} \right)}{|SN - \{u_i\}|} \tag{3}$$

where Ea_{ij} (Ia_{ij}) stands for the explicit (implicit) link from user u_j to user u_i. $Ea_{ij} = 1$ ($Ia_{ij} = 1$) if the link exists; $Ea_{ij} = 0$ ($Ia_{ij} = 0$), otherwise. $|SN - \{u_i\}|$ represents the total number of users in the social network except user u_i.

3.3.2. Review rating

Most of the websites supporting social creation provide an article rating mechanism, for example Wikipedia, Epinions, Yahoo Answers, and so on. It is a direct way to evaluate whether an article is good or bad. However, the perspective of experts may differ from the view of ordinary reviewers [20]. Therefore, expert reviewers should be given higher importance than regular reviewers. In this research, we also include a review rating factor for measuring the authority of a review. Denote $rate(r_j, u_k)$ as the review rating of review r_j given by user u_k.

Eq. (4) is used to calculate the collected reviewer rating score, which considers the importance of reviewers.

$$RS(u_i) = \frac{\sum_{u_k \neq u_i \in SN, r_j \in \Phi_i} rate\left(r_j, u_k \right) \times ES(u_k)}{|SN - \{u_i\}| \times |\Phi_i|}, \tag{4}$$

3.4. Sentiment analysis module

This module aims to appraise the public's impression of product features. Consumers usually use subjective words (e.g. good, perfect, bad, terrible) to express their emotion in reviews. To predict whether the customers like particular features or not, sentiment analysis is conducted. Subjective words would be recognized as positive, negative, or neutral and labeled as 1, −1, and 0, respectively. Accordingly, there are two major problems to solve first:

(1) How can we extract the subjective words correctly?
Here are some examples of sentences expressed in online reviews:

- "The sound on this TV is pretty decent."
- "The TV works perfectly with video game and DVD player."
- "The TV has many inputs."

The first and second sentences use subjective words (decent and perfectly) to describe the pros of TV, but they use different POSs (adjective and adverb). The third sentence illustrates that the TV is easy to use because it supports many different devices, but this sentence does not use any subjective words. To identify the subjective words, this study uses POS tagging to recognize the part of speech firstly, and then acquires the adjective and adverb that appear within three words' distance from the feature. We discovered that reviewers usually use adjectives and adverbs to express their emotion and those subjective words commonly appear within three words of the feature. Notice that a whole sentence that does not contain any subjective words will be reserved in the database and will be recognized manually.

(2) How can we classify and quantify the subjective words?
In this step, we use Senti-Wordnet to recognize the adjectives and adverbs produced previously. Senti-Wordnet is based on Wordnet, which classifies words into three different types (positive, negative, and objectivity) [12]. The distance between two different opinion words is measured by applying the polarity analysis interface provided by SentiWordNet. For identifying the sentiment of the review, we first deal with POS tagging processes on the product reviews and get 3657 adjectives and adverb terms. Then each term was fed into the query interface of SentiWordnet to calculate the polarity. If the positive score is greater than the negative score, the word will be tagged as 1. Otherwise, the word will be tagged as −1. Also, if the positive score is equal to the negative score, the world will be tagged as 0.

3.5. Product portfolio construction module

Although users usually depict the pros (satisfied) and cons (unsatisfied) of the products in their reviews, these perceptions are probably affected by several factors. Therefore, the proposed system considers and aggregates the quantitative measurements from the authority analysis, knowledge analysis, and sentiment analysis modules to identify the qualified reviews and evaluate the polarity of each review. In order to understand the market and make better decisions, this study integrates the factors of knowledge (fitness analysis and expertise analysis), authority (influence analysis and rating analysis) to evaluate the importance weighting of the reviewers. We select the reviewers with importance weighting score higher than 0 and smaller than 1 is to balance the quality of reviews ("wisdom of professionals") and the quantity of reviews ("wisdom of crowds"). According to our experiments, we found the threshold value 0.5 can better balance the quality and quantity effects and generate the best results. Therefore, we use it for further analysis. We further identify the qualified reviews from them. In the sentiment analysis module, we evaluate the polarity of each review. We reserve those positive comments and construct the interval of satisfied feature specification.

Finally, the enterprise will receive a proposal containing recommendations for product improvements and development priorities. Specifically, this proposed system will generate the inferred product feature specification and product feature importance distribution. The inferred product feature specification shows the consumers' acceptable scope for every feature specification of the product. Additionally, the product feature importance distribution indicates the public discussion proportion of each feature, which can be inferred as the priority of consumer decision making. With these outputs, the enterprise can better understand the market by knowing which features of their product are vital and what the customers want the product to be. The enterprise can use the result to develop the next generation of the product.

3.5.1. Feature importance analysis

This sub-module will identify the product improvement priorities by calculating the number of discussions on a specific product feature. When features are discussed, the discussions may be likely to influence the purchase intention of consumers. Since there are restrictions on the selling price and production cost, the enterprise can further evaluate the feature importance the system generated, and then make a decision regarding which specifications should be added to the next generation of the product first.

3.5.2. Feature specification inference

This sub-module will examine the polarity of the corresponding features and generate the interval of satisfied specification of each product feature. Specifically, the system first searches the units of product specification listed in the product dictionary, then grabs the digital number in front of it. Then, it recognizes the polarity of a sentence. Those positive comments are reserved and the negative ones are removed. Finally, the maximum and minimum numbers are used to construct the interval of satisfied feature specification. Here are some examples:

- "The 4.3 in. touch screen is really good."
- "4.0 in. screen is perfect for me."
- "I like the 3.7 in. screen."

The inferred feature specification for the screen size of smart phones will be [3.7 to 4.3].

4. Experiments

4.1. Experiment setting

In this study, we choose Epinions.com as the experimental platform; this platform is one of the most popular online view website. It contains multiple categories of reviews, such as television, cell phone, camera, and even movie and computer software. It includes a large number of reviews and several different mechanisms (trust list, rating, advisor, and so on). The product reviews on Epinions.com are separated into two groups: express reviews and regular reviews. This work collects regular reviews as the analysis objects because people usually express their real emotions and uses their experience in the reviews. A regular review, containing more than 200 words, allows users to describe the characteristics of products in a more detailed way [35]. The review data was collected from April 2010 to April 2012 and totally 938 reviews were gathered. The information extracted from the reviews includes the product, review context, reviewers, trust list, comment list, rating score, etc.

In this research, we use the following steps to construct a feature dictionary: POS (part of speech) tagging, feature extracting, synonym recognizing, and the tree structure constructing. The system obtains a three-tier structure tree and 321 classifications. This study chooses "cell phone" as the target product category for analysis because cell phones combine multiple features (bluetooth, camera, mp3 player, and so on), and the differentiation can easily be observed when the product has diverse characteristics. We totally analyze 33,502 sentences and 3657 opinion words. The system uses the Senti-Wordnet to analyze the category of opinion words to calculate the polarity. If the positive score is greater than the negative score, the word will be tagged as 1; otherwise, the word will be tagged as −1. Also, if the positive score is equal to the negative score, the word will be tagged as 0. The trust list and comment list are used to construct the relationship network. Epinions.com uses five different levels (off topic, somewhat helpful, helpful, very helpful, and most helpful) to review the reviews. This study transforms the grading mechanism into the range of integers from −2 to 2 to measure the reviewers' credibility. Finally, we use up-to-date of top five sales information from Amazon.com to evaluate the accuracy of the inferred feature specification.

4.2. Data collection and preprocessing

This study chooses "cell phone" as the target product category for analysis because cell phones combine multiple features (bluetooth, camera, mp3 player, and so on), and the differentiation can easily be observed when the product has diverse characteristics. The review data were collected from April 2010 to April 2012 and in total 938 reviews were gathered. The information extracted from the reviews includes the product, review context, reviewers, trust list, rating score, etc.

First of all, the proposed system acquires the features' specification and classification from Epinions.com. For example, the display entry contains the screen size, display size, and touch screen. The battery entry contains stand-by time, talk time, and power. Furthermore, some product features are divided into smaller portions (digital and unit) for the sentence recognition. For example, the screen size marked as 4.3 in. will be divided into 4.3 in.

To amplify the feature dictionary, the system splits the review context into several sentences and uses BFSU Stanford POS Tagger, developed by Xu and Jia [39], to recognize the part of speech in the sentence. This study acquires nouns to detect the features that do not appear in the dictionary or have different expressions. The system analyzes 33,502 sentences and abstracts 10,841 nouns. The result of the frequency noun mining is shown in Table 1. The dictionary has eight main categories, which can be subdivided into several items, and the words in the brackets after the items stand for the synonyms. We also deal with singular and plural nouns. The last field is the unit of feature specification and is used to produce the inferred feature specification.

To establish a feature dictionary with higher accuracy and more completeness, we extract product feature from Amazon.com as the basic product features structure and extract the features from the up-to-date reviews from Epinions.com which are not contained in the dictionary. We adopted the POS tagging approach to identify the part of speech of each word in a sentence. After the POS tagging step, the system uses the frequency nouns to identify the product feature.

Table 1
The dictionary of product features.

Category	Detail	Unit
Screen	LCD (TFT)	Inch (in.)
	Touch screen	
	3D	
Processor	Storage (RAM, memory, built-in memory)	GB/MB
	Processor	GHz/MHz
Communication	3G	
	4G	
	Wi-Fi (wlan, wireless)	
Camera	Picture (photo, graphic, photograph)	Megapixels
	Zoom	
	Autofocus	
Multimedia	Video	
	HD	
	Player (mp3)	
	Headphones (stereo jack)	
	Speakerphone (audio, voice, volume, sound)	
Battery	Talk time	Hours
	Stand-by time	Hours
	Charger	
Operating system	Application (app)	
	Software (email, calendar, browser)	
	Interface (UI, settings)	
Other feature	USB (micro-usb)	
	SD (micro-sd)	
	Keyboard (keypad, qwerty)	
	Bluetooth	
	GPS	

Table 2
Principal component analysis.

Component	Total	% of variance	Cumulative %
Influence score	1.560	38.994	38.994
Review rating	1.057	26.428	65.421
Fitness score	0.89	22.249	87.671
Expertise score	0.493	12.329	100

Table 3
The feature importance distribution (cell phone).

Feature	Percentage (%)
Screen	15.45
Processor	4.90
Communication	7.0
Camera	16.26
Multimedia	17.21
Battery	11.58
Operating system	22.42
Other feature	5.18
All	100

4.3. Knowledge and authority analyses

4.3.1. Fitness score computing

To predict a reviewer's preference and the review fitness, this study needs to construct the review category tree first. As Epinions.com has already organized its reviews into a tree structure, we acquire the classification and reviews category tags to calculate the average preference distance of each reviewer. The system obtains a three-tier structure tree and 321 classifications.

4.3.2. Expertise score computing

All the reviews of a specific reviewer in the cell phone category are aggregated into one document. Then we detect the feature words defined in the feature dictionary from the aggregated reviews. The ratio of sentences containing the feature words is used to analyze the expertise of the reviewer.

4.3.3. Influence score computing

This research uses the trust list and comment list of Epinions.com to construct the relationship network. In Epinions.com, the trust list has two types: "who trusts the user" and "who is trusted by the user." The list of "who trusts the user" lets users follow specific reviewers. When they decide whether or not to buy a specific product reviewed by some reviewers, they can choose to show the reviewer's review first. The list of "who is trusted by the user" shows how many users trust a specific reviewer and the list will influence the advisor selection process in Epinions.com. The system uses the list of "who trusts the user" to analyze the users' popularity because the "who is trusted by the user" list can easily be manipulated by the users.

4.3.4. Review rating computing

Epinions.com provides a rating mechanism that helps users express which kinds of reviews they would like to read in the future and suggest to other reviewers which reviews should be read and which should not. It uses five different levels (off topic, somewhat helpful, helpful, very helpful, and most helpful) to review the reviews. This study transforms the grading mechanism into the range of integers from −2 to 2 to measure the reviewers' credibility.

Finally, we aggregate the scores generated from the four modules in the framework (i.e., *IS*, *ES*, *FS* and *RS*) to calculate the importance weighting of each review author. The aggregated importance weighting score of an author u_i is formulated as:

$$AS(u_i) = \alpha IS(u_i) + \beta ES(u_i) + \gamma FS(u_i) + \delta RS(u_i), \quad (5)$$

where α,β,γ, and δ are different weights of the factors (e.g. influence, expertise, fitness, review). In this research, we use principal component analysis (PCA) [18] to obtain the weight setting and select the reviewers whose importance weighting score is higher than 0.5 as domain experts. The total variance explanation table is shown in Table 2.

4.4. Sentiment classification

To identify the sentiment tendency of the sentence, we have to acquire the Senti-Wordnet database first. The database contains the POS of each word and its positive or negative score. This study uses adjectives and adverbs for sentiment analysis, and there are in total 18,156 adjectives and 3621 adverbs. After the POS tagging process, the system obtains a total of 3657 adjectives and adverbs from the product reviews. The data of Senti-Wordnet are used in a comparison with the adjectives and adverbs recognized by the proposed system in order to calculate their polarity. If the positive score is greater than the negative score, the word will be tagged as 1; otherwise, the word will be tagged as −1, and if the positive score equals the negative score, the word will be tagged as 0.

4.5. Product portfolio generation

This proposed system also generates the portion distribution of discussions (as a percentage). The ratio of discussion is shown in Table 3. A feature that has a higher discussion ratio implies that a customer may consider it as a purchase consideration of high priority. This can help developers to adjust the development strategies for different product features when producing new products.

According to the feature dictionary (inches, megapixels, and so on) and the polarity words recognized by the sentiment analysis, the inferred feature specification is produced and outlined in Table 4.

Table 4
The inferred feature specification (cell phone).

Category	Range
Screen size	3.7–4.3 in.
Processor	1–1.2 Ghz
Camera	5–8 megapixels
Built-in memory	512 MB–1 GB
Battery	6–8 h

5. Results and evaluation

5.1. Market prediction accuracy

To verify the effectiveness of the proposed mechanism, we compare each feature contained in the inferred feature specification with the sales amount of the top five smart phone touchpads on Amazon.com on 31 May 2012. As Amazon.com is the largest electronic commerce website, its sales rankings have high representativeness and credibility. The features of the top five cell phones are shown in Table 5.

Table 5
The features of the top five cell phones.

	Product A	Product B	Product C	Product D	Product E
Screen size	4.3 in.	4.65 in.	4.3 in.	4.52 in.	4.5 in.
Processor	1.2 Ghz	1.2 Ghz	1.4 Ghz	1.2 Ghz	1.5 Ghz
Camera	8 megapixel	5 megapixel	8 megapixel	8 megapixel	8 megapixel
Built-in memory	1 GB	1 GB	512 MB	1 GB	1 GB
Battery	6.3 h	8 h	6.7 h	8 h	6.7 h

Table 6
The feature importance distributions generated by different approaches.

Feature	No filter	Authority	Knowledge
Screen	17.50	16.30	15.75
Processor	3.81	3.79	5.54
Communication	6.22	6.92	7.27
Camera	15.53	15.94	17.09
Multimedia	17.28	16.73	16.14
Battery	11.09	11.43	8.05
Operating system	22.63	23.26	23.02
Other feature	5.95	5.64	7.14
All	100	100	100

We also generate the feature specification and importance distribution based on other benchmark approaches—(1) traditional voting (no filter), (2) reviewer's authority based, and (3) reviewer's knowledge based. The feature importance distributions and the inferred feature specifications, which are produced by the no filter, authority, and knowledge approaches, are shown in Table 6 and Table 7.

The effectiveness comparisons are conducted in the following steps:

(1) Examine whether the features generated in the inferred feature specification are contained in the top five cell phones. If a feature is contained, the feature item will be tagged with a score using the weight (percentage) specified in the feature importance distribution table.
(2) Sum the scores of all the contained features and normalize the total score into a five-point Likert scale.
(3) Compare the Likert score produced by different approaches with the product ratings shown on Amazon.com.

The comparisons of rating results with respect to different products are shown in Fig. 5. We can observe that the prediction of our proposed approach (social design) is most close to the actual user evaluation for all the products. Fig. 6 shows the MAEs (mean absolute errors) generated by the different approaches. It is apparent that the proposed social design approach outperforms all the other benchmark approaches.

We further compare the accuracy of the predicted range of inferred feature specifications by the distance between the placement position and the upper and lower bounds, as shown in Eq. (6).

$$\text{distance} = \sqrt{(p-l)^2 + (u-p)^2}, \tag{6}$$

where p stands for the placement position and l and u represent the upper and lower bounds. If an inferred feature specification contains the feature and has a smaller range, the recommendation has higher accuracy. After normalization, the results are shown in Fig. 7. The segments closer to zero stand for the predicted scope being closer to the target value. Fig. 7 indicates that the proposed mechanism has a higher level of accuracy than the other benchmark approaches. Notice that a narrower range of the inferred feature has significant economic implications. For example, the narrower range of the predication means the smaller range of inferred features, which have already matched the requirements of the majority of consumers. Hence, the product provider only needs to consider this small range of this feature when producing a new product. It can save a lot of production costs.

Table 7
The inferred feature specifications generated by different approaches.

Category	No filter	Authority	Knowledge
Screen size	3.1–4.3 in.	3.5–4.3 in.	3.7–4.3 in.
Processor	600–1.2 Ghz	800–1.2 Ghz	600–1.2 Ghz
Camera	3–8 megapixels	5–8 megapixels	5–8 megapixels
Built-in memory	256 MB–1 GB	256 MB–1 GB	256 MB–1 GB
Battery	5–12 h	5–8 h	6–10 h

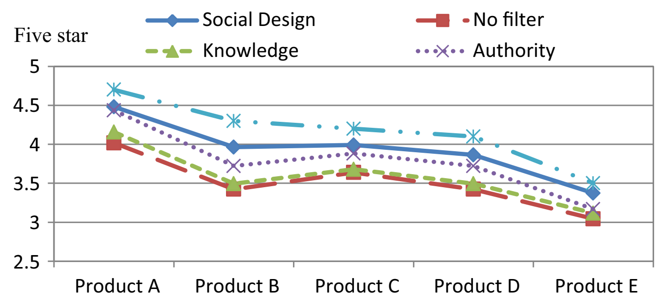

Fig. 5. Effectiveness ratings of different approaches.

By observing the reviews, we find that regular people usually express their feelings directly, such as why they chose this phone or to illustrate some specific features, but people who have related knowledge describe the overall feature in more detail. Furthermore, they may list the pros and cons of the product. Thus, they can provide better opinions than the results presented in Figs. 7 and 8. The authority approach has greater performance. An influential person can publish persuasive reviews affecting other people's purchasing intention and be followed by a large number of people. However, the knowledge approach does not perform as well as the authority approach.

We further evaluate whether the type of product will influence the result of the social intelligence mechanism or not. We select two other products (cameras and PC laptops) to measure the difference. According to the electronics product launch schedule, the system collects one year of data, from April 2011 to April 2012, and acquires the sales of the top five products on 31 May 2012 from Amazon.com in the same way. This research calculates the MAE of five-star rating and the score that is produced by the mechanism (Fig. 8). The study also calculates the placement position of the product features with the upper and lower bound distances due to the scope size problem mentioned above (Fig. 9).

The results indicate that the MAE and the distance scope of PC laptops are higher than those for the other two products. Because the average price of PC laptops is higher than the price of the other items, customers perform more consideration (gaming, small business, family use, etc.) when purchasing. For example, customers want the processor of their PC laptop to have high performance when playing video games. In contrast, office workers would prefer a smaller screen size for ease of carrying. This situation may decrease the forecast of the feature importance distribution and inferred feature specification. Despite the demand difference problem, the accuracy of the proposed social intelligence mechanism is still higher than that of the other approaches.

5.2. Consumer acceptance evaluation

We further used questionnaires to evaluate whether the result is consistent with the consumer acceptance. We invited people who have bought a smart phone or been willing to buy a smart phone recently to be respondents. The questionnaires were collected from 5 May 2012 to

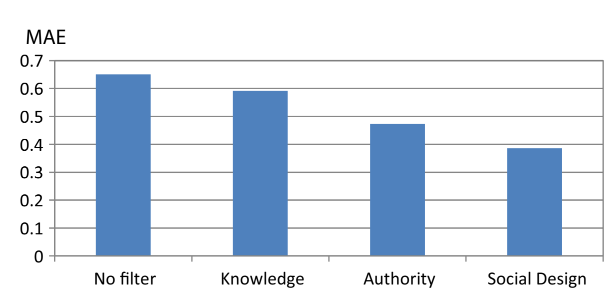

Fig. 6. The MAEs of different approaches.

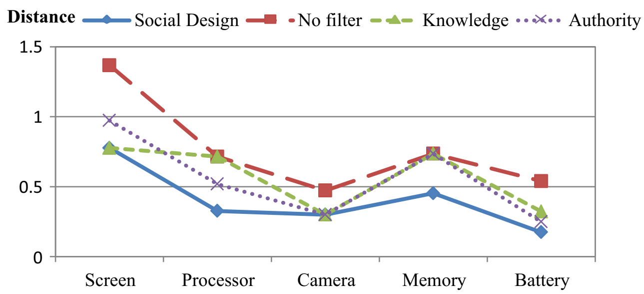

Fig. 7. The comparison of inferred feature specification.

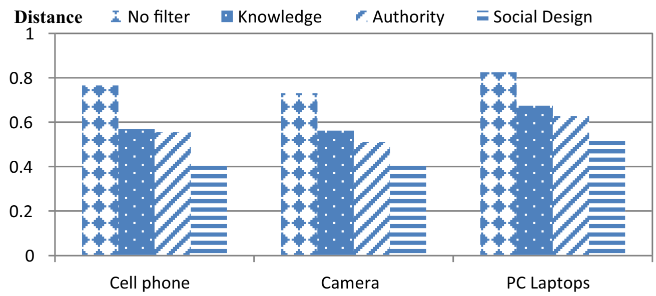

Fig. 9. The distances of different products.

10 June 2012 and 103 responses were received in total. The questionnaire has two parts: the first one requires the respondents to write down the specification of their own cell phone, and then they rate the specification from 1 to 5 points. It is assumed that people like this specification when the point is greater than or equal to 3. On the contrary, a point is less than 3 implies that they do not like this specification and 0 represents no opinion. The preference of respondents for cell phones is shown in Fig. 10.

This study regards people as likely to accept the result of inferred feature specification in the following two conditions:

(1) The specification is in the scope of the inferred feature specification and people express that they like this specification.
(2) The specification is not in the scope of the inferred feature specification and people express that they dislike this specification.

This research calculates the acceptance rate of each feature according to the rule mentioned above. The result is shown in Table 8 and the average acceptance rate reaches 75%.

Another part of the questionnaire gives the respondents eight categories of features and requires them to rank the categories according to their considerations when purchasing a cell phone. This research sets the first rank as eight points and the rest may be deduced by analogy. After summing all the scores, the results are shown in Fig. 11.

This work compares the rank of the feature importance distribution produced by the social intelligence mechanism with the questionnaire (see Table 9).

We use the Kendall tau rank correlation coefficient (Kendall's τ) [19] to compare the ranking similarity between the questionnaire and the social design mechanism, as shown in Eq. (7).

$$\tau(r_a, r_b) = \frac{n_c - n_d}{n_c + n_d}, \tag{7}$$

where r stands for the different rankings and n_c and n_d represent the number of concordant and discordant pairs. The value lies between -1 and 1, and it equals to 1 when the ranking results are completely the same. The statistic result is shown in Table 10, and it indicates that the ranks of the questionnaire and the social design mechanism are similar. The result also shows that when talking about cell phones, most of the consumers are concerned with the features of cell phones' multimedia, camera, and operating system. It is very intuitive; nowadays, the smart phone is becoming increasingly popular. People can download different applications at will and receive emails, watch videos, and browse the Internet anytime and anywhere. They also use cell phones to take pictures and upload the pictures to Facebook or Twitter. Thus, the multimedia, camera, and operating system have a higher discussion percentage.

To summarize the evaluation discussion, the results of the sales amount and questionnaires show that the feature importance distribution and the inferred feature specification are not only consistent with the market, but also consistent with the consumer need. The proposed system can produce accurate suggestions for developers to make better decisions on designing their product portfolio.

6. Discussion and conclusion

Social media empower customers to share their comments on brands or products instantly and provide an open and accessible resource to allow enterprises to become closer to their customers by analyzing customer opinions and deriving market intelligence. Taking the opinions of customers into the business as beneficial resources will gradually become an essential trend. It will greatly enhance enterprises' competitive advantage when they utilize collective reviews from social media and derive insights for new product development.

In order to discover relevant market intelligence on product portfolios, this study analyzes the opinions expressed by influential reviewers who have high levels of knowledge and experience of the target products. Through the relationship of people and their profile, the proposed system could identify the appropriate domain experts of high authority to filter meaningless data and improve the quality of the information, thus helping enterprises to reduce their cost and time consumption. Our experimental results show that the proposed social intelligence mechanism outperforms that of other benchmark approaches in

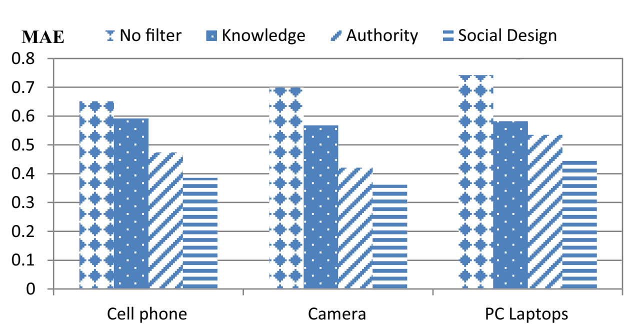

Fig. 8. The MAEs of different products.

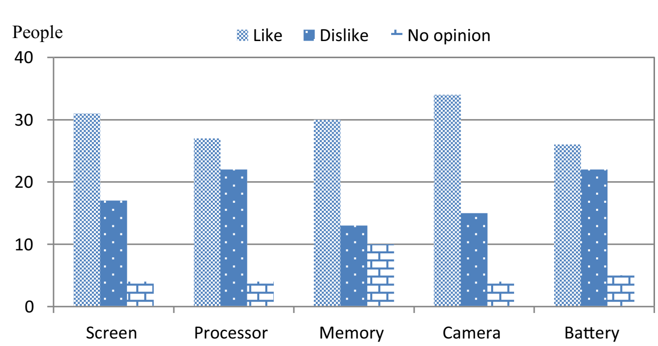

Fig. 10. User acceptance evaluation.

Table 8
The accuracy of each feature.

	Screen	Processor	Memory	Camera	Battery
Acceptance rate (%)	75.47	81.13	73.58	71.70	71.70

predicting market trends and customer acceptance in new-generation product design. According to the individual opinions and experiences expressed in social media, this proposed framework can effectively produce the feature importance distribution and the inferred feature specification for enterprises to make decisions for new product portfolio development. The proposed framework is particularly applicable to those product types with clear product features, shorter product generation life cycle, and trending products.

6.1. Research contributions

The research contributions and practical contributions of this paper are summarized as follows. Firstly, from the system design perspective, we integrate the opinion mining and social network analysis techniques to build social intelligence from social media which can help enterprises observe the trend of next-generation product development. The major applications of opinion mining are analyzing product reviews [3,5], or deriving market intelligence for supporting decision-makers [23]. However, the support of practical applications for new product development is still insufficient. The proposed system can automatically generate insights (inferred product feature specification and feature importance distribution). According to the generated information of product features, developers can refer to the cost and the selling price of the product to decide which features should be added into the next-generation of product first, and then formulate the development of product specifications. Secondly, from the methodological perspective, the proposed social intelligence mechanism allows us to identity the important reviews so as to balance the quality (wisdom of the professionals) and quantity (wisdom of the crowds) of online reviews by synthesizing the components of knowledge analysis and authority analysis and to analyze corresponding opinions to derive the popular product features by use sentiment analysis. Comparing with most of the existing works, the proposed mechanism considers more comprehensive factors such as review's fitness, reviewer's expertise, reviewer's influence, and review rating to evaluate the importance of a reviewer. The authority approach can make up some drawback from the knowledge approach. For example, the advertisers who mention a lot of product features and publish the reviews in the same category will cause the high knowledge score. By combining with the authority approach, this kind of reviews will be filtered effectively so as to further enhance the overall quality of information. Thirdly, from the practical perspective, the method of focus groups used in traditional market research present problems the problems of small data quantity, huge cost of expense, and long-time consumption. The proposed mechanism can automatically gather opinions expressed in public and analyze preference information from social media to improve the effectiveness of new product design. The inferred product feature specification shows the consumers' acceptable scope of every feature specification of the product. The product feature importance distribution indicates the public discussion proportion of each feature which can be inferred as the priority of consumer decision making. With these outputs, the enterprise can understand the market, knowing which features of their product are vital and how the customers want it to be. Fourthly, from the empirical perspective, utilizing the Epinion.com and Amazon.com platform, we verify that our proposed mechanism can effectively suggest the cell phone product's feature specification and importance. The product feature specification generated form the reviews on Epinions.com is consistent with the market evidences according to the sales amount (sales of top five cell phone form Amazon.com). For cell phones, the features of their multimedia, camera, and operating system are the top priority for consideration.

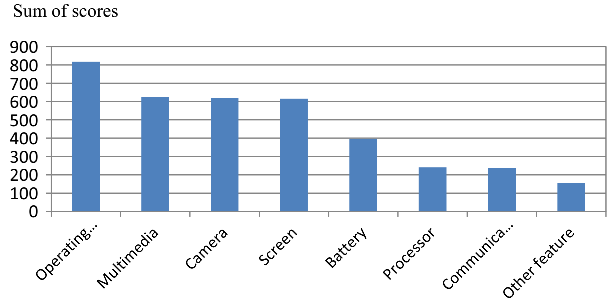

Fig. 11. Score ranking of product features.

Table 9
The rank of the feature category.

Rank	Questionnaire	Social design
1	Operating system	Operating system
2	Multimedia	Multimedia
3	Camera	Camera
4	Screen	Screen
5	Battery	Battery
6	Processor	Communication
7	Communication	Other feature
8	Other feature	Processor

6.2. Limitations and future studies

There are some research limitations and several extended studies that can be further undertaken. Firstly, while this research recognizes the domain experts to measure the importance degree of each opinion, it does not take their basic profiles (occupation, age, gender) into account. In the real world, the product design process may focus on different customer groups. Secondly, in this study, we discover that several products in the top sales belong to the same brand and some brands generate a relatively large number of discussions. Although this study does not discuss brand loyalty and the snowball effect, we believe that the brand could be one of the important factors that influence the purchase intention. Thirdly, utilizing the proposed mechanism, enterprises can improve and refine their products and produce the next generation of products for the market. In this cycle, the system plays a role of product evolution. However, the system cannot achieve a product revolution. For example, this study cannot propose the suggestion of new technology that can be added to the next generation of products. The issue of how to achieve the product revolution is worth studying. Lastly, the experimental data are acquired from Epinions.com. There are still several online review websites like Amazon.com or CNET that have some different properties. An extended work could conduct experiments over different social media and heterogeneous products. Diverse social media data can help enterprises make a comprehensive survey of the market.

Table 10
Kendall's tau rank correlation coefficient.

		Questionnaire	Social intelligence
Questionnaire	Correlation coefficient	1.000	0.857**
	Sig. (2-tailed)		0.003
	N	8	8
Social intelligence	Correlation coefficient	0.857**	1.000
	Sig. (2-tailed)	0.003	
	N	8	8

** The mean difference is significant at the 0.05 level.

Acknowledgment

This research was supported by the National Science Council of Taiwan (Republic of China) under the grant NSC 101-2410-H-009-009-MY3.

References

[1] A.S. Abrahams, J. Jiao, W. Fan, G.A. Wang, Z. Zhang, What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings, Decision Support Systems 54 (4) (2013) 871–882.
[2] M. Abramovici, A. Lindner, Providing product use knowledge for the design of improved product generations, CIRP Annals-Manufacturing Technology 60 (1) (2011) 211–214.
[3] B. Allan, Social enterprise: through the eyes of the consumer (prepared for the National Consumer Council), Social Enterprise Journal 1 (1) (2005) 57–77.
[4] J. Allan, J. Carbonell, G. Doddington, J. Yamron, Y. Yang, Topic detection and tracking pilot study: final report, Proceedings of the DARPA Broad-cast News Transcription and Understanding Work- shop, 1998.
[5] N. Archak, A. Ghose, P.G. Ipeirotis, Deriving the pricing power of product features by mining consumer reviews, Management Science 57 (8) (2011) 1485–1509.
[6] D. Ben-Arieh, T. Easton, Multi-criteria group consensus under linear cost opinion elasticity, Decision Support Systems 43 (3) (2007) 713–721.
[7] Q. Cao, A.M. Thompson, Y. Yang, Sentiment analysis in decision sciences research: an illustration to IT governance, Decision Support Systems 54 (2013) 1010–1015.
[8] L.C. Chen, P.Y. Chu, Developing the index for product design communication and evaluation from emotional perspectives, Expert Systems with Applications 39 (2) (2012) 2011–2020.
[9] L. Chen, P. Goes, J.R. Marsden, Z. Zhang, Design and use of preference markets for evaluation of early stage technologies, Journal of Management Information Systems 26 (2009) 45–70.
[10] R. Decker, M. Trusov, Estimating aggregate consumer preferences from online product reviews, International Journal of Research in Marking 27 (4) (2010) 293–307.
[11] X. Ding, B. Liu, L. Zhang, Entity discovery and assignment for opinion mining applications, International conference on Knowledge discovery and data mining ACM, Paris, France, 2009. 1125–1134.
[12] A. Esuli, F. Sebastiani, Sentiwordnet: A Publicly Available Lexical Resource for Opinion Mining, Citeseer, 2006.
[13] Gartner, The Future of the Social Customer, 2012.
[14] S.W. Hsiao, F.Y. Chiu, S.H. Lu, Product-form design model based on genetic algorithms, International Journal of Industrial Ergonomics 40 (3) (2010) 237–246.
[15] M. Hu, B. Liu, Mining opinion features in customer reviews, American, Association for Artificial Intelligence (2004) 755–760.
[16] B.J. Jansen, M. Zhang, K. Sobel, A. Chowdury, Twitter power: tweets as electronic word of mouth, Journal of the American Society for Information Science and Technology 60 (2009) 2169–2188.
[17] H. Jin, Study for Industrial Design Based on Knowledge Management, 2008.
[18] I.T. Jolliffe, MyiLibrary, principal component analysis, Wiley Online Library, 2, 2002.
[19] F.N. Kerlinger, H.B. Lee, Foundations of Behavioral Research, Wadsworth Publishing, 1999.
[20] C. Kiss, M. Bichler, Identification of influencers—measuring influence in customer networks, Decision Support Systems 46 (1) (2008) 233–253.
[21] A. Kittur, E.H. Chi, B.A. Pendleton, B. Suh, T. Mytkowicz, Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie, World Wide Web 1 (2) (2007) 19.
[22] Y.C. Ku, C.P. Wei, H.W. Hsiao, To whom should I listen? Finding reputable reviewers in opinion-sharing communities, Decision Support Systems 53 (2012) 534–542.
[23] Y.M. Li, T.Y. Li, Deriving market intelligence from microblogs, Decision Support Systems 55 (1) (2013) 206–217.
[24] N. Li, D. Wu, Using text mining and sentiment analysis for online forums hotspot detection and forecast, Decision Support Systems 48 (2010) 354–368.
[25] H. Liu, H. Lieberman, T. Selker, A Model of Textual Affect Sensing using Real-World Knowledge, ACM, 2003.
[26] J. Liu, Y. Cao, C.Y. Lin, Y. Hung, M. Zhou, Low-quality product review detection in opinion summarization, Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), 2007, pp. 334–342.
[27] G.A. Miller, WordNet: a lexical database for English, Communication of the ACM 38 (11) (1995) 39–41.
[28] C.B. Myers, How to Effectively Crowdsource Product Design, http://thenextweb.com/socialmedia/2010/11/19/how-to-effectively-crowdsource-product-design/ 2010.
[29] J. Ness, C. Hoffman, Putting Sense into Consensus: Solving the Puzzle of Making Team Decisions, Judy Ness, 1998.
[30] K.C. Ng, B. Abramson, Consensus diagnosis: a simulation study, IEEE Transaction on Systems, Man and Cybernetic 22 (5) (1992) 916–928.
[31] Nielsenwire, Global Advertising: Consumers Trust Real Friends and Virtual Strangers the Most, 2009.
[32] OTX, The Impact of Social Media on Purchasing Behavior, 2008.
[33] A. Reyes, R. Paolo, Making objective decisions from subjective data: detecting irony in customer reviews, Decision Support Systems 53 (2012) 754–760.
[34] Y. Rogers, H. Sharp, J. Preece, Interaction Design: Beyond Human-Computer Interaction, Wiley Publishing, 2011.
[35] C. Scaffidi, K. Bierhoff, E. Chang, M. Felker, H. Ng, C. Jin, Red opal: product-feature scoring from reviews, 8th ACM Conference on Electronic Commerce ACM New York, NY, USA, 2007, pp. 182–191.
[36] R.R. Seva, H.B.L. Duh, M.G. Helander, The marketing implications of affective product design, Applied Ergonomics 38 (6) (2007) 723–731.
[37] M.A.M. Shaikh, H. Prendinger, M. Ishizuka, SenseNet: A Linguistic Tool to Visualize Numerical-Valence Based Sentiment of Textual Data, 2007.
[38] Wikipedia, Epinions, available from: http://en.wikipedia.org/wiki/Epinion 2006.
[39] J. Xu, Y. Jia, BFSU Stanford POS Tagger: A Graphical Interface Windows Version, Beijing: National Research Center for Foreign Language Education, Beijing Foreign Studies University, Beijing, 2011.
[40] J. Yi, T. Nasukawa, R. Bunescu, W. Niblack, Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques, in: J. Shavlik (Ed.), Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, Florida, USA, 2003, pp. 427–434.
[41] V. Zwass, Co-creation: toward a taxonomy and an integrated research perspective, International Journal of Electronic Commerce 15 (1) (2010) 11–48.

Yung-Ming Li is a Professor at the Institute of Information Management, National Chiao Tung University in Taiwan. He received his Ph.D. in Information Systems from the University of Washington. His research interests include network science, Internet economics, and business intelligence. His research has appeared in IEEE/ACM Transactions on Networking, INFORMS Journal on Computing, Decision Sciences, European Journal of Operational Research, Decision Support Systems, International Journal of Electronic Commerce, Information & Management, ICIS, WITS, among others.

Hsuan-Ming Chen received his M.S. degree from the Institute of Information Management, National Chiao Tung University in Taiwan and B.S. degree in Information Management from the National Dong Hwa University, Taiwan. His research interests focus on electronic commerce and social computing.

Jyh-Hwa Liou is a Ph.D. student at the Institute of Information Management, National Chiao Tung University in Taiwan. Her research interests include electronic commerce and business intelligence.

Lienfa Lin is a Ph.D. student at the Institute of Information Management, National Chiao Tung University in Taiwan. He received his M.S. in Computer Science from Yuan-Ze University and B.S. degrees in Information Management from Fu-Jen University in Taiwan. His research interests include electronic commerce, mobile computing, and network economics. His research has appeared in International Journal of Electronic Commerce, Information Sciences, and Information Processing and Management.