Engineering Applications of Artificial Intelligence 47 (2016) 38–48

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

What makes consumers unsatisfied with your products: Review analysis at a fine-grained level

Jian Jin [a,b,*], Ping Ji [b,c], C.K. Kwong [c]

[a] Department of Information Management, School of Government, Beijing Normal University, Beijing 100875, China
[b] The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China
[c] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Available online 30 May 2015

Keywords:
Customer requirement
Sentiment analysis
Review analysis
Product design
Conceptual design
Text mining

ABSTRACT

Online product reviews contain valuable information regarding customer requirements (CRs). Intelligent analysis of a large volume of online CRs attracts interest from researchers in various fields. However, many research studies only concern sentiment polarity in the product feature level. With these results, designers still need to read a list of reviews to absorb comprehensive CRs. In this research, online reviews are analyzed at a fine-grained level. In particular, aspects of product features and detailed reasons of consumers are extracted from online reviews to inform designers regarding what leads to unsatisfied opinions. This research starts from the identification of product features and the sentiment analysis with the help of pros and cons reviews. Next, the approach of conditional random fields is employed to detect aspects of product features and detailed reasons from online reviews jointly. In addition, a co-clustering algorithm is devised to group similar aspects and reasons to provide a concise description about CRs. Finally, utilizing customer reviews of six mobiles in Amazon.com, a case study is presented to illustrate how the proposed approaches benefit product designers in the elicitation of CRs by the analysis of online opinion data.

1. Introduction

The rapid development of information and communication technology encourages an increasing number of more consumers to shop online on such websites like JD.com and Amazon.com. According to a news report in Forbes,[1] the total transaction of Alibaba Group Holding Ltd. reported a record of USD 9.3 billion in sales on November 11, 2014. A large volume and variety of consumer data are generated constantly online, including customer search logs, purchase behaviors, and customer reviews, which provide helpful information for potential consumers and product designers.

An exemplary category of consumer data, online reviews contain valuable customer requirements (CRs). These reviews help designers understand CRs, which alleviate them from performing time-consuming investigations. In the field of computer science, opinion mining is a trendy research topic. Two popular research problems are how to extract product features and how to identify sentiment polarity in the product feature level from textual data (Hu and Liu, 2004; Moghaddam and Ester, 2012). The primary concerns of understanding the problems are which product features are mentioned and is the opinion positive, negative or neutral. Consider the following three review sentences of the Nokia N8 smart phone in Amazon.com for example.

S1[2]: "The battery life is a horrible."
S2[3]: "The on-board battery meter can be misleading."
S3[4]: "My only complaint here is that the battery is difficult to remove."

Using sentiment classification techniques, opinionated sentences on a specific product feature can be identified. In this example, most opinion mining techniques are capable of recognizing that consumers are writing negative opinions regarding the battery of a mobile phone. However, despite these results, designers still need to consolidate all of the reviews and read these sentences consecutively to understand customer expectations clearly. For example, various aspects of the battery are critiqued in S1–S3 and these details provide more

* Corresponding author at: Department of Information Management, School of Government, Beijing Normal University, Beijing 100875, China. Tel: +86 10 5880 5387.
E-mail address: jinjian.jay@bnu.edu.cn (J. Jin).

[1] http://www.forbes.com/sites/hengshao/2014/11/11/9-3-billion-sales-recorded-in-alibabas-24-hour-online-sale-beating-target-by-15/.

[2] http://www.amazon.com/review/RSZY6GDKAWEGL/ref=cm_cr_rdp_perm.
[3] http://www.amazon.com/review/R1FU6HHDLLE9MU/ref=cm_cr_rdp_perm.
[4] http://www.amazon.com/review/R2QQ2YOO7HU54T/ref=cm_cr_rdp_perm.

http://dx.doi.org/10.1016/j.engappai.2015.05.006

instructive suggestions to designers. However, many approaches in opinion mining fail to differentiate aspects of the battery and instead note the detailed reasons associated what makes consumers unsatisfied.

Accordingly, the review analysis is conducted at a fine-grained level to explore the reasons why consumers are unsatisfied with products. To obtain valuable CRs, in this research, product features and sentiment polarity are first identified from online reviews with the help of pros and cons reviews. Based on the sentimental information, an approach based on conditional random fields (CRFs) is developed to label each word in online reviews. These tags help to discern different aspects of product features, as well as detailed reasons written by consumers automatically. Moreover, a co-clustering algorithm is devised to cluster aspects of product features and reasons jointly with the help of their inter-relations, which aims to provide a brief description about CRs.

The rest of this research is organized as follows. In Section 2, relevant work in summarization of online opinion data and how online opinion data are utilized by designers are briefly reviewed. Next, the problem to be studied is clarified in Section 3. In Section 4, the technical details of the proposed approaches are explained. Section 5 presents experimental results to show how these approaches benefit product designers. Section 6 concludes this research.

2. Related work

2.1. Summarization of online opinion data

A review summarization framework was proposed at a sentence level by Zhuang et al. (2006). Opinions were captured from the expansion of seed words from the WordNet. Next, dependency relation templates were utilized to detect feature-opinion pairs. Finally, organized sentences were recognized as the review summary. Another summarization approach was reported to analyze the topic structure of online reviews by Zhan et al. (2009). In this approach, important topics were extracted and aggregated from online reviews. The final summary of reviews was clustered by the topic structure and different clusters were ranked according to the topic importance. A probabilistic mixture model was initially proposed to analyze topics and sentiments in online reviews by Mei et al. (2007). In this model, a document was considered to be generated by background words and other words, which were generated from one of many subtopics. Next, a sentiment word was utilized to describe the topic. Finally, a HMM (Hidden Markov Model) model was employed to analyze the dynamic change of the sentiments in online reviews.

The CRFs approach is widely utilized to summarize online reviews. Jakob and Gurevych (2010) utilized tokens, POS (Part of Speech) tags, short dependency paths, word distances between opinion words and other features to define features of online reviews. With these features, the approach of CRFs was to detect opinion targets in online reviews. Chen and Qi (2011) conducted a comprehensive user study examining the reasons that may lead consumers to reach final decisions. These researchers claimed that both static features and social features of products impacted consumers' decisions. The authors argued that there are three stages when consumers make decisions, (1) to filter alternatives and select ones for in-depth study, (2) to view the product's details and save it in a favorite list and (3) to compare candidates and make final decisions. This study also confirmed the previous argument that “when information about an object or firms comes through the opinions of another person, negative information can be credible and generalizable than positive information” (Mizerski, 1982). Accordingly, a framework using the approach of CRFs was built to identify sentiment polarity of product features by tagging sentimental words of product features. Additionally, four types of CRFs models were compared to identify product features and related opinion words (Li et al., 2010a), which include the linear CRFs, the skip-chain CRFs, the tree CRFs and the skip-tree CRFs. Moreover, some researchers proposed an opinion summarization for Bengali news articles (Das and Bandyopadhyay, 2010). In this approach, an SVM classifier was utilized to identify subjective sentences. Next, a model of CRFs was utilized to recognize theme words. Finally, sentences were clustered together according to their cosine similarity, and a Page Rank algorithm selected representative sentences for each cluster.

In addition to online reviews, the summarization of other opinion data is also investigated. An opinion summarization system for tweets was proposed by Meng et al. (2012). In this summarization system, a hashtag graph was built, and the relatedness between two hashtags was calculated by their concurrent relation, their contextual similarity and their topic-aware similarity. Next, an algorithm for the affinity propagation was employed to cluster hashtags, which were regarded as topics. A pattern based method was then employed to identify insightful tweets. The opinions conveyed by tweets were identified through a lexicon-based method. Finally, an optimization problem was formulated to select representative tweets. To cluster reader comments in a news article, two probabilistic graph models were compared by Ma et al. (2012). In these models, news articles and reader comments were regarded as master documents and slave documents, respectively. In the first model, topics in reader comments were confined to the topics in the news article. In the second model, topics in reader comments were derived from topics in news articles and all comments themselves. Additionally, in these models, representative sentences were selected by the approach of Maximal Marginal Relevance.

2.2. Online opinion data for product design

A framework was presented by Decker and Trusov (2010) to aggregate CRs from online reviews for product design. This framework was utilized to infer the relative effect of product features and effect of different brands on overall customer satisfaction. A system that monitors customer opinions from textual data was built by Goorha and Ungar (2010). First, frequent phrases and phrases near the terms of interest were extracted from textual data. These phrases were then utilized to identify which of them will emerge dramatically. Additionally, to determine whether a phrase was interesting depended on the frequency to which it was referred, its previous referred frequency and the level of specificity at which it refers to a topic. In this system, to present the results in an interactive user interface effectively, TFIDF weights and the cosine similarity method were utilized to cluster relevant terms. Several categories of dimensions that relate to the usability and the user experience were defined by Hedegaard and Simonsen (2013). According to such criteria, review sentences were manually labeled. A SVM-based method was then utilized to classify review sentences into the category of usability and the category of user experience.

Notably, one objective of CRs extraction and aggregation is to make new products to be utilized by potential customers. Accordingly, Miao et al. (2013) proposed to identify opinion leaders in a specific domain. Observing the fact that customers post several reviews and that the reviews may belong to different domains; accordingly, in this approach, the number of reviews in the same domain was utilized to define the similarity of consumers. Additionally, consumers might present different interests on different product aspects. To cluster consumers with similar interests, a permutation-based structural topic model was proposed by Si et al. (2013). By using this model, the frequency of different product aspects and the occurrence ordering were presented. Additionally, a feature-based product ranking method was developed by Zhang et al. (2010) based on consumers' diversified CRs. Subjective and comparative sentences were identified from online reviews in this method. A weighted directed graph was then built to model product relationships. With this graph, products were compared at the feature level.

To capture the rapid change of CRs, a two-stage hierarchical process was built from online reviews (Lee, 2007). At the first stage, related product attributes and CRs were clustered into hyper-edges by an association rule algorithm. At the second stage, hyper-rules were applied on hyper-edges to track CRs. Similarly, online reviews were utilized to predict product design trends (Tucker and Kim, 2011). The sentiment polarity in the feature level was extracted from product reviews. Next, the Holt-Winters exponential smoothing method was used to model product preference trends. Additionally, several researchers utilized online reviews to predict the rank of products in the near future. The affinity rank history, average ratings and affinity evolution distances were extracted from review sites (Li et al., 2010b). Next, the AutoRegressive model with exogenous inputs was applied to forecast the rank of products.

To make online reviews directly usable by product designers, finding how to prioritize engineering characteristics from online reviews was investigated (Jin et al., 2012). The rating values of online reviews were regarded as the overall customer satisfaction. The sentiment polarity over engineering characteristics were utilized as features. Next, an ordinal classification approach was proposed to prioritize engineering characteristics for designers. Additionally, a three-step method was proposed for customer-driven product design selection using online reviews (Wang et al., 2011). In the first step, product attributes were extracted from online reviews. In the second step, a hierarchical customer preference model was built by using Bayesian linear regression. Product ratings, category ratings, attribute ratings and product specifications were taken into consideration in this hierarchical model. Finally, in the last step, an optimization problem was formulated to maximize potential profit by taking engineering constraints into considerations. There are few studies concerning the quality of online reviews from the perspective of designers. The helpfulness of online reviews was initially defined from the perspective of designers (Liu et al., 2013). According to designers' arguments, four categories of features were extracted from online reviews. With these features, the helpfulness of online reviews was inferred by using a regression method. Additionally, without domain-dependent features, it is found that there was no significant loss in terms of the helpfulness prediction.

Online opinion data are also utilized to make comparisons between products. Utilizing online reviews, researchers extracted product names and product features using CRFs (Netzer et al., 2012). With the proposed approaches, two applications were described to compare products. A graph propagation method was also proposed to compare products by using online reviews and community-based question answering by Li et al. (2011). In this method, comparative sentences were extracted from reviews. Next, weights of product pairs in the graph for information propagation were defined according to the number of options between products. A product comparison network was also created by exploiting comparative sentences in online reviews (Zhang et al., 2013). A single-link graph, a dichotomic-link graph and a multi-link graph were built according to the overall sentiments. Additionally, different regression models were utilized to analyze factors that influence a product's rank.

3. Problem statement

To understand what makes consumers unsatisfied with products, product features and corresponding sentiment polarity are extracted from online reviews. However, the review analysis in the feature level does not provide sufficient information. As presented in the examples of the three review sentences in Section 1, the battery life and the battery meter are different aspects of the battery. Automatically identifying different aspects of product features from online reviews helps to clarify CRs for designers further. In addition, consumers may offer detailed reasons to support their arguments in online reviews. The analysis of such information is also helpful for designers to comprehend CRs directly without reading all relevant online reviews.

Accordingly, four types of messages should be exploited from online reviews. Generally, let Q= < F, S, A, R > be a quadruple, which includes product features (F), sentiment polarity (S), aspects of product features (A) and detailed reasons (R). Specially, an aspect refers to one property or parameter of the product feature, such as battery life, battery indicator, screen size, screen sensitivity, etc. The detailed reasons refer to the explanations that are written by consumers to support their arguments regarding their sentiment polarity.

Take the three sentences in Section 1 for example. In the first sentence, a negative polarity is presented on the battery life and no extra details are provided. Then, the extracted quadruple will be Q= < "battery", "negative", "life", NULL > . In the fourth element of this quadruple, NULL denotes that no additional reasons are provided by the consumer. In the second sentence, the consumer complaints about the battery meter and the consumer utilizes the word "misleading" to support his/her argument. Hence, the quadruple of the second sentence is Q= < "battery", "negative", "meter", "misleading" > . For the third sentence, the complaint is concerning the difficulty of removing the battery. Thus, the quadruple is Q= < "battery", "negative", NULL, "remove" > .

Specially, in this research, the problem to be investigated is the extraction of four types of information from online reviews to help designers on understanding CRs. Four tasks need to be conducted, including (1) to extract product features, (2) to identify sentiment polarity, (3) to recognize aspects of product features and (4) to discern detailed reasons of consumers.

4. Methodology

4.1. A framework for the identification of feature aspects and consumer reasons

In Fig. 1, a framework for the identification of product feature aspects and consumer detailed reasons from online reviews is presented.

As seen from this figure, POS tags are initially obtained from online reviews. These tags are utilized for the identification of product features and the analysis of sentiment polarity. In this study, with the help of pros and cons reviews, a supervised learning approach is devised to identify product features and analyze the corresponding sentiment polarity from online reviews. By utilizing this approach, product feature related review sentences can be categorized according to the sentiment polarity.

Note that, concentrations of this study are to recognize aspects of product features and to discern detailed reasons of consumers from a large number of online concerns. Specially, in this research, a two-phase approach was suggested. In this first phase, a CRFs model was trained for the recognition of feature aspects and consumer detailed reasons from each product feature related sentences. In this CRFs based model, two categories of features were utilized. The node features characterized the word related information and three types of node features were extracted to describe each word. The edge features were employed to depict

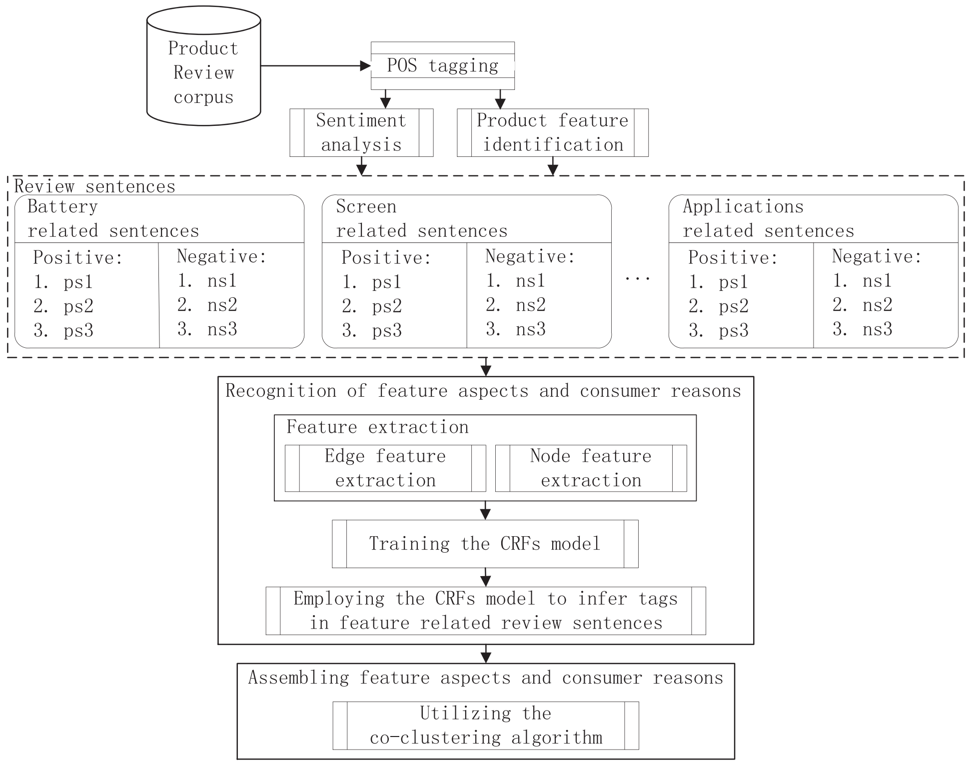


Fig. 1. A framework for the identification of product feature aspects and consumer detailed reasons from online reviews.

the association between hidden tags and the association between a hidden tag and a visual word and two types of edge features are defined to explain these associations. Next, the trained CRFs model was utilized to infer a quadruple $Q = <F, S, A, R>$ from each product feature related sentence. All of these technical details are explained in Section 4.3. In the second phase, a co-clustering algorithm was proposed. This algorithm was to jointly assemble all recognized similar product feature aspects and consumer detailed reasons in product feature related sentences. This co-clustering algorithm provides a brief summary towards consumers' concerns on different aspects of product features.

4.2. Extracting product features and identifying sentiment polarity

In the area of opinion mining, different approaches are proposed to extract product features and identify sentiment polarity. Many state-of-the-art approaches are developed by researchers in computer science (Ding et al., 2008; Lin and He, 2009; Zhai et al., 2010). However, some of these approaches are slightly difficult for data practitioners in other fields to understand and implement. Additionally, similar to other different types of supervised learning approaches, a large number of labeled data are required for model training, which is generally time-consuming to obtain. In this research, the extraction of product features and the identification of sentiment polarity are conducted with the help of pros and cons reviews. Similar approaches for extracting features and identifying customer sentiments are also elaborated upon in the literature cited in Kim and Hovy (2006) and Yu et al. (2011).

4.2.1. Extracting product features

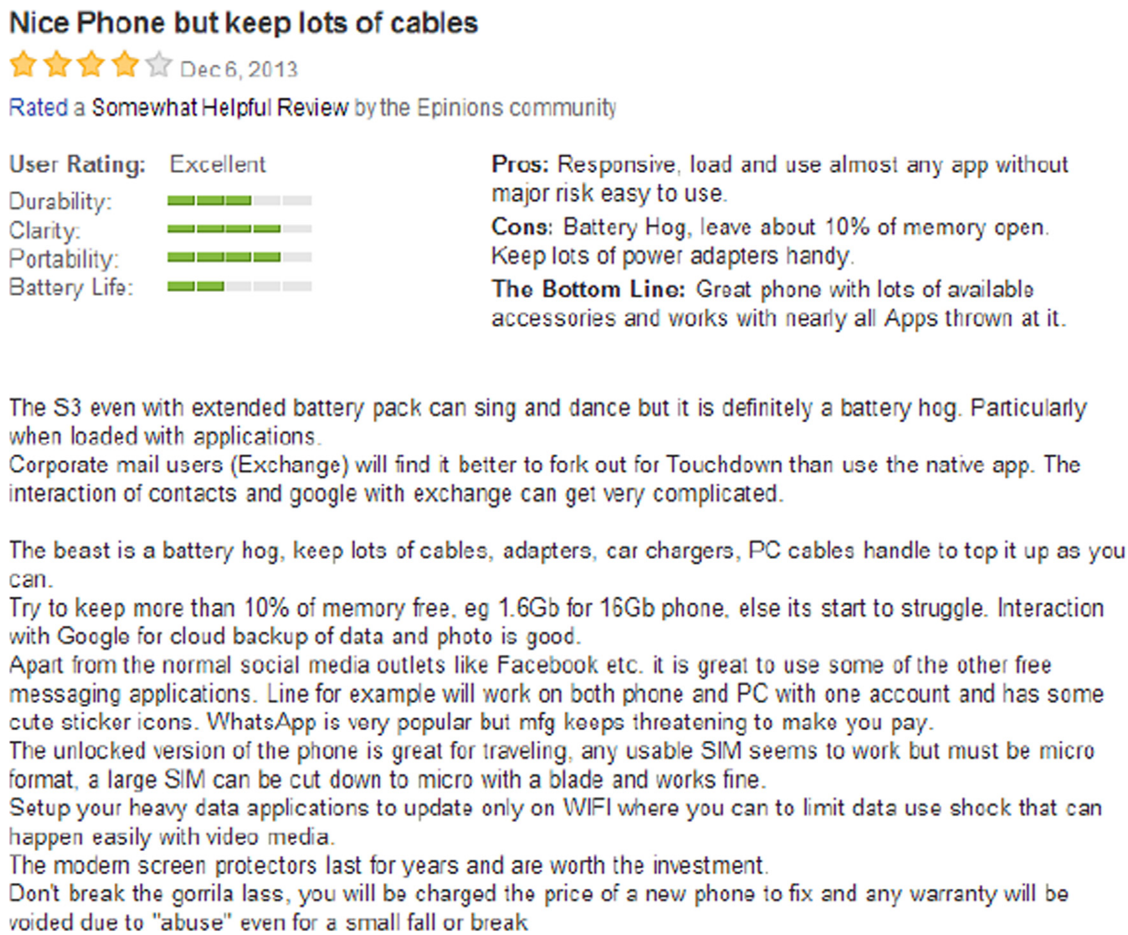
In this research, pros and cons reviews are utilized to extract product features. A representative pros and cons review is shown in Fig. 2. As seen from this figure, product features are highlighted clearly in the pros and cons lists. These lists provide valuable training data and help to extract product features from reviews. Accordingly, with the POS tagging conducted on reviews, the frequently mentioned nouns or noun phrases in the pros and cons list are used to identify product features.

Additionally, different words are employed by different consumers to describe the same product features. For instance, both "memory" and "storage" are often utilized interchangeably to denote the same product feature. With the help of WordNet, synonyms of nouns are extracted and these synonyms are clustered to identify the same product feature. To avoid imprecise expanded synonyms of words, in this research, only synonyms within two WordNet distances are considered. In addition, abbreviated forms of words are often utilized in reviews. For instance, "apps" is used to denote the "applications" of mobiles. However, these words are seldom defined in WordNet or other thesauruses. Thus, some manually defined rules are made to supplement synonyms from the WordNet expansion. Finally, product features are extracted from online reviews and clustered by using pros and cons reviews.

4.2.2. Identifying sentiment polarity

In this research, it is assumed that this consumer holds a neutral sentiment if one sentence has an objective opinion toward product features. Accordingly, two subtasks need to be conducted for the identification of sentiment polarity. One is to know whether a subjective or objective opinion is expressed. The other is to identify whether consumers hold a positive or negative opinion.

In 2004, Pang and Lee (2004) built a publicly available subjective dataset, which included 5000 subjective and 5000 objective sentences. Hence, for the first subtask, with the help of this dataset and denoting each sentence as a bag of words (BOWs), a binary Naive Bayes classifier can be constructed to distinguish subjective and objective sentences in online reviews.


Nice Phone but keep lots of cables

Dec 6, 2013

Rated a Somewhat Helpful Review by the Epinions community

User Rating: Excellent

Durability:

Clarity:

Portability:

Battery Life:

Pros: Responsive, load and use almost any app without major risk easy to use.

Cons: Battery Hog, leave about 10% of memory open. Keep lots of power adapters handy.

The Bottom Line: Great phone with lots of available accessories and works with nearly all Apps thrown at it.

The S3 even with extended battery pack can sing and dance but it is definitely a battery hog. Particularly when loaded with applications.
Corporate mail users (Exchange) will find it better to fork out for Touchdown than use the native app. The interaction of contacts and google with exchange can get very complicated.

The beast is a battery hog, keep lots of cables, adapters, car chargers, PC cables handle to top it up as you can.
Try to keep more than 10% of memory free. eg 1.6Gb for 16Gb phone, else its start to struggle. Interaction with Google for cloud backup of data and photo is good.
Apart from the normal social media outlets like Facebook etc. it is great to use some of the other free messaging applications. Line for example will work on both phone and PC with one account and has some cute sticker icons. WhatsApp is very popular but mfg keeps threatening to make you pay.
The unlocked version of the phone is great for traveling, any usable SIM seems to work but must be micro format, a large SIM can be cut down to micro with a blade and works fine.
Setup your heavy data applications to update only on WIFI where you can to limit data use shock that can happen easily with video media.
The modem screen protectors last for years and are worth the investment.
Don't break the gorrila lass, you will be charged the price of a new phone to fix and any warranty will be voided due to "abuse" even for a small fall or break

Fig. 2. One typical pros and cons review.

If a sentence is predicted to be subjective, the next subtask is to judge whether it is positive or negative. Notice that besides the product features listed in pros and cons reviews, the positive and negative sentimental information about features is also provided. Then, utilizing sentences with sentiment polarity as training data, a binary classifier is built to discriminate whether a positive or negative opinion is expressed in each review sentence. To build this sentiment classifier, rather than a BOWs representation, sentimental terms are considered in review sentences. Notably, in this research, each sentence is represented by sentimental terms in the subjective lexicon provided by the MPQA project (Wilson et al., 2005). Using sentences with sentiment polarity as training data and MPQA representation, a binary classifier is built to discern the sentiment polarity of product features.

4.3. Recognizing aspects of product features and detailed reasons of consumers

Given extracted product features and sentiment polarity, the next critical task is to recognize different aspects of product features and the reasons of consumers from online reviews.

As shown in Section 3, different aspects of product features are usually described with individual nouns or noun phrases. Additionally, reasons are often reported with the least number of words that support consumers' arguments about their sentiment polarity. Correspondingly, the extraction of aspects and reasons requires identifying informative expressions or a sequence of words that are evaluated as properties of features or arguments to back up consumers' opinions. In this research, an approach of CRFs is utilized to generate a label for each word in the review sentences, which help to identify a sequence of words as aspects of features or reasons of consumers. CRFs is a discriminative probabilistic graphical model, which involves types of contextual node features and edge features. It is often utilized to model the relationship between observations and generate sequences of labels for structured prediction in, for example, natural language processing and biological sequences mining.

4.3.1. Labeling scheme

Three types of expressions are required to be identified from online reviews, including product features, aspects of product features and reasons of consumers. As noted, these expressions can be words or a sequence of words. In this research, the conventional BIO encoding for tag representation is utilized to label each sentence containing a product feature. Specially, product features are represented as “F”s and aspects of product features are “A”s. For a sequence of words that describes reasons, “RB” is to represent the beginning of one reason and “RI” is for the inside of the reason. All other words and punctuations are then labeled as “O”s. Finally, in total, five tags are utilized to denote each word in a review sentence. Following this labeling scheme, the previous three sentences in Section 1 can be denoted as follows:

S1: The/O battery/F life/A is/O a/O nightmare/O./O
S2: The/O on-board/O battery/F meter/A can/O be/O misleading/O./O
S3: My/O only/O complaint/O here/O is/O that/O the/O battery/F is/O difficult/O to/O remove/RB./O

Other two complex examples are also shown as follows:
S4: I/O might/Oget/O 2/RB days/RI of/O battery/F usage/A before/O I/O charge/RB it/O./O
S5: Using/O features/O such/O as/O playing/RB games/RI or/O watching/RB movies/RI can/Oreally/O drain/RB the/O battery/F life/A./O

Now, to recognize aspects of product features and to discern detailed reasons of consumers are jointly modeled as a tagging problem on review sentences and the approach of CRFs is utilized to predict the tag for each word.

4.3.2. Feature extraction

In this subsection, features extracted from online reviews for tag labeling in CRFs are described. Notably in CRFs, both node features and edge features are considered.

All node features are listed in Table 1. Position features are to judge the position of the current token. In considerations of the current token, the previous token and the following token, six types of word features are utilized, including original form of token, lemma of the token, token category, POS tag, prefixes and suffixes. The token category is considered as one type of word features, which is determined according to the shape of a token. In this research, the 19 categories in LingPipe are utilized, which include whether it is a single digit, whether it contains letter only, whether it is a single uppercase letter, etc. Moreover, four types of prefixes and suffixes are also extracted. Four types of prefixes and suffixes include sequences containing one, two, three and four characters, respectively. Finally, four types of dependency relation features are extracted regarding the current token. Dependency relation features provide detailed information about the parsed dependency tree. In this research, Stanford typed dependencies representation is utilized. This representation is often employed to analyze grammatical relationships in a sentence. In this representation, a triple denotes the relationship between a pair of words, which involves the name of this relation, the governor of this relation and the dependent of this relation.

For edge features, they are presented in Table 2. The aim of tag features provides the tag and category information of the previous token. Similarly, dependency relation features define the tag and the name about the relation for the case in which a token is involved as a governor or a dependent of one grammatical relationship.

4.4. A co-clustering algorithm for aspects of features and consumer reasons

In the previous section, product features, sentiment polarity, aspects of product features and detailed reasons of consumers are exploited from online reviews by using the approach of CRFs.

One observation is that consumers tend to employ different words to describe similar detailed reasons. For instance, different words are used to complain about the battery life, such as “a few hours”, “less

Table 1
Node features.

Position features	Is the first token
	Is the last token
	Is not the first token or the last token
	The position of the current token
Word features	Current token
	Previous token
	Next token
	Lemma of the current token
	Lemma of the previous token
	Lemma of the next token
	Token category of the current token
	Token category of the previous token
	Token category of the next token
	Current POS tag
	Previous POS tag
	Next POS tag
	Prefixes of the current token
	Prefixes of the previous token
	Prefixes of the next token
	Suffixes of the current token
	Suffixes of the previous token
	Suffixes of the next token
Dependency relation features	Is a governor
	Name of relation if it is a governor
	Token of the dependent if it is a governor
	Name of relation and token of the dependent if this node is a governor
	Is a dependent
	Name of relation if it is a dependent
	Token of the governor if it is a dependent
	Name of relation and token of the governor if this node is a dependent

Table 2
Edge features.

Tag features	Tag of the previous token
	Token category and the tag of the previous token
Dependency relation features	Tag of the dependent if it is a governor
	Name of relation and tag of the dependent if the node is a governor
	Tag of the governor if it is a dependent
	Name of relation and tag of the governor if the node is a dependent

than a day" and "drain". In this case, with the previous results from the approach of CRFs, designers still need to read a bundle of similar extracted reasons one by one. A more instructive result is to provide a concise description regarding CRs, which cluster similar reasons, and it will clearly enable product designers to comprehend CRs efficiently. A similar scenario is that different words are also employed to describe the same aspect of product features. Thus, it is necessary to cluster aspects of features and consumer reasons for product designers.

A simple approach is to cluster consumer reasons by using the WordNet distance. For instance, two reasons are expected to be grouped into the same cluster if the WordNet distance between reasons is small. However, this approach cannot be applied directly. As noted in the previous observations, both phrases and words are utilized by consumers to describe reasons. Because the distance between phrases cannot be evaluated by WordNet, the WordNet approach fails to be applied to cluster reasons. Additionally, this approach cannot handle the previous use case where "few hour", "less than a day" and "drain" are employed interchangeably in different sentences, although consumers are complaining about the battery life.

In this research, a co-clustering algorithm is proposed to jointly cluster reasons of consumers and aspects of product features. The intuition behind this algorithm is that similar words that describe reasons are often derived from the same aspect and, likewise, the same aspect often derives to similar words. For example, the followings are four typical sentences with labeled tags to describe the battery of one mobile.

S1: The/O on-board/O battery/F meter/A can/O be/O misleading/RB
S2: The/O battery/F indicator/A is/O deceptive/RB
S3: It/O can/O really/O KO/RB the/O battery/F life/A
S4: The/O battery/F life/A is/O not/O what/O promised/O no/O matter/O if/O you/O find/O a/O way/O to/O charge/RB it/O

In S1 and S2, it can be concluded that "meter" and "indicator" are the same aspect because the reasons of "misleading" and "deceptive" are semantically similar. Additionally, in S3 and S4, "KO" and "charge" are regarded to be similar because both words point to the battery life. Accordingly, these detailed reasons are utilized to cluster the corresponding aspects and, likewise, the aspects of product features are utilized to cluster the corresponding reasons. The details are described in Algorithm 1.

Algorithm 1. A co-clustering algorithm for aspects of features and consumer reasons

Input:
Map of aspects $A_s \leftarrow <\text{Aspect}, <\text{Document ID}>>$
Map of reasons $R_s \leftarrow <\text{Reason}, <\text{Document ID}>>$
Output:
Clusters of aspects C_a, cluster of consumer reasons C_r
Steps:
1 DO
2 FOR any two reasons R_a and R_b in R_s
3 IF R_a and R_b are similar
4 $IDs(R_a) \leftarrow$ Extract all document ID of R_a
5 $IDs(R_b) \leftarrow$ Extract all document ID of R_b
6 $As(IDs(R_a)) \leftarrow$ Extract all aspects of $IDs(R_a)$
7 $As(IDs(R_b)) \leftarrow$ Extract all aspects of $IDs(R_b)$
8 $A_s \leftarrow$ Group $As(IDs(R_a))$ and $As(IDs(R_b))$
9 END IF
10 END FOR
11 $C_a \leftarrow$ Extract clusters of aspects from A_s
12 IF A_s changed
13 FOR any two aspects A_p and A_q in A_s
14 IF A_p and A_q are similar
15 $IDs(A_p) \leftarrow$ Extract all document ID of A_p
16 $IDs(A_q) \leftarrow$ Extract all document ID of A_q
17 $Rs(IDs(A_p)) \leftarrow$ Extract all reasons of $IDs(A_p)$
18 $Rs(IDs(A_q)) \leftarrow$ Extract all reasons of $IDs(A_q)$
19 $R_s \leftarrow$ Group $Rs(IDs(A_p))$ and $Rs(IDs(A_q))$
20 END IF
21 END FOR
22 END IF
23 $C_r \leftarrow$ Extract clusters of reasons from R_s
24 While A_s or R_s changed, return to Line 1

The objective of Algorithm 1 is to cluster documents according to both aspects of product features and reasons of consumers. First, if two reasons are semantically similar, the corresponding documents are extracted (line 3–5). Aspects of product features that connect to these documents are considered to be semantically similar and these aspects are put into the same cluster (line 6–8). Likewise, if two aspects are semantically similar, according to the referred documents, the corresponding reasons are extracted and they are grouped into the same cluster (line 12–20).

5. Case study

5.1. Data preparation

To clarify how the proposed approach will help product designers to identify CRs efficiently, a case study is shown in this section.

In this case study, 475 pros and cons reviews of 13 smart phones were collected from Epinions.com. These pros and cons reviews are utilized as training data to extract product features and identify the sentiment polarity from reviews in a general format. Specially, 5730 reviews of six mobile phones in Amazon.com are utilized. In consideration of data privacy, the names of these products are represented as P_1, P_2, P_3, P_4, P_5 and P_6. The number of reviews is shown in Table 3 and some statistics about these reviews are listed in Table 4.

On the average, there exist 142.16 words in each review. However, they are not distributed evenly, with the maximum of 3379 words in a single review. A similar phenomenon is also found in terms of the sentence number per review, with an average 7.32 and a maximum 239.

For the tagging task to identify aspects of product features and reasons of consumers, three annotators were hired to label reviews manually. Aspects of product features and reasons of consumers need to be labeled. The labeling scheme is introduced in Section 4.2.1. Each review was labeled by two annotators. Conflicts of review labeling were examined and determined by the third annotator.

5.2. Product feature extraction and sentiment identification

The techniques about how to extract product features and identify sentiment polarity from online reviews are explained in Section 4.1. Using these techniques, experimental results are presented in this subsection.

Table 3
The number of reviews.

Name	P_1	P_2	P_3	P_4	P_5	P_6
Num. of reviews	1108	784	880	965	905	1088

The top five frequently discussed features of 5730 mobile reviews are shown in Table 5. As seen from this table, screen, application and battery become hot features. Generally, all of six selected products are smart phones. Perhaps the first impression of a smart phone is of its screen. Indeed, a large and clear screen makes it attractive to consumers. Additionally, applications in the phone and its operating system are other critical factors that consumers use to make purchasing decisions. For instance, comparisons of applications in different operating systems, such as Android or Windows, affect decisions that consumers make on their selection of brands and product models. Moreover, if one mobile phone's battery causes consumers to have to charge regularly, it can be expected that it will result in negative comments.

Accordingly, the battery and the screen are chosen as two exemplary product features. The objective is to exemplify the extraction of aspects and the identification of reasons that lead consumers to provide a negative sentiment. In Table 6, there are some statistics concerning about the screen and the battery of six products.

Compared with Table 5, more than 35% consumers prefer to talk about either of these two product features. However, the sentiment polarity of consumers towards two features are not the same. As seen from Table 6, the percentage of negative reviews holds an indispensable share. To investigate what make consumers unsatisfied regarding a specific product in feature level, designers still need to read these negative opinions and explore reasons for understanding CRs.

In Table 7, several exemplary reviews of complaints for the battery and the screen are sampled from the negative reviews of six products. Consider the screen, for instance. Different aspects of the screen are mentioned, including response time, automatic locking, ease of scratched, size, resolution, and rotation. This finding further confirms the argument that it is the different aspects of product features that lead to negative opinions and the detailed reasons provide several instructive suggestions to product designers.

5.3. *Exploration of aspects of product features and reasons of consumers*

In Section 4.2, a model based on CRFs is developed, which mainly aims to recognize aspects of product features and detailed reasons of consumers from online reviews. To evaluate the performance of the proposed method, in this case study, review sentences that are talking about the battery life and the screen of P_1, P_2 and P_3 are examined.

Table 4
Statistics of 5719 reviews of six mobile phone reviews in Amazon.com.

Sentences	0–10	11–20	21–30	31–40	41–50	51–60	61+
Num.	4514	751	249	93	45	28	28
Words	0–100	101–200	201–300	301–400	401–500	501–600	601+
Num.	3572	1027	415	234	151	85	224

Table 5
Top five frequently discussed features of 5730 mobile phone reviews.

Top referred product features	% of reviews referred features
Cover screen screens	24.41%
App application applications apps	18.38%
Batteries battery	17.96%
Android androids os symbian window windows	17.68%
Internet net network networks web wi-fi wifi	15.70%

5.3.1. *Evaluation metrics*

Five widely utilized classification evaluation metrics are employed, including precision, recall, F_1, true positive rate (TP Rate) and false positive rate (FP Rate). Precision and recall are generally defined according to the measure of relevance. Precision is the fraction of retrieved instances that are relevant. Precision is denoted as follows:

$$\text{Precision} = \frac{|\{\text{relevant documents}\}| \cap |\{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

Recall is the fraction of relevant instances that are retrieved,

$$\text{Recall} = \frac{|\{\text{relevant documents}\}| \cap |\{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}$$

F_1 is an evaluation metric that combines both precision and recall, which is the harmonic mean of precision and recall.

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

True positive rate (TP Rate, also called sensitivity in some research fields) evaluates the fraction of actual relevant instances that are correctly identified as such. It usually equals to the value of Recall.

False positive rate measures the proportion of non-relevant documents that are retrieved, out of all non-relevant documents available, which is denoted as follows:

$$\text{FPRate} = \frac{|\{\text{non} - \text{relevant documents}\}| \cap |\{\text{retrieved documents}\}|}{|\{\text{non} - \text{relevant documents}\}|}$$

5.3.2. *Evaluation results*

The following results are reported on the average of three-fold cross evaluations over products. Specially, it takes reviews of two products as training data and reviews of the third product as testing data.

The performance of the proposed approach, based on CRFs, to recognize aspects and reasons about the battery and the screen are listed in Tables 8 and 9, respectively. As seen from the two tables, words that refer to product features are recognized accurately in both datasets.

In terms of the identification of aspects, a relatively higher performance is achieved in the battery dataset. As noted, words that describe aspects of battery tend to be limited. For instance, these phrases include “battery life”, “battery charging”, “battery drain”, “battery alert” and “battery maintenance”. However, for phrases that describe a screen, they tend to be fuzzy and include “screen space”, “screen lock”, “screen size”, and “screen display”. In addition, the complexity of phrase structures that are utilized to describe aspects of two features is also different. For example, in “life of battery”, it is relatively easy to infer that the word “life” tends to be one aspect of the battery. Similarly, “size” in “the size of screen” can be deduced as one aspect of the screen. However, in another case, “corner” in “the right corner of screen” is not an aspect, although this phrase has a similar structure. The diversity of words and the complexity of structures make it somewhat difficult to detect aspects accurately.

For the identification of the detailed reasons, compared with the recall, a higher precision is obtained in both datasets. A wild range of words are utilized to describe the detailed reasons, which make them tend to be quite literally dissimilar with each other, although the same aspect of product features is pointed to. Additionally, occasionally, complaints of consumers are twisted with different product features. For instance, one consumer complaint said that the “QUERTY keyboard is too small for efficient usability, Camera is

Table 6
Statistics about the screen and the battery of six products.

Product	Feature	Num. of reviews referred the feature	% of reviews referred the feature	% of negative towards the feature
P_1	Screen	324	29.24%	51.23%
	Battery	105	9.48%	50.48%
P_2	Screen	218	27.81%	44.95%
	Battery	206	26.27%	56.31%
P_3	Screen	211	23.98%	50.24%
	Battery	172	19.55%	40.12%
P_4	Screen	188	19.71%	42.55%
	Battery	149	15.62%	48.99%
P_5	Screen	165	18.23%	45.45%
	Battery	154	17.02%	50.65%
P_6	Screen	290	26.65%	24.48%
	Battery	241	22.15%	33.61%

Table 7
Some exemplary negative reviews concerning the screen and the battery.

Screen	1. The touch screen is so slow and so hard to use.
	2. The biggest problem with this phone is that the screen automatically locks after you dial a phone number to call.
	3. Plastic touch screen can be easily scratched, use caution in pocket with keys or coins.
	4. The screen is so small, that even with my reading glasses, I have a horrible time reading anything.
	5. Resolution of the screen again is not as good as the iphone.
	6. Sometimes when I flip the phone, the screen will not rotate.
Battery	1. Biggest complaint about this phone is it is ridiculously terrible battery life.
	2. They thought the battery was too thick.
	3. Some of the more advanced features are particularly battery draining.
	4. Had problems with the battery after a solid year of use, which I had to replace but it was not expensive (16 bucks) .
	5. In my honest opinion, it looked like a cheap battery.
	6. I am also extremely disappointed in the fact that the battery is embeded and therefore not replaceable.

Table 8
Aspects and reasons identification over reviews about the battery.

Battery	Feature	Aspect	Reason
Precision	1.000	0.936	0.787
Recall	0.998	0.854	0.453
F_1	0.999	0.893	0.575
TP Rate	0.998	0.854	0.453
FP Rate	0	0.001	0.007

Table 9
Aspects and reasons identification over reviews about the screen.

Screen	Feature	Aspect	Reason
Precision	0.990	0.580	0.569
Recall	0.985	0.444	0.306
F_1	0.988	0.503	0.398
TP Rate	0.985	0.444	0.306
FP Rate	0	0.002	0.004

difficult can't see the icons on the screen". Generally, these complaints describe that the icon formatting on the screen makes it difficult for consumers to use the phone, although the screen of this mobile is mentioned. Another case is that some consumers explain the whole scenario about what makes him/her unsatisfied, such as "from having the phone in your pocket, buttons are pressed and it turns the screen on, which uses some of your battery". This example is describing the sensitivity of the screen, though no specific words are employed to describe the aspects of the product features directly. All of these complex structures make it difficult to identify reasons at a high recall. It leaves some space to develop sophisticated models to enhance the overall performance.

Notice that, in this study, the objective is to recognize aspects of product features and the detailed reasons of consumes from product feature related sentimental review sentences. As noted in Section 3, it is modeled as a sequential label inference problem for textual data. Notably, the HMM is also a frequently utilized approach to infer hidden states for sequential tagging. It is known for the applications in speech, POS tagging, handwriting, bioinformatics, etc. Accordingly, to make comparisons with the proposed CRFs based approach, an HMM based approach was conducted. Specifically, battery related review sentences and screen related review sentences were analyzed. The performance was evaluated in terms of precision, recall and F_1 measures, which are reported in Fig. 3.

As seen from this figure, the proposed CRFs based approach outperformed the HMM based approach in both two datasets. In these datasets, high precision and recall were gained by both approaches for the recognition of product features from review sentences. However, for the recognition of aspects of product features and detailed reasons of consumes, the CRFs based approach were seen to perform much better than the HMM based approach. Especially, in the screen dataset, the HMM based approach nearly failed to recognize aspects of product features. As explained, only a few words were employed to denote a specific product feature within these two feature related datasets, which makes CRFs and HMM successfully label the correct words or phrases. However, fuzzy phrases and words tends to be utilized to describe aspects of product features and detailed reasons of consumers. Additionally, in this study, five tags were utilized to label different perspectives of consumers' concerns and only the transition probability between two successive tags are reckoned in the HMM based approach. Sufficient correlated information in review corpus is not captured. It might be the major reasons that leads to a relative poor performance on the recognition of aspects of product features and detailed reasons of consumes. Besides, more diversified phrases tends to be utilized in the screen dataset, which aggravate the poor performance.

In the previous sections, the objective of clustering identified aspects of product features and detailed reasons of consumers is highlighted to provide a concise description about CRs. In the following, battery and screen reviews of P_1 and P_2 are utilized as

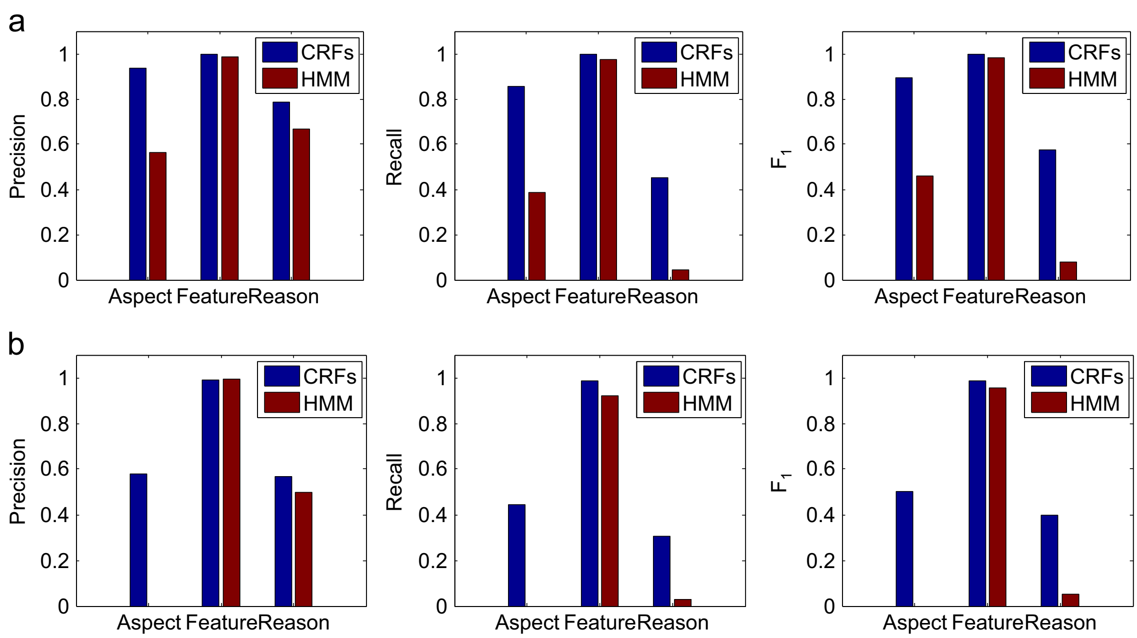


Fig. 3. Comparison of the CRFs based approach and the HMM based approach. (a) Performance comparison on the battery dataset (b) performance comparison on the screen dataset.

Table 10
Identified aspects of battery and the detailed reasons of consumers.

Aspects	{indicator}, {model}, {meter}, {maintenance}, {LIFE; life}, {power}, {load}, {saver}, {design}
Reasons	{GPS}, {wifi; internet}, {texting}, {application}, {call}, {charge; charging; charged; drained; drain; last; recharging; recharge; charger; discharge; KO}, {game}, {# hour; #-# hour; few hour; a couple of hour}, {email}, {direction}, {less than a day; a night; long more than # day; a day; # day; #.# day; #-# day; day}, {# pm}, {dead; died; run; bother; change; replace; get; replaceable}, {save}, {# year}, {widget}, {remove; removable}

Table 11
Identified aspects of screen and the detailed reasons of consumers.

Aspects	{brightness}, {keyboard}, {suck}, {protector; saver}, {side}, {lock}, {quality}, {prorifery}, {resolution}, {size}
Reasons	{resolution/brightness}, {see; Playing; freezed; used; flip; dying; rotate; Locked; read; created; turned; show; freeze; unlock}, {saver}, {responsiveness; response; unresponsive; scratch; responding; respond to touch; stuck}, {plastic}, {located}, {bright}, {impressive}, {bigger; small; big; large; larger}, {sluggish}, {resistive}, {cracked}, {sensitivity; sensitive}, {capacitive}, {defective}, {black}, {advertised}, {corrupted}, {blank; off}, {clean}

training data and reviews of P_3 are utilized as testing data. With the proposed co-clustering algorithm, aspects of battery and screen as well as detailed reasons of consumers are listed in Tables 10 and 11, respectively.

A number of detailed reviews that complain about the battery life are presented as clusters of words in Table 10. In particular, all the digits were replaced by the symbol "#" before the co-clustering algorithm was applied. These clusters are employed to describe CRs, such as "charge", "KO", "hour" and "day". Additionally, other reasons are frequently mentioned by consumers, such as "GPS", "internet", "texting", "application", "call", "game" and "email". These reviews actually point to the battery consumption by some specific functions of P_3. It suggests that battery designers of P_3 should check why these functions drain the battery quickly or provide a solution to prevent the battery being consumed too fast. There also exist consumer complains about the difficulty of replacing or removing the battery. Thus, a user-friendly design of battery replacement is recommended. Compared with reviews that point to the battery of a mobile, reviews that complain about the screen are diverse. A number of consumers disagree with the size of the screen by describing it as, for example, "small", "big" and "larger". Certain consumers disappointed with the response of the screen with descriptive words, such as "response" and "stuck". Intuitively, consumers refer to the screen's "sensitivity". However, generally, the word "sensitivity" and the word "response" are not literally defined to be similar to each other. Accordingly, as shown in Table 11, they are inferred as two clusters. However, in a specific domain, like one where reviews refer to the screen of mobile phones, these words should be grouped in the same cluster. Hence, in the future, some domain-specific similarities between words need to be considered and estimated in sophisticated algorithms.

6. Conclusions and future work

Online opinions are generated from time to time, which contain valuable CRs about products. Effectively understanding CRs at a fine-grain level based on online opinions exert an influential aspect on the improvement of products in market-driven design.

The objective of this research is to extract various aspects of product features and investigate detailed reasons regarding what make consumers unsatisfied with products. Particularly, product features and correlated sentiment polarity are identified with the help of pros and cons reviews. Next, an approach based on CRFs is used to pinpoint different aspects of features, as well as reasons of consumers from online reviews. Furthermore, a co-clustering algorithm is proposed to group jointly both aspects of features and reasons of consumers to provide a concise description regarding CRs for product designers.

In addition to the extraction of sentiment polarity in product feature level, this research enables designers to obtain much more insightful and critical suggestions from online reviews. This study facilitates designers to absorb CRs from big opinion data efficiently. In the future, a number of sophisticated models will be devised to enhance the extraction performance of aspects and reasons. Additionally, with the help of proposed approaches, a number of dedicated applications are expected to be developed, evaluated and applied in many real scenarios of product design to alleviate the burden of understanding CRs from a large volume of online opinion data.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. NSFC 71271185/G011201), the Fundamental Research Funds for the Central Universities (No. SKZZX2013091) and the Research Project of the Hong Kong Polytechnic University (No. G-YBE3).

References

Chen, L., Qi, L., 2011. Social opinion mining for supporting buyers complex decision making: exploratory user study and algorithm comparison. Soc. Netw. Anal. Min. 1 (4), 301–320.
Das, A. and Bandyopadhyay, S.,2010. Topic-based bengali opinion summarization. In: COLING'10, pp. 232–240.
Decker, R., Trusov, M., 2010. Estimating aggregate consumer preferences from online product reviews. Int. J. Res. Mark. 27 (4), 293–307.
Ding, X., Liu, B., and Yu, P.S. , 2008. A holistic lexicon-based approach to opinion mining. In: WSDM'08.
Goorha, S. and Ungar, L.,2010. Discovery of significant emerging trends. In: KDD'10, pp. 57–64.
Hedegaard, S. and Simonsen, J.G., 2013. Extracting usability and user experience information from online user reviews. In: CHI'13, pp. 2089–2098.
Hu, M. and Liu, B.,2004. Mining and summarizing customer reviews. In: KDD'04, pp. 168–177.
Jakob, N. and Gurevych, I., 2010. Extracting opinion targets in a single- and cross-domain setting with conditional random fields. In: EMNLP'10, pp. 1035–1045.
Jin, J., Ji, P. and Liu, Y., 2012. Product characteristic weighting for designer from online reviews: an ordinal classification approach. In: EDBT/ICDT'12 Workshop: BEWEB, Berlin, Germany,.
Kim, S.M. and Hovy, E., 2006. Automatic identification of pro and con reasons in online reviews. In: COLING'06, pp. 483–490.
Lee, T., 2007. Needs-based analysis of online customer reviews. In: ICEC'07, pp. 311–318.
Li, F., Han, C., Huang, M., Zhu, X., Xia, Y.-J., Zhang, S., and Yu, H., 2010a. Structure-aware review mining and summarization. In: COLING'10, pp. 653–661.
Li, H., Bhowmick, S.S., and Sun, A., 2010b. Affinity-driven prediction and ranking of products in online product review sites. In: CIKM'10, pp. 1745–1748.
Li, S., Zha, Z.J., Ming, Z., Wang, M., Chua, T.S., Guo, J., and Xu, W., 2011. Product comparison using comparative relations. In: SIGIR'11, pp. 1151–1152.
Lin, C. and He, Y.,2009. Joint sentiment/topic model for sentiment analysis. In: CIKM'09, pp. 375–384.
Liu, Y., Jin, J., Ji, P., Harding, J.A., Fung, R.Y.K., 2013. Identifying Helpful online reviews: a product designer's perspective. Comput. Aided Des. 45 (2), 180–194.
Ma, Z., Sun, A., Yuan, Q., and Cong, G., 2012. Topic-driven reader comments summarization. In: CIKM'12, pp. 265–274.
Mei, Q., Ling, X., Wondra, M., Su, H., and Zhai, C., 2007. Topic sentiment mixture: modeling facets and opinions in weblogs. In: WWW'07, pp. 171–180.
Meng, X., Wei, F., Liu, X., Zhou, M., Li, S., and Wang, H., 2012. Entity-centric topic-oriented opinion summarization in twitter. In: KDD'12, pp. 379–387.
Miao, Q., Zhang, S., Meng, Y., and Yu, H., 2013. Domain-sensitive opinion leader mining from online review communities. In: WWW'13 Companion, pp. 187–188.
Mizerski, R.W., 1982. An attribution explanation of the disproportionate influence of unfavorable information. J. Consum. Res. 9, 301–310.
Moghaddam, S. and Ester, M., 2012. Aspect-based opinion mining from product reviews. In: SIGIR'12, pp. 1184-1184.
Netzer, O., Feldman, R., Goldenberg, J., Fresko, M., 2012. Mine your own business: market-structure surveillance through text mining. Mark. Sci. 31 (3), 521–543.
Pang, B. and Lee, L., 2004. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: ACL'04.
Si, J., Li, Q., Qian, T., Deng, X., 2013. Users interest grouping from online reviews based on topic frequency and order. World Wide Web 17, 1321–1342.
Tucker, C. and Kim, H.M., 2011. Predicting emerging product design trend by mining publicity available customer review data. In: ICED'11, pp. 43–52.
Wang, L., Youn, B.D., Azarm, S., and Kannan, P.K., 2011. Customer-driven product design selection using web based user-generated content. In: IDETC'11, pp. 405–419.
Wilson, T., Wiebe, J., and Hoffmann, P., 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In: EMNLP'05, pp. 347–354.
Yu, J., Zha, Z.J., Wang, M., and Chua, T.S., 2011. Aspect ranking: identifying important product aspects from online consumer reviews. In: ACL'11, pp. 1496–1505.
Zhai, Z., Liu, B., Xu, H., and Jia, P., 2010. Grouping product features using semi-supervised learning with soft-constraints. In: COLING'10, pp. 1272–1280.
Zhan, J., Loh, H.T., Liu, Y., 2009. Gather customer concerns from online product reviews—a text summarization approach. Expert Syst. Appl. 36 (2 Part 1), 2107–2115.
Zhang, K., Narayanan, R. and Choudhary, A., 2010. Voice of the customers: mining online customer reviews for product feature-based ranking. In: WOSN'10, Boston, USA,.
Zhang, Z., Guo, C., Goes, P., 2013. Product comparison networks for competitive analysis of online word-of-mouth. ACM Trans. Manag. Inf. Syst. 3 (4), 1–20.
Zhuang, L., Jing, F., and Zhu, X.Y., 2006. Movie review mining and summarization. In: CIKM'06, pp. 43–50.