

Customer Requirements Elicitation based on Social Network Service

Yoon-kyu Lee, Neung-Hoe Kim, Dohoon Kim, Dong-hyun Lee and Hoh Peter In

Department of Computer Science & Engineering, Korea University

Anam-Dong, Sungbuk-Gu, Seoul 136-713, Korea

[e-mail: {beatbox3, nunghoi, karmy01, tellmehenry, hoh_in}@korea.ac.kr]

*Corresponding author: Hoh Peter In

*Received April 7, 2011; revised July 24, 2011; accepted August 18, 2011;
published October 31, 2011*

Abstract

In the early stages of a software project, it is critical to understand the needs of the customers and elicit their customer requirements. Various requirements elicitation methods have been proposed. However, existing methods still have the limitations such as a limited number of target customers, limited expression of customers' opinions, and difficulty in collecting the customers' opinions continuously. A novel method for eliciting customer requirements is proposed by utilizing a social network service (SNS), which is a shared source of raw information of the customers' needs and opinions. The proposed method is validated to show its effectiveness in overcoming the limitations of existing methods.

Keywords: Software engineering, customer requirements, requirements engineering, SNS, social network

The part of this paper was presented in the ICONI (International Conference on Internet) 2010, December 16-20, 2010, Philippines. This research was supported by the Korea University (T1001531), the MKE (the Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2011-(C1090-1131-0008), the Korea Research Foundation Grant funded by the Korean government (MOEHRD, Basic Research Promotion Fund) (KRF-2007-331-D00360), and the Mid-career Researcher Program through NRF grant funded by MEST (No.2010-0000142).

DOI: 10.3837/tiis.2011.10.003

1. Introduction

“Understanding what customers really want” is a critical factor in the success of a software project. If customer requirements are insufficiently comprehended, the software may not meet the needs of the customers, and, eventually, the software will be abandoned [1][2]. According to the Standish Group CHAOS report, about 50% of projects were scrapped due to failure of good requirements elicitation [3].

Because of the importance of customer requirements elicitation (CRE), various methods involving interviews, surveys, observations, and simulations under different scenarios have been proposed [4]. However, these methods still have the following limitations:

- *Limited number of target customers*—Most of the aforementioned methods first choose a sample customer group and then collect the requirements. This is mainly done because of temporal and spatial constraints. Sample customer groups may not represent the entire customer population. Therefore, the collected requirements are necessarily limited, and the opinions of latent customers are overlooked.
- *Limited expression of customers’ opinions*—The existing methods force customers to answer predefined questions. Therefore, the results may not escape from intention of the questioners/interviewers. This also induces a “socially desirable response”¹ [5], which results from the artificial setting of the questions. This makes it difficult to elicit the inner needs of the customers.
- *Difficulty in collecting customer opinions continuously*—Through the existing methods, it is difficult to collect customers’ follow-up opinions on a continuous basis, after the interview or survey. This results in a potential loss of additional information and requirements.

Considering these problems, this study proposed a social network service (SNS)-based CRE, which utilizes the SNS to overcome the existing CRE limitations. Through the SNS, individuals actively build relationships, share relevant information, and exchange opinions, including customer needs. The SNS-based CRE enables us to overcome the aforementioned limitations of the existing methods and offers the following advantages:

- *Access to a wider range of customers*—In contrast to the existing methods, the SNS-based CRE enables the collection of relevant data from a larger pool of customers. Because it is possible to include a diverse range of customers without temporal and spatial constraints, it is also possible to derive a variety of new perspectives that are unbounded by sample sets representative of specific group(s).
- *Unlimited expression of opinions without a predefined format*—SNS-based CRE enables researchers to avoid the “socially desirable response” as it allows customers to freely express and share their opinions via the SNS without a predefined format. Many studies have focused on how to gather system requirements with well-defined formats from the opinions of specific users. However, our proposed method aims to elicit early business requirements from a diverse set of nonspecific users, in an informal environment.

¹ *Socially Desirable Response*: Socially desirable response has been of long-standing interest in the field of marketing, especially in terms of survey methods. When stating opinions, respondents show a tendency of not responding truthfully but simply providing answers that will make them look good [5].

- *Collection of continuous customer opinions*—SNS forms a continuous connection between the customers and the requirements elicitors. Thus, the customers can continuously express their follow-up requirements.

The rest of the paper is organized as follows. Section 2 discusses related work on the SNS and CRE methods. Section 3 introduces the proposed SNS-based CRE method. Section 4 presents a case study and analyzes the results. Section 5 presents the validation of our method, and Section 6 presents the conclusions of our study and discusses future work.

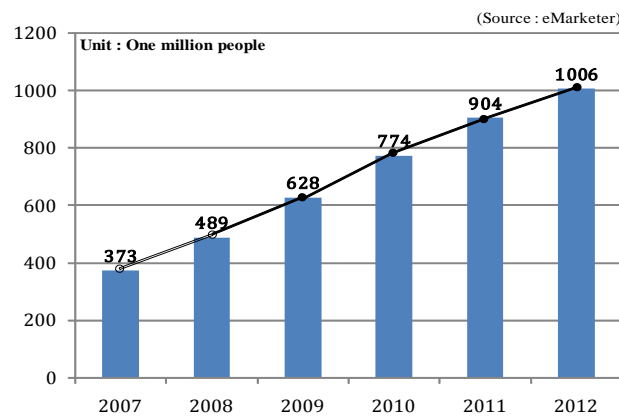


Fig. 1. Statistics of SNS users around the world

2. Related Work

2.1 Social Network Service

A social network is defined as a social relationship structure in which individuals are related (directly or indirectly) to each other based on interest, friendship, trust, etc. From the perspective of computation theory, the analysis of a social network is a complex domain because of the numerous nodes and the large number of relationships among them. Given that SNSs have been receiving considerable attention recently, active research in various areas such as liberal arts, economics, and engineering is underway [6][7].

An SNS is an online system for building personal relationships. By building relationships and posting opinions or experiences on particular digital spaces, an SNS continuously provides news about an individual to the communities and users that share interests with this individual. Most SNSs are Web-based services; however, recently, mobile-based services have begun to be supported. The most commonly used SNSs are Facebook, Twitter, MySpace, and Linknow [8].

The influence of such SNSs is increasing in proportion to the rapid increase in the number of SNS users (Fig. 1) [9]. A recent research shows that 175 million people use Twitter worldwide, and 200 million tweets are exchanged daily. In the case of Facebook, studies show that over 700 million users also share 30 billion pieces of content daily. Another research conducted a survey of 48,000 Internet users and showed that 6 out of 10 people have SNS accounts that they use on an average of 4.6 h per week (Web browsing: 2.3 h; e-mail: 4.4 h, multimedia viewing: 3.9 h) [10][11][12].

As the number and influence of SNS users are increasing, SNSs have evolved into an appropriate medium to collect various individual opinions based on data from the participants.

Initially, SNSs prioritized the social purpose of amity among users but, currently, they are used to evaluate the demographics of particular goods and services as well as the preferences/tastes of certain groups of users. A recent research showed that users posted product reviews on SNSs at least once a month, and 1/4 of them responded to these reviews [13]. SNS users express their judgment, experience, and level of satisfaction/dissatisfaction without inhibitions, aiming to prove the objectivity and reliability of their opinions. Consequently, many similar opinions form a public opinion and a large knowledge base [6].

2.2 Customer Requirements Elicitation Methods

Various methods to elicit customer requirements have been devised. The most commonly used CRE methods are as follows.

The interview technique is the most general method in which users, stakeholders, and domain experts are questioned directly to elicit their needs and requirements. If the respondents have enough chances to answer the questions, the interview can prove to be an excellent source of customer requirements. A “structural interview” (a sort of interview using prepared questions) is commonly used and is effective in obtaining answers that interviewers expect [14]. There is an obvious limitation in terms of actively eliciting new requirements from customers: a structured interview tends to limit the customers’ opinions within the scope set by the interviewers. Besides, the target customer sample is limited owing to time and space constraints.

A *survey* is a method by which particular customers, who are assumed to have a certain amount of information or knowledge about a subject, are asked questions. A survey can be carried out via telephone, e-mail, or online, and its results are integrated. A survey is commonly used in a situation where a researcher needs to collect information effectively and he/she is bound by time constraints [15]. However, since the subjects of a survey are mostly selected based on certain conditions, the results do not reflect the opinions of a wide range of customers. In addition, it is difficult to collect customers’ follow-up opinions because there is usually no connection between the selected customers and the questioners/interviewers. It is also difficult to collect active opinions from customers because those questioned answer according to predetermined questionnaires; thus, their answers are restricted.

Contextual inquiry is a user-centered method that is a part of the contextual design method. Contextual inquiry involves collecting detailed information about a user’s work practice by observing and interviewing the user while he/she actually works. The researcher stays in the background and lets the user lead the situation as much as possible. This method is effective for comprehending a certain user and his/her needs [16]. However, this method cannot target a wide range of users.

Existing CRE methods have the above limitations. Therefore, it is necessary to access a wider range of customers without temporal and spatial constraints, collect opinions freely and without predefined questionnaires, and form continuous connections for collecting additional opinions.

3. SNS-based Customer Requirements Elicitation

In this study, we introduce the guidelines of the SNS-based CRE method, specifically how to approach an SNS, how to apply it to the collection of customer requirements, and how to refine and prioritize these collected requirements. The entire process is divided into three phases (*preparing, gathering, and refining*). Each of these three phases is further divided into specific stages (Fig. 2). The stages of each phase are explained in detail below. Moreover, examples

and parts of the results of case studies are appended to facilitate comprehension.

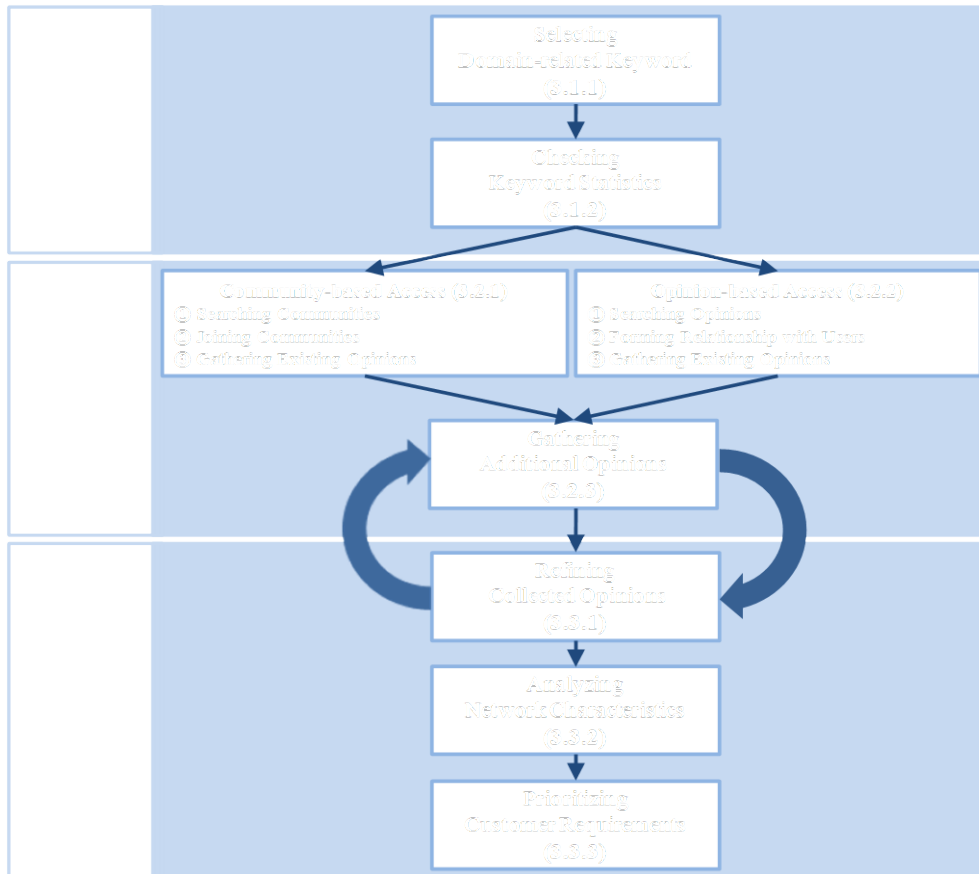


Fig. 2. Process of SNS-based CRE

3.1 Preparing Phase

This phase describes the preparation for CRE through SNS. To approach the extensive number of user opinions, a keyword-based approach is used. Domain-related keywords are selected to understand user trends, which is fundamental to finding what users want. A keyword-based approach helps researchers to understand the trend of the users' needs in the network more easily and effectively. Therefore, it is universally adopted in various domains like marketing and social statistics, which aims to comprehend people's interests [17]. This phase is further divided into the following two sub stages:

3.1.1 Selecting Domain-Related Keywords

In this stage, domain-related keywords that represent features, functions, or quality of the product are selected. When selecting keywords, it is necessary to select those words that carry the least possibility of vagueness or confusion. **Table 1** lists examples of keywords related to smartphones.

3.1.2 Checking Keyword Statistics

In this stage, the frequency of the selected domain-related keywords on an SNS is checked for a certain period. After checking keyword statistics, highly ranked keywords are chosen in ord-

Table 1. Ex

Keywords	Smartphone App	Smartphone OS	Smartphone Design	Smartphone Music	Smartphone Price	Smartphone Camera	Smartphone Weight
Frequency	254	127	115	115	63	38	28

er to focus on the general trend of user needs. Through keyword statistics, the degree of users' interests is understood; based on this, the overall users' needs are comprehended and a concrete direction for performing a CRE is set. **Table 1** shows part of the statistics on smartphone-related keywords that were collected from Twitter from June 15 to July 15, 2010.

3.2 Gathering Phase

During the gathering phase, the SNS and its users are approached and their opinions are collected. This phase is divided into two types of access according to the subjects and the direction of approach: community-based access and opinion-based access. Because the social network is a space comprising complex relationships among various communities and users, two different ways are set to approach the SNS more effectively.

3.2.1 Community-based Access

Community-based access targets diverse communities in SNS. Communities that are concerned with highly ranked keywords are chosen, and then, the existing opinions of the users in these communities are collected. This access is appropriate for collecting more concrete and reliable opinions because most members of these communities tend to have an interest and knowledge in the relevant domain. This access consists of three sub stages.

- ① **Searching Communities:** In this stage, communities that have interests in the highly ranked keywords chosen in the preparation phase are selected. Most of the SNSs provide community services and a keyword-based community search function. The criteria to select the communities are the recent trends in activity and member participation.
- ② **Joining Communities:** Requirements elicitors join each of the selected communities as a member and form relationships with members of these communities. Users with common interests are able to join the communities and to build relationships in the SNS without a complex processes.
- ③ **Gathering Existing Opinions:** In this stage, the opinions exchanged among the members of the selected communities are gathered. Based on the highly ranked keywords, the opinions, posts, and comments within the communities are first comprehended, and then, the relevant opinions are gathered.

3.2.2 Opinion-based Access

Opinion-based access targets individual users and not SNS communities. The opinions with the highly ranked keywords are searched, and the opinions of the users who suggested them are collected. In opinion-based access, individual users are approached based on their opinions rather than on the particular users or groups; therefore, it targets unlimited potential users, enabling the collection of a wide variety of opinions from diverse users.

- ① **Searching Opinions:** In this stage, searches are conducted for the opinions that users type on an SNS, along with highly ranked keywords. Included within the scope of the search are the typed-in opinions that also have their matches. As most SNSs provide the keyword-based search function, it is not difficult to search opinions that include such

keywords.

- ② **Forming Relationships with Users:** Relationships are formed with users who have provided the searched opinions. Through the SNS, people from different national, cultural, ethnic, and occupational backgrounds can build relationships without inhibitions.
- ③ **Gathering Existing Opinions:** In this stage, the existing opinions, which were posted on the SNS by the relevant users, are collected. Based on the highly ranked keywords, the associated opinions, posts, and comments by the relevant users on the SNS are collected.

3.2.3 Gathering Additional Opinions

First, extra opinions, which are for supplementing existing opinions, are collected from users who have a relationship with requirements elicitors through *community-based access* and *opinion-based access*. The newly collected opinions supplement the previous vague opinions with improved details. After the completion of the first round of collecting opinions, additional opinions are collected through the continuous relationship with the users during the CRE period. In contrast to the existing CRE methods, this method enables the collection of additional opinions by continuously communicating with the users at any time.

Gathered opinions are organized and managed in tabular format (**Table 2**). This table lists the types of opinions, version information, and related keywords; the SNS user ID is used as the table key. To maintain traceability, the time at which the opinion was posted should be appended to each opinion. Based on type, opinions are divided into “existing opinion” or “additional opinion.” As each opinion is elaborated upon and specified, they are managed according to their version.

Table 2. Example of tabular format for managing gathered opinions

User ID	Type	Opinions					Related Keyword
		Original	Version_1	Version_2	...	Final	
Sy**	Existing Opinion	My battery goes dead tooooo soon...but an external battery'd look horrible..:((2011/01/22)	They must last longer. Longer than a day at least (2011/01/30)	It should be slim too (2010/01/31)		Smartphone batteries need to be more durable and slim.	Smartphone Design
	Additional Opinion	Smartphones all look the same and dull!! (2011/01/30)	Custom-design may help to satisfy customers (2011/01/31)	-		Custom-design needed!	Smartphone Design

3.3 Refining Phase

Refined and prioritized final customer requirements are elicited in this phase. Refinement and prioritization are carried out based on the collected opinions and information on users' and opinions' network characteristics. Additional opinions are gathered simultaneously through established relationships with the users, and these opinions are also refined. This phase is divided into the following three sub stages.

3.3.1 Refining Collected Opinions

First, in this stage, we filter the most helpful opinions from among those gathered in the previous stages, and we assign attributes to these filtered opinions. Assigning attributes facilitates categorizing and refining the opinions as well as obtaining the objectivity and traceability of the requirements [14]. Attributes are divided into three classes: *positive*, *negative*, and *informative*. Each attribute class is named after its connotation. Opinions about a certain product suggested by users through the SNS usually fall within the boundary of these

three classes; those falling outside the boundary are mostly not related to customer requirements [18]. **Table 3** shows examples of how the gathered opinions are organized.

Table 3. Example of the organization of searched and collected opinions

Opinion	Attribute
The best: The YouVersion Bible app is the best thing smartphones have to offer. Love my Bible app!" < #Cosign	Positive
Great ideas for new smartphone apps—Live soil map for farmers showing soil types, best local plants, value, etc.	Informative
Is Nokia problem not app: 'm so convinced. Gmail application on Nokia Smartphones can be so unstable!	Negative
Android is the best smartphone OS I have exp till today. And it rocks on the Samsung Galaxy s!	Positive
Palm and HP are betting a lot on their new web-based OS. Expanding it way beyond smartphones and into slates and peripherals.	Informative

After assigning attributes to each opinion, a *transformation* is carried out. Because the collected opinions on the SNS are not in a standard form, they are standardized. This *transformation* includes two major steps (**Fig. 3**):

- *Restoration of Abridgment/Abbreviations*

Due to the spread of online jargon, a considerable number of opinions on an SNS are difficult to understand as they are expressed in very original forms. To accurately understand user opinions, therefore, it is necessary to alter these online jargon and abbreviations and form grammatically correct phrases. Various dictionaries for online jargon and abbreviations already exist. **Fig. 3** shows an example of the correction of an abbreviated opinion.

- *Standardization*

The assigned attributes and the main keywords constituting opinions serve as the standardization criteria of the written forms. **Fig. 3** exemplifies how the collected opinions are transformed into standard requirement forms.

The syntactical complication renders the *transformation* difficult. However, most opinions exchanged on an SNS are generally not syntactically complex.

Original Opinion		
IMO, problem with the design of all other smartphn are butt ugly carrier logos. why is it only palm and apple can gafi?		
Restoration		
Category	Attribute	Restored Opinion
Smartphone Design	Negative	In my opinion, problem with the design of all other smartphone are butt ugly carrier logos. why is it only palm and apple can get away from it?
Standardization		
Category	Attribute	Customer Requirement
Smartphone Design	Negative	Ugly butt carrier logos should be avoided

Fig. 3. Example of *transformation* of collected opinion

Table 4. Network characteristics of users and their opinions

Objective	User	Opinion
Evaluation Criteria	The degree of influence	
Related factors in network theory	Centrality ² , Reachability ³	Ripple Effect
Factors in SNS	Number of Followers, Following, and Friends	Number of Citations, Retweets, and Views

3.3.2 Analyzing Network Characteristics

In this stage, the network characteristics of each user and his/her opinions are analyzed. As shown in **Table 4**, the criteria include how influential the users are and how influential their opinions are on a certain social network.

$I_i = H_{k_i} \times O_i$	I_i	Importance of i
	H_{k_i}	Influence of k_i
	O_i	Ripple effect of opinion i
	i	Customer requirement identifier
	k_i	User identifier who suggested i
$H_{k_i} = \frac{1-d}{N} \times d \sum_{k_j \in M(k_i)} \frac{H_{k_j}}{F_{k_j}}$	d	Damping factor (default: 0.85)
	N	Total number of users in relevant SNS
	F_{k_i}	Number of people who have relationship with k_i
	$M(k_i)$	Set of users who have relationship with k_i
$O_i = \frac{\omega_i}{T}$	t	Time when factor O_i is measured
	ω_i	Number of citations of i at point t
	T	Period from when i was suggested to

Fig. 4. Formulas for calculating *importance*

3.3.3 Prioritizing Customer Requirements

When resources are limited, prioritizing customer requirements helps manufacturers to reflect the requirements to the software project in order of importance. In this stage, elicited requirements are first prioritized on the basis of the keywords; then, the requirements, which belong to each keyword, are prioritized according to their *importance*.

Fig. 4 shows a formula to determine *importance*, which is based on the network characteristics analyzed in the previous stage. By prioritizing each customer requirement according to its *importance*, it becomes possible to determine how influential and important that requirement is. In addition, it enables us to judge which customer requirement is more representative of the needs of a multitude of users.

² *Centrality*: In graph theory and network analysis, diverse methods are used to measure the centrality of a vertex (degree, closeness, and betweenness) and to determine the relative importance of a vertex within a graph [19].

³ *Reachability*: In graph theory, reachability refers to the ability to move from one vertex to another in a directed graph [20].

The *importance* of customer requirements consists of the “influence of the user” who suggested the relevant opinion (factor H) and the “ripple effect of the opinion” (factor O). In a social network, the node’s influence and the opinion of the relevant node are related; therefore, both factors are considered together.

Though a variety of methods have been suggested to measure users’ influence in a social network, thus far, the superiority of one method compared to the others has not been verified. Therefore, in our study, the page rank algorithm⁴ is chosen to calculate factor H , which is general and produces a similar output to other methods but is relatively easy to calculate [21]. The greater this factor is, the larger is the influence of the relevant user and the more influential are his/her opinions.

In the case of factor O , the ripple effect of opinion i at point t is calculated with the quotient of the number of citations divided by T . As the number of citations or references increases in a short period, factor O increases; thus, we can judge whether the opinion has a large ripple effect.

4. Case Study

4.1 A Case Study: Smartphone Requirements Elicitation

A case study was carried out through the analysis of Twitter and Facebook, two well-known SNSs. Twitter was chosen as the target of the experiments since it provides various applications for its diverse functions based on an open API (Application Programming Interface), and it retains a user pool of more than 200 million people. Facebook was chosen because it too has a user pool of more than 700 million people and provides diverse community services.

The case study on the smartphone domain was carried out over three weeks, from January 20 to February 10, 2011. Five thousand and ninety-one tweets and posts were analyzed and, as a result, 67 customer requirements were finally elicited. The case study was conducted in English. Specific results in each phase are as follows.

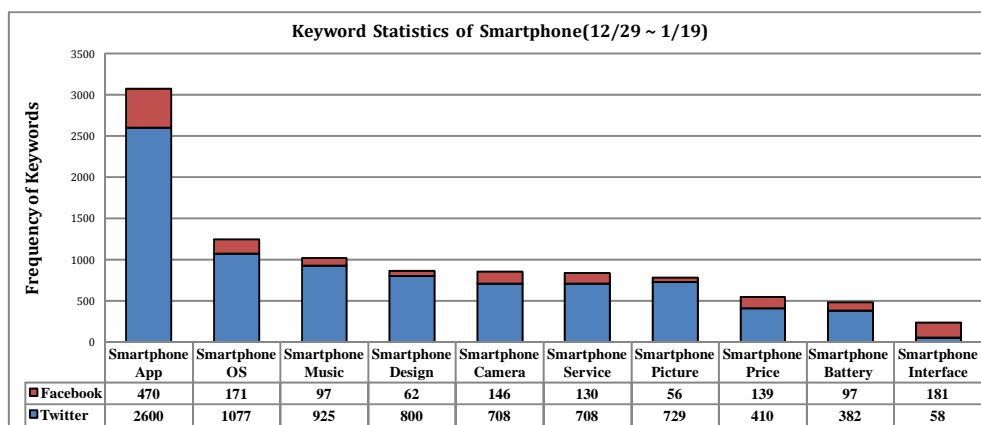


Fig. 5. Keyword statistics for smartphones

⁴ *PageRank Algorithm*: The algorithm used by the Google search engine. The importance of a Web page is judged by the number of hyperlinks from other pages. It is based on the premise, prevalent in the world of academia, that the importance of a research paper can be judged by the number of citations the paper has from other research papers [22].

1. Preparing Phase

Thirty-one smartphone-related keywords were selected, and keyword statistics pertaining to the keywords were collected from December 29, 2010, to January 19, 2011, through Facebook and Twitter. Keyword statistics were collected through Facebook and Twitter statistics APIs, and the top 10 keywords are shown in Fig. 5.

2. Gathering Phase

Community-based access was performed through Facebook, which provides various functions related to communities, whereas *opinion-based access* was performed through Twitter, which supports diverse APIs and functions associated with opinion searching.

- *Community-based Access*

Five communities with active communication and participation within the previous 30 days were selected from among the top 10 keyword-related communities on Facebook. We joined the communities and established relationships with 65 active members. Based on the top 10 keywords selected from the previous phase, 93 existing opinions were collected and 28 extra opinions for supplementing existing opinions were collected in the first gathering.

- *Opinion-based Access*

Opinions on Twitter were searched based on the top 10 keywords. We searched 4888 existing opinions, and 115 opinions remained after excluding duplicates and unrelated opinions. We established relationships with 44 users who suggested relevant opinions, and we collected 27 extra opinions for supplementing existing opinions in the first gathering.

Table 5. Part of the final results by SNS-based CRE

Keyword Priority	Keyword	Attribute	Req. Priority	Requirement	User	Importance	ID
1	Smartphone App	Positive	1	An application that translates user's talk directly into another language is needed.	Fu**	3.63	01_P_001
			2	An application that manages user's pay, usage, and deposits is very useful.	Xe**	3.24	01_P_002
			3	An application that completely manages all types of messaging applications is needed.	En**	3.02	01_P_003
		Negative	1	Bank applications are vulnerable to attack.	Jo**	2.69	01_N_001
		Inform	1	An application that manages users' credit cards is released.	Ys**	2.61	01_I_001
2	Smartphone Design	Positive	1	More slim and powerful battery design is needed.	Sy**	2.56	02_P_001
		Negative	1	Designs of existing smartphones look remarkably similar	Mj**	2.87	02_N_001
			2	Built-in battery design(iPhone) needs to be avoided	Ke**	1.41	02_N_002

3. Refining Phase

In the *refining phase*, additional opinions were gathered by establishing continuous relationships with users; simultaneously, these opinions were also refined. Through continuous relationships, 75 additional opinions were gathered.

The most helpful opinions were filtered and duplicated, or similar opinions were combined. Attributes were assigned to each opinion, and then, *transformation* was performed.

After the *transformation*, we prioritized the final customer requirements on the basis of *importance*. A part of the final results is shown in Table 5.

Table 6. Numerical results of the case study

	Community-based Access	Opinion-based Access
# of Searched Objectives	5 communities (105 users)	4888 opinions (193 users)
# of Objectives with Relationships	65 users	44 users
% of Response	61.9%	22.79%
# of Searched Existing Opinions	203	4888
# of Filtered Existing Opinions	93	115
# of Extra Opinions for Supplementing Existing Opinions	28	27
# of Additional Opinions through Continuous Relationships	34	41
# of Final Elicited Requirements	14	53

4.2 Analysis

Four distinct features are found from the results of the case study (**Table 6**).

- The number of useful opinions is significantly smaller than that of searched opinions. This is because cited, duplicated, similar, and unrelated opinions were searched.
- The percentages of the responses of the *community-based access* and those of the *opinion-based access* differ greatly. This difference is attributed to the fact that Facebook and Twitter offer different service types and access methods was different. Regarding service types, it is difficult to build relationships and to gather opinions in a short period because relationships are only one-way on Twitter, as compared to Facebook, on which relationships are two-way. In contrast, with regard to access methods, it is easier to gather opinions through *community-based access* because community members tend to be more active than ordinary users when exchanging opinions.
- The number of additional opinions gathered through continuous relationships is over 22% of the total number of opinions. This means that a considerable number of user opinions were gathered through continuous relationships even after the first opinion gathering had finished.
- The number of opinions as well as that of final elicited requirements is larger when we use *opinion-based access* than *community-based access*. This shows that opinion-based access that approaches general users who do not belong to a specific community can supplement opinions that can be easily overlooked through monitoring specific users in a particular community. It also shows that opinions gathered from *opinion-based access* are indeed more diverse and useful.

5. Validation

To validate the method proposed in our study, we qualitatively compared the existing CRE method carried out by a Korean smartphone company and the SNS-based CRE. The existing method was performed through an online survey and in-depth interview. As shown in **Fig. 6**, both methods were related to the smartphone domain. In the case of the existing method, the online survey for investigating customers' needs was performed on a target of 1200 customers in North America, 48 of which were subsequently selected for eliciting customer requirements through an in-depth interview. On the other hand, the SNS-based CRE investigated the users' needs from 9946 users who used the main keywords of users' needs, and we elicited customer requirements from 298 users who suggested opinions.

In terms of capturing customers' needs, the result shows that the main keywords that are related to the customer needs, which were gathered from each method, are mostly similar or

the same. This means that our method is as effective as the existing method in capturing customer needs.

However, in terms of eliciting customer requirements, the result shows quite a big difference in the number of final requirements and the content. Yet, the difference in the number of the final requirements could be mainly attributed to the period and time of the investigation; therefore, we estimated the includedness of the requirements elicited from each method based on their contents. Thirty percent of the requirements collected from the SNS-based CRE and 55% from the existing method were shown as similar or the same requirements. Considering this, the requirements set from the SNS-based CRE encompasses a more diverse and wider range of customer requirements than those of the existing method.

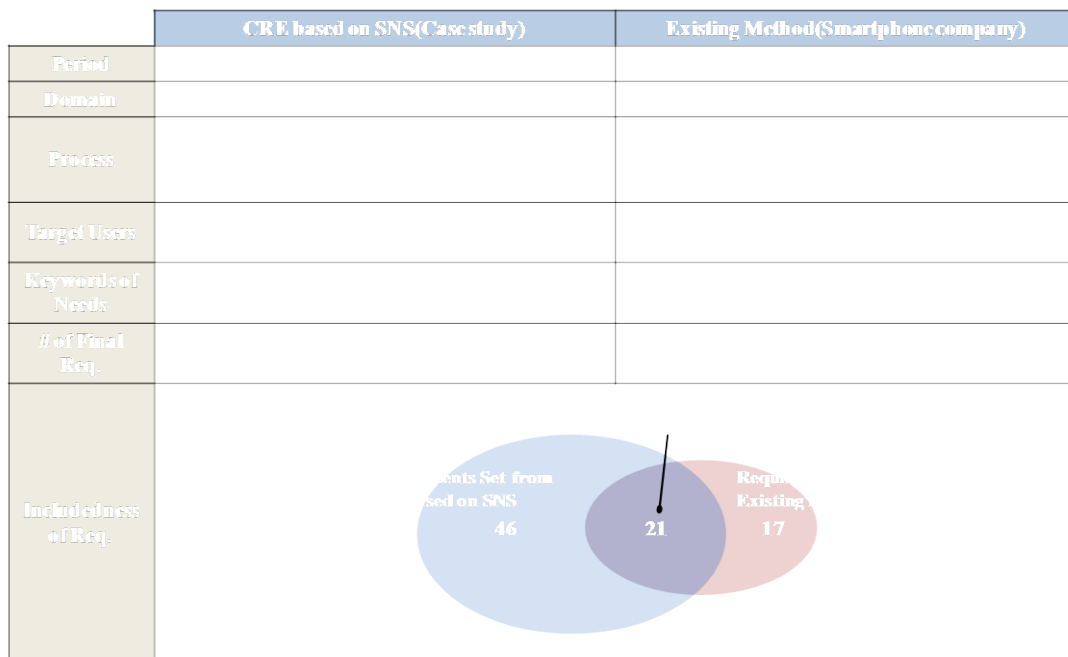


Fig. 6. Qualitative comparison

5.1 Validation of Claims

We presented three advantages of the SNS-based CRE in Section 1. To validate these three advantages, we asked questions to users with whom we had established relationships, and subsequently used the results of the case study.

① Access to a wider range of customers

The SNS-based CRE uses all SNS users around the world, and does not depend on a limited pool of users. The characteristics of the SNS are unbounded by time and space and enable access to a diverse pool of potential customers. **Fig. 7** shows the diversity in the nationalities and occupations of the users who participated in the case study (298 users). It is safe to say that the customer requirements of the SNS-based CRE were indeed elicited from a wide range of potential customers.

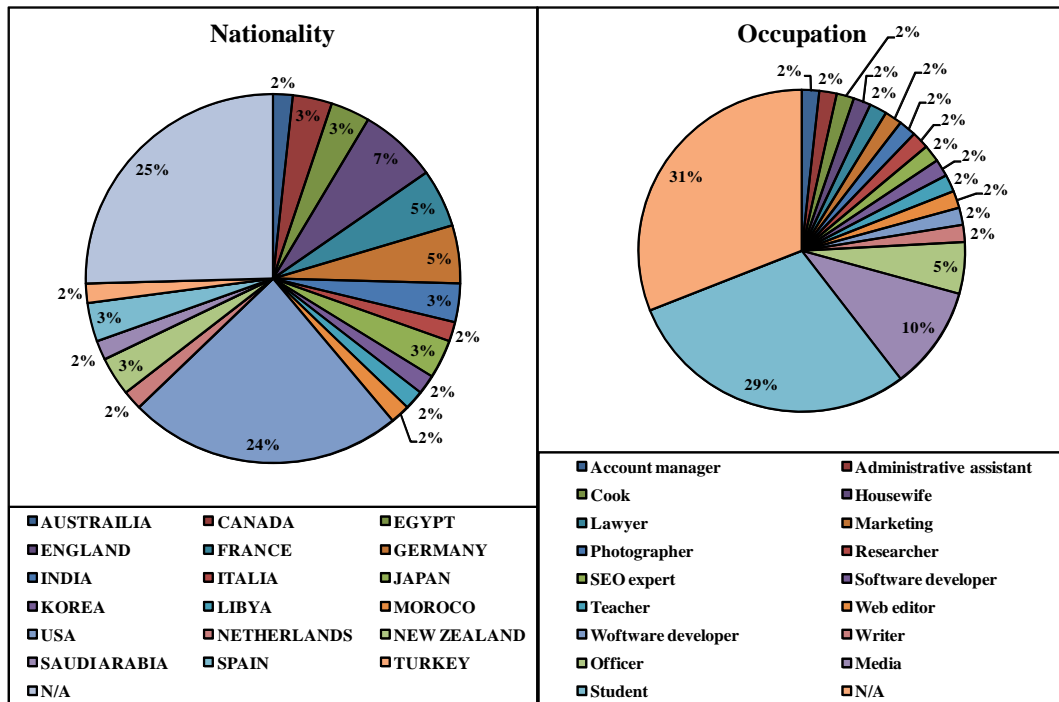


Fig. 7. Nationality and occupation of users

② Unlimited expression of opinions without predefined format

For 37 users with whom we established relationships, we investigated their preferences to the CRE methods and the reasons for it. As shown in Fig. 8, 46% of the users prefer the SNS-based CRE, and they responded with a variety of reasons for their preference. Most users who prefer the SNS-based CRE mentioned that they could answer without any predefined format and restriction. Other reasons were that users could express their thoughts without inhibition, access the SNS easily, and share their opinions easily.

Judging from the users’ preferences and their reasons, we found that the SNS-based CRE could be more preferred by users than the existing methods, and they think that the SNS-based CRE allows them to express their opinions freely.

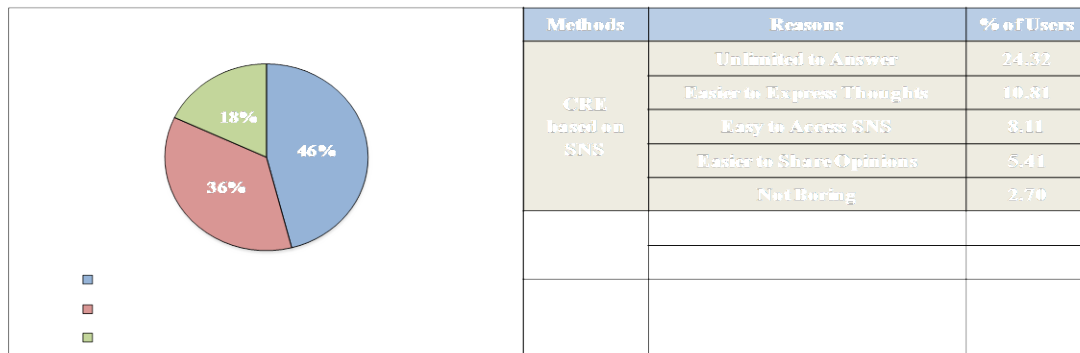


Fig. 8. Users’ preferences and reasons

③ Collection of customers’ opinions continuously

As shown in the results of the case study (Table 5), the number of opinions gathered through

continuous relation
opinions were refl
channels of contin
collecting opinions
opinions and ideas
opinions.

< Descriptive Statistics >

ID	Requirements	Avg. of Priority	Std. Deviation	Min. of Priority	Max. Of Priority
1	A battery with lower discharge rate is required.	1.5	0.71	1	2
2	2/3 of present price of smart phone would be reasonable.	3.5	2.12	2	5
3	It needs to get rid of restraints of recharging through gender.	4.5	2.12	3	6
4	Flash or pop-ups should work normally when using the internet.	3.5	0.71	3	4
5	An application which combines and enables to manage Kakao Talk, SMS and Whatsapp together is necessary.	3	2.83	1	5
6	A function to offer various special characters and keyboard is required.	5	1.41	4	6
7	A new design differentiated from existing ones is needed (query keyboard+touch screen).	9	2.83	7	11
8	It is required to do the internet shopping and pay more freely.	9	1.41	8	10
9	An appropriate distribution of functions between button input system and touch screen input system is needed	10.5	2.12	9	12
10	A function as a R-HDD (external hard drive) is necessary.	12	2.83	10	14
11	An application which interconnects online telephone directory with a mobile phone.	13	2.83	11	15
12	Setting groups within address book is needed.	14.5	3.54	12	17
13	Recording during phone call is needed.	15.5	3.54	13	18
14	An application which is connected with public transportation and traffic lights is necessary.	10.5	4.95	7	14
15	It is required to make it free to register as an application developer.	17	2.83	15	19
16	An application to view events and amenities on main streets at a look.	16	0	16	16
17	A function is required to prevent damages of mobile phones by maintaining it within certain temperature.	15	2.83	13	17
18	A function to measure temperature change or heartbeats is needed.	13.5	6.36	9	18
19	An application to help shopping is needed.	19.5	0.71	19	20
20	An application to check time to take medicine and whether the medicine is taken is necessary.	14	8.49	8	20

< Test Statistics >

Kendall's W	0.828
N	2
Chi-square	31.457
Degree of Freedom	19
Significance level	0.05

Table 7. Validation of prioritization by Kendall's W

5.2 Validation of Prioritization

To validate the prioritization based on the *importance* of the customer requirements proposed in this paper, we performed an experiment to compare it with the order of the priority of the needs of the users. For this experiment, 30 Facebook users were selected randomly and included in a small network to collect opinions about smartphones from February 12 to February 27, 2011. After collecting the opinions, with the process proposed in this paper, 20 prioritized final requirements were elicited. Then, these requirements were sent back to the users, and we asked them to prioritize them based on their own needs.

The result of comparing the priorities in the needs of the users with those in our method using Kendall's W^5 is shown in **Table 7**. As shown by the value of W (0.828), there is a high agreement between the priorities, which means that the priorities established by our method

⁵ Kendall's W (Kendall's Coefficient of Concordance): By normalizing the statistic, it becomes possible to assess the agreement among raters. It is typically used to compare the consistency between two rankings. Kendall's W ranges from 0 (no agreement) to 1 (complete agreement) [23].

are indeed in agreement with those based on users' needs. As a result, we concluded that our prioritization scheme can calculate the priorities of the customer requirements closer to the users' needs without asking them directly.

6. Conclusions

In software projects, a variety of methods has been studied to understand the customers' needs and to elicit the requirements from the customers. Existing methods, however, have some limitations—they focus on a limited range of target customers and pose a limited expression of customer's opinions; further, it is difficult to collect the opinions of customers continuously. Therefore, we propose an SNS-based CRE. By utilizing an SNS that encompasses a larger amount of information and a greater number of users, this study shows how to access SNS effectively in order to elicit customer requirements and refine them.

Our proposed method facilitates access to a wider range of customers and enables the unlimited expression of customers' opinions without using a predefined format; it also allows us to collect customer opinions continuously. With these advantages, it enables us to reflect the diverse requirements of new perspectives to software products by gathering uninhibited opinions from the users. Moreover, SNS-based CRE is a useful method to access users, especially when specific users are not identified in a new market or when there is a need to collect novel opinions from undefined users or when the desired features of the users in the general software domain are needed.

A case study and an evaluation were carried out to validate the SNS-based CRE method and demonstrate its usefulness. The case study on the smartphone domain was conducted through Facebook and Twitter, based on the proposed method. As a result, 67 customer requirements were elicited, and it was possible to elicit a wider range of customer requirements than that elicited by the existing method. We also conclude that the three advantages of the SNS-based CRE are valid. Moreover, the calculated priorities of the customer requirements with our method considerably correspond with the needs of the users.

Future studies should focus on improving opinion refining techniques, developing formal methods to motivate social network members to participate more efficiently in the proposed method, and exploring the applicability of our method to an even larger number of users and to a more linguistically diverse environment to make the proposed method more effective and rapid.

References

- [1] Karl E. Wiegers, "Software Requirements 2nd Edition," *Microsoft Press*, pp.21-33, 2003.
- [2] O. Dieste, N. Juristo, F. Shull, "Understanding the Customer: What Do We Know about Requirements Elicitation?," *IEEE Software*, vol. 25, no. 2, pp. 11-13, Mar.-Apr. 2008. [Article\(CrossRef Link\)](#)
- [3] "CHAOS summary 2009," The Standish Group, 2009.
- [4] K. Aoyama, T. Ugai, S. Yamada, A. Obata, "Extraction of Viewpoints for Eliciting Customer's Requirements based on Analysis of Specification Change Records," in *Proc. of 14th Asia-Pacific Software Engineering Conference*, pp. 33-40, Dec. 2007. [Article\(CrossRef Link\)](#)
- [5] J.-B. E.M. Steenkamp, M. G.de Jong, Hans Baumgartner, "Socially Desirable Response Tendencies in Survey Research," *Journal of Marketing Research*, vol. 47, no. 2, pp. 199-214, 2010. [Article\(CrossRef Link\)](#)
- [6] D. Kim, K.-P. Lee, J. Lee, "Design Needs Research of Social Network Service Development," *Korean Society of Design Science*, pp. 24-25, May, 2009.

- [7] O. Kwon, Y. Wen, "An Empirical Study of the Factors Affecting Social Network Service Use," *Journal of Computers in Human Behavior*, vol. 26, no. 2, pp. 254-263, Mar. 2010. [Article\(CrossRef Link\)](#)
- [8] "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210-230, Oct. 2008. [Article\(CrossRef Link\)](#)
- [9] DMC Media, "SNS users around the world," <http://www.dmcreport.co.kr>, 2010.
- [10] Tweetreports, "The Twitter Clock: Real-time Twitter Statistics," <http://www.tweetreports.com>, 2011
- [11] CheckFacebook.com, "Global Audience," <http://www.checkfacebook.com>, 2011
- [12] eMarketer, "Types of Online Activities Performed Daily or Ever, Sep 2010 (% of internet users worldwide)," <http://www.emarketer.com>, 2010
- [13] George Colony, "Social Sigma Rising," <http://blogs.forrester.com>, 2010.
- [14] Ian F. Alexander, Richard Stevens, "Writing Better Requirements," *Addison-Wesley Professional*, pp. 31-39, 90-95, 2002.
- [15] F. Scheuren, "What is a Survey?," <http://www.whatisasurvey.info>, 2004
- [16] K. Holtzblatt, S. Jones, "Contextual Inquiry: A Participatory Technique for System Design," *Participatory Design: Principles and Practices*, pp. 177-209, 1993
- [17] Yuri, "How to do keyword research and learn what your customers need," <http://www.improvetheweb.com>, 2007.
- [18] Glen Parker, Lindsey Thomas, "Wave 6 – The Socialization of Brands," UM EMEA research team, 2010.
- [19] T. Opsahl, F. Agneessens, J. Skvoretz, "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths," *Social Networks*, vol. 32, no. 3, pp. 245-251, July, 2010. [Article\(CrossRef Link\)](#)
- [20] T.W. Valente, R.K. Foreman, "Integration and Radiality: Measuring the Extent of an Individual's Connectedness and Reachability in a Network," *Social Networks*, vol. 20, no. 1, pp. 89-105, Jan. 1998. [Article\(CrossRef Link\)](#)
- [21] H. Kwak, Changhyun lee, Hosung Park, Sue Moon, "What is Twitter, a Social Network or a News Media?," in *Proc. of the 19th international conference on World wide web*, pp. 591-600, 2010. [Article\(CrossRef Link\)](#)
- [22] L. Page, S. Brin, R. Motwani, T. Winograd, "The Pagerank Citation Ranking: Bringing Order to the Web," Stanford InfoLab, 1998.
- [23] M. Kendall, "A New Measure of Rank Correlation," *Biometrika*, pp. 81-93, 1938.



Yoon-kyu Lee is an M.S. candidate with the Department of Computer Science at Korea University, Seoul, South Korea. He received his B.S. degree from the College of Information and Communications, Korea University, in 2010. His research interests include social network, requirements engineering, value-based software engineering, and embedded software engineering.



Neung-Hoe Kim is a Ph.D. candidate with the College of Information and Communications, Korea University. His research interests include requirements engineering, value-based software engineering, software engineering economics, and embedded software engineering. He received his M.S. degree in Computer Science from Korea University.



Dohoon Kim is a Ph.D candidate in the Department of Computer Science and Engineering at Korea University in Seoul, Korea. His research interests include network security, risk management, software engineering, and forecasting engineering. He received his B.S. and M.S. degrees in Computer Science and Engineering from Korea University in 2005 and 2007, respectively.



Dong-hyun Lee is a Ph.D. candidate with the College of Information and Communications at Korea University. His primary research interests include software engineering on embedded systems, value-based software engineering, and ubiquitous computing. He received his M.S. degree in Computer Science from Korea University.



Hoh Peter In is the Vice Dean of the College of Information and Communication at Korea University in Seoul, South Korea. His primary research interests are embedded software engineering, social media platform and service, and software security management. He earned the most influential paper award for 10 years in ICRE 2006. He has published over 100 research papers. He was an Assistant Professor at Texas A&M University. He received his Ph.D. in Computer Science from the University of Southern California (USC).