Ruijia Zha

IMPROVING THE USABILITY OF MOBILE APPLICATION USER REVIEW COLLECTION

Faculty of Information Technology and Communication Sciences
M. Sc. thesis
February 2019
User reviews play an important role in the contemporary mobile application industry. Users, developers and application marketplaces can all get benefit from user reviews. The existing methods of mobile application user review collection are not convenient to use for users. Meanwhile, users express their feedback in text on the heterogeneous level of quality, which brings difficulty for application producers to extract needed information for software maintenance and evolution. This thesis aims to improve the usability of mobile application user review collection.

This thesis introduces the existing review collection methods, and analyzes the usability of these methods. In addition, by applying user-centered design method, the author develops a review collection tool named Reviewer app, and compares it with the existing review collection methods in the usability evaluation.

The study shows that the Reviewer app has significant improvement in the effectiveness and efficiency of review collection. It can help users express their thinking precisely and explicitly, and can lead users to reach the reviewing page quickly and directly. The design could be implemented in the actual application marketplaces.

Keywords: review collection, mobile application, usability, user-centered design

The originality of this thesis has been checked using the Turnitin Originality Check service.
Table of Contents

1. Introduction ........................................................................................................ 1

2. User Review ......................................................................................................... 4
   2.1 What is a User Review? ............................................................................... 4
   2.2 Mobile Application User Review ............................................................... 6
   2.3 Types of Mobile Application User Reviews ........................................... 8
      2.3.1 Bug Reporting ....................................................................................... 8
      2.3.2 Feature Request ..................................................................................... 9
      2.3.3 User Experience .................................................................................... 9
   2.4 User Review Collection Methods .............................................................. 10
      2.4.1 Online Feedback Forum ....................................................................... 10
      2.4.2 Bug Tracking System ........................................................................... 11
      2.4.3 Website Feedback Channel .................................................................. 11
      2.4.4 Review Channels on the Application Marketplaces .............................. 12
      2.4.5 Built-in Review Channel ...................................................................... 13
   2.5 Summary ................................................................................................... 14

3. Consideration of Usability in User Review Collection Methods ...................... 16
   3.1 Usability ..................................................................................................... 16
   3.2 Discussion of Usability in User Review Collection Methods .................... 19
      3.2.1 Effectiveness ....................................................................................... 19
      3.2.2 Efficiency ............................................................................................ 20
      3.2.3 Learnability ......................................................................................... 20
      3.2.4 Errors Covering ................................................................................... 20
   3.3 Summary ................................................................................................... 20

4. Design the Reviewer app ................................................................................... 22
   4.1 Background of the Reviewer app.............................................................. 22
   4.2 User-centered Design ................................................................................. 22
      4.2.1 Principles of User-centered Design ...................................................... 23
      4.2.2 User-centered Design Process .............................................................. 24
   4.3 Design of the Reviewer app ....................................................................... 26
      4.3.1 Implementation of User-centered Design ............................................. 26
      4.3.2 Design of the Reviewer app ................................................................. 29

5. Usability Evaluation .......................................................................................... 33
   5.1 Objectives of Usability Evaluation ............................................................... 33
5.2 What is the Usability Evaluation? ............................................................ 33
5.3 Participants of Usability Evaluation ......................................................... 34
5.4 Procedures of Usability Evaluation .......................................................... 35
  5.4.1 Tasks ................................................................................................... 36
  5.4.2 Questionnaire ...................................................................................... 38
  5.4.3 Interview ............................................................................................. 39

6. Discussions ......................................................................................................... 41
  6.1 Results ........................................................................................................ 41
    6.1.1 Result of Questionnaire .............................................................. 41
    6.1.2 Answers to Interview Questions ............................................... 42
  6.2 Results Analysis ......................................................................................... 43
    6.2.1 Analysis of Results .............................................................. 43
    6.2.2 Potential Improvements .............................................................. 44
  6.3 Limitations ................................................................................................. 44

7. Conclusion ......................................................................................................... 45

References ................................................................................................................. 46

Appendix 1: Consent Form ...................................................................................... 53
Appendix 2: Questionnaire ....................................................................................... 54
1. Introduction

During the recent years, the smartphone has become an essential part of contemporary life. Hundreds of new applications are released every day in application marketplaces, which are usually operated by the owner of the mobile operating system and where users can download applications. In addition to finding a desired application, users can also leave their reviews of the application. The reviews include praise, complaint, bug reporting and even any kind of feeling, and provide valuable information to both application developers and users [Vasa et al., 2012]. From the developers’ perspective, the reviews could provide good suggestion on design of functions as well as a direct way about promoting and updating their products constantly. From the users’ perspective, the reviews can be seen as references when other users making decision if to buy the application or not. For example, a negative review may reveal the limitations of the application. In addition, Fu et al. [2013] addressed that market operators such as Google Play [Google Play 2019] and Apple App Store [Apple App Store 2019] can benefit from the user reviews, which can ensure safe and quality content and steer the market towards greater prosperity by collecting and analyzing the reviews of different categories of mobile applications. Therefore, the user, developer and application marketplace can all get benefit from user reviews.

In the modern application marketplace, the most pervasive way of leaving reviews is through the review channel. Having bought and used an application, users are able to provide feedback or express their opinion on it. Users can reach the review page by searching the specific application. In addition, some applications have their own channel to collect user reviews. In general, a user may proceed with many steps to reach the reviewing page to write down the review in text. Therefore, users are likely to give up the reviewing tasks due to the effort and time spent on reaching the reviewing page. Furthermore, it is difficult for users to express their opinion explicitly only in text, which leads to the difficulty for application producers to extract accurate information from textual reviews. To deal with these problems, this thesis aims at developing a mobile application user review collection tool with improved usability.

Usability is defined by the ISO 9241 [1997] as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." The main properties include effectiveness which concerns the result of interaction between the user and system, efficiency which pays attention to the process of the interaction, and satisfaction which is about the feeling from users about their use of the system [Abran et al. 2003]. In the context of a user review collection tool, the effectiveness is the validity and availability of collected reviews. The validity emphasizes that the user is able to give a review correctly reflecting
his or her opinion on the application. The availability means the ability to make review and related information accessible as needed by application developers. Moreover, the efficiency is the agility of user review collection process. The agility is an ability to move quickly and flexibly, and the agility of review collection requires to simply steps and operations. Obviously, the usability of current review channels, operated by either application marketplaces or application producers, still have room for improvement.

This research aims to design the function of mobile application user review collection with high usability, which ensures both the validity of collected reviews and the convenience for the user to offer reviews. This thesis tries to answer the following question:

- What are the usability issue identified in the current user review channels?
- How can we improve the above identified issues?

As the overall goal of this research is to improve the usability of user review collection function, the approach should be understanding the usability issues in the current review channels and design a new method of review collection. User-centered design (UCD), originally coined by Dr. Donald Norman [1986], is applied in the study. UCD is a design method that designers pay attention to user characteristics, environments, tasks and workflow of a product at each step of the design process, to ensure the high usability of the product. Additionally, in order to concentrate on the user review collection function itself and facilitate the research, the author developed a demo application called Reviewer app, which only contains the function of collecting user reviews. UCD is applied in the design process of the Reviewer app. After the demo has been published, the usability of the application is evaluated. The study is to compare the usability of the existing review collection with the Reviewer app. Participants are required to complete the tasks of providing reviews of given applications, answer a questionnaire and attend the interview during the evaluation process. The evaluation results show the improvement of user review collection process. With additional effort, such a design can be applied in practice by marketplaces or application producers.

This thesis includes seven chapters. Chapter 2 starts with an introduction to user reviews and mobile applications’ user review. User reviews are classified into different types on the basis of the review content. This chapter also studies several popular user review collection methods. Chapter 3 discusses issues related to usability, which is one of the most important attributes for software products. The usability attributes of the existing review collection methods are analyzed and discussed. Chapter 4 covers the design process including the implementation of UCD and the working flow of the design. Chapter 5 continues with the evaluation of the review collection tool. The usability
evaluation includes tasks, questionnaire and interview. Chapter 6 discusses the results, including the analysis and potential improvements of the tool, as well as the limitations of this study. Chapter 7 draws a conclusion of the thesis.
2. User Review

This chapter introduces the definition of the user review and the existing mobile applications of user review. User reviews are classified into different types according to their content. This chapter also addresses the different ways of user review collection.

2.1 What is a User Review?

The definition of review by Cambridge Dictionary [2019] is:

“A report in a newspaper, magazine, or programme that gives an opinion about a new book, film, etc.”

The review is given to a published product or public service such as a book (book review) and a film (movie review), and is often presented in the public media. In addition, an opinion or feedback written by a user or consumer for a product or a service is regarded as a user review. Users may provide their reviews from different viewpoints including not only praise and criticism toward the object to be reviewed, but also the reports of users’ experience in a particular context [Pradhan et al. 2016; Guzman et al. 2015; Bakiu and Guzman 2017]. In addition, user reviews are commonly given voluntarily.

In software industry, user reviews play a pivotal role in different development methods. Agile development methodologies, which encourage rapid and flexible responses to changes [Alliance 2015], advocate a variety of practices for the constant user review on requirements, technical decisions and management constraints. Specifically speaking, the dynamic systems development method (DSDM) recommends short-cycle user prototyping, when SCRUM features end-of-iteration reviews with user focus groups [Highsmith and Cockburn 2001]. Consequently, the good collection and use of user reviews will facilitate software development.

The user review is rolling like online Word-Of-Mouth (WOM), which is recognized as influential in information transmission [Godes and Mayzlin 2004; Vasa et al. 2012]. Davis [1989] presented the perceived usefulness and perceived ease of use and explained why users accept or reject information technology. Perceived usefulness refers to “the degree to which a person believes that using a particular system would enhance his or her job performance”; perceived ease of use refers to “the degree to which a person believes that using a particular system would be free of effort”. The user review could be regarded as the communication intermedium that reflects quality perceived by users. Vasa [2012] indicates that the user review gives users a chance to publicize and promote the applications they like, and allows them to warn the others about the possible problems, which creates a positive review loop in the following software development iterations.
The structures of user reviews for different types of software are similar [Vasa 2012]. The main body of the reviews is written in textual feedback. Most application marketplaces require reviewers to give the numerical ratings at the same time. Both the ratings and textual reviews are visible to the public including other users and developers. The reviewing dates and times are recorded automatically and presented together with the reviews. Meanwhile, the username of the registered user will be shown as well. In addition, as the new version of the application is released at intervals, the version information of the application from the device of the user is also saved. Some web-based applications also require the users to leave contact information, for the purpose of acquiring more detailed feedback from reviewers. Many websites only enable registered users to leave reviews.

A user review is a form of co-value creation. The users can add values in the following three perspectives: [Tan and Vasa 2011]:

- User reviews are accessible to both the application producers and the application marketplaces.
- The reviews, such as praise and recommendation, improve the chance of an application being discovered and downloaded by the potential users. A large number of good reviews can encourage the developers.
- The reviews, such as warnings and critics to the specific features, inform the application producers about where they need to pay attention to, and also warn other users about the possible limitation of the application.

As shown in Figure 1, the user, developer and application marketplace can all benefit from this value-adding process. From the perspective of developers, the reviews provide good suggestions on guiding the future maintenance and evolution work [Li et al. 2018] as well as a direct way to receive bug reporting. In order to fix bugs, implement new features and improve user experience, mobile application developers tend to collect timely and constructive feedback from users [Chen et al. 2014; Pagano and Bruegge 2013].

From the perspective of users, they are able to warn others about the limitations of the application by contributing reviews, which can be seen as the references when other users making a decision regarding whether they buy the application or not. Feedbacks offered by previous buyers affects the relationship between the seller and potential customers [Woodside and Delozier 1976].

From the perspective of application marketplaces, the results of statistical analysis about the user reviews are helpful to evaluate and rank the applications. For example, the applications with a mass of praise, recommendation and higher ratings also rank higher in “hot lists”, which could increase the exposure rate and the number of downloads.
A mobile application is a software artifact that is specifically developed for handheld devices such as a smartphone, tablet or watch [Hoehle and Venkatesh 2015]. Mobile application development has undergone an exponential growth since mobile application marketplaces opened in 2008 [Wasserman 2010]. App Store and Google Play are two official mobile application marketplaces for the IOS and Android mobile operating system respectively. Harman et al. [2012] indicate that mobile application marketplace is a new form of software repository and vastly different from traditional ones. Applications are available to download through marketplaces either free of charge or at a cost. By the end of 2017, over 3.5 million and 2.1 million applications were published on Google Play and App Store respectively. The growing popularity of marketplaces, the ease of sale and convenience of deployment, as well as the huge communities of registered users make them very attractive and admirable for mobile application producers [Pagano and Maalej 2013].

The original intention of mobile application development is an extension to basic functions available in a mobile device, such as calendar, E-mail, clock and contacts. With the technical improvement of the mobile device, more exclusive functions such as shopping, video playing, gaming, and location-based services, are on demand. That is the reason why millions of applications available nowadays. As mobile applications become more complex, numbers of mobile application producers established, it is essential to apply software engineering processes to assure the high-quality of produced mobile applications [Wasserman 2010]. However, the traditional software development is cumbersome to support the mobile application development [Konig-Ries 2009]. The mobile applications mostly have more frequent releases than traditional software. Normally, the popular mobile applications such as Facebook [2019] and YouTube [2019] are updated nearly every month. In addition, the ranking mechanism in application marketplaces leads to a better customer accessibility but also the inevitable competition.
among application producers [Holzer and Ondrus 2011]. The ranking mechanism also intensifies the competition that diversified demands of users shall be reached in varying situations [Li and Zhang 2015]. In this fiercely competitive environment, the user review enables the application producer to acquire feedback timely and quickly from user. Therefore, the user review plays an irreplaceable role in the mobile application industry. Similarly, the mobile application user, developer and vendor, and marketplaces can all get benefit from the mobile application user review.

Figure 2 Screenshots of user reviews for TikTok

Figure 2 shows three screenshots of user reviews for the application named TikTok[2019] on Google Play. Figure 2 (a) shows several reviews given by users of the TikTok. The reviews contain information such as usernames of the reviewers, numerical ratings, reviewing dates and textual reviews. Among them, username is given by the registered user name or shown as the platform user such as “Google user” if the reviewer chooses to be anonymous. In addition, the date of review is stamped by the system automatically. The key points of a mobile application user review are textual review and numerical ratings. As shown in Figure 2 (b), users who are browsing reviews can view the edit history of each review if there is any, and can also give reviews the “thumb up” or mark them as “unhelpful”, “inappropriate” or “spam”. This function is helpful to filter out useful reviews from such a big number of them. Figure 2 (c) shows the overall rating, top feature and review highlight. The overall rating includes average rating, rating volume and rating distribution which respectively represents the average star rating, the total number of ratings and the distribution of the ratings that the application has received across all ratings submitted. The top feature is a list of the best features the application has, approved by users who give it a thumbs-up or a thumbs-down. The review highlight
provides a quick glance at the most popular themes or topics in the application’s reviews. The review function presented above build a channel for the mobile application marketplace. Both users and application producers can gain information through the review channel.

### 2.3 Types of Mobile Application User Reviews

As the mobile application marketplace does not let users clarify the types of information written in the user review, no category has pre-defined officially. Researchers identify different types of review information, including praise, helpfulness, feature information and shortcoming [Pagano and Bruegge 2013]. According to the information included, Maalej and Nabil [2015] categorize user reviews into bug reports, feature requests, user experiences and ratings. In the research of Panichella et al. [2015], for the purpose of application maintenance and evaluation, user reviews are classified as information giving, information seeking, feature request and problem discovery. Guzman et al. [2015] define seven categories for user reviews: bug reporting, feature strength, feature shortcoming, user request, praise, complaint and usage scenario. This thesis adopts the user reviews classification of three types: bug reporting, feature request and user experience.

#### 2.3.1 Bug Reporting

Bug reporting is the review that describes issues with the application or unexpected behaviors [Panichella et al. 2015], such as display misplacement, faulty behavior and application no response or crash. Users usually describe those issues in a few sentences. For example,

> “Everything is so good the only problem is that when I use Tiktok for after 10-15 minutes my phone warm up so easily and gets really hot that makes it not enjoyable to use.”


This is a review for Tiktok which reports the problem that the mobile device heats up after a period of use. Normally, bug reporting can help application producers to discover errors and bugs as soon as possible. However, the textual reviews only contain the description from users, which is not enough for developers and analysts of the application to solve the problems under certain circumstances. Developers can hardly use user feedback without context information [Bettenburg et al. 2008]. As presented in the review above, the reviewer intends to report the issue, but there is no way of collecting and presenting any contextual information related to the issue, such as model and brand of mobile device or how much the battery life is left when the issue occurs. Although the
problem was reported on the review channel, it is still difficult for developers to understand the situational environment in which the issue occurs. Therefore, it is helpful for application developers and analysts to collect the feedback with contextual information including the mobile device brand and model, network status, remaining battery time and the length of use on the application.

2.3.2 Feature Request

The feature request contains the suggestions on how to improve the application by adding new features in the future release [Maalej and Nabil 2015]. Users are likely to point out the missing functionalities or content by comparing with other applications in the same genre. In addition, users also ask for improvement of the existing features [Guzman et al. 2015].

“Thank you for providing 1.75x speed option. I requested this option in previous review and you have provided the option in this update. Also I want 3x and 4x speed options. Please consider also. Second point is that I’m not interested in the videos I’ve already watched. Please don’t recommend them in feed of home page or as a suggestion because I like the videos which I find useful and save in my Playlist for later viewing if required and dislike which I don’t find useful.”


This is a typical review of feature request for YouTube which contains a couple of requests, such as adding more options to the video playing speed, stopping recommending the watched videos, etc. Some of the requested features are very subjective, such as the speed request above. Nevertheless, this type of reviews may offer creative ideas which are valuable to the application developers.

2.3.3 User Experience

User experience, including mostly praises or complaints, reflects the experience of users with the application or the specific features in certain situations [Maalej and Nabil 2015]. The emotional expression such as aggressive words and sentences are somehow used.

“This is a very easy and useful app. Good for communicating and it’s the best social network ever!”

“This app is very poor. Can’t stop intrusive ‘people you may know’ messages. Sharing memories format has changed and can’t be reverted. I have removed this rubbish from my phone.”
Listed above are two reviews for Facebook which express the opposite attitudes. In the first one, the reviewer conveys his or her appreciation and satisfaction to Facebook. In contrast, the reviewer expresses his or her anger. Unlike the bug reporting and feature request, the user experience always includes evaluation of the application or its features. The application marketplaces can extract useful information such as keywords and top features by analyzing this type of reviews. Additionally, for the potential users, reading user experience from other users can also have an impact on their purchasing decision [Vasa et al. 2012].

2.4 User Review Collection Methods

To obtain and digest user reviews in an effective and efficient manner forms a challenge for many application developers [Chen et al. 2014]. Online feedback forums and bug tracking systems are used to collect and extract user reviews in the traditional software development. As mentioned in Section 2.2, Apple App Store and Google Play offer a much easier way for users to rate and post reviews for the applications. In addition, many application producers own the built-in functionality of user review collection in their products.

2.4.1 Online Feedback Forum

Before the application marketplace appeared, many companies maintain online feedback forums through their websites to collect reviews from their customers [Lee and Lee 2006]. Users exchange information and experience, learn about the software company and its products or services through feedback forums [Catterall and Maclaran 2002]. For example, as a digital distribution platform for purchasing and playing video games, Steam [2019] is operating its own forum (shown in Figure 3). Moreover, the official staff from Steam will join the discussion, gathering the helpful ideas and answering the doubts. The online feedback forum is beneficial to users and software producers that they can interact with each other directly. However, maintenance and operation of the forum cost a lot of time and money that only the large software company can afford. The review channel supported by the application marketplace can be regarded as an upgraded and incorporative version of the original online forum.
2.4.2 Bug Tracking System

Bug tracking systems play a crucial role in many software projects. Users can communicate with developers to let them know about erroneous behaviors or performance issues as well as to request new features. In addition, through bug tracking systems, developers are able to not only track unresolved bugs which could be continuously tracked, but also request more information from users [Just et al. 2008]. For example, Usersnap [2019] is a bug tracking system which can be embedded into any web-based application, enabling web developers to have better quality assurance process in web projects. This system can track the web-based application automatically, reporting bugs when any occurring. Feedback tool is also included in Usersnap, where users can send their comments to the application producers directly.

2.4.3 Website Feedback Channel

A website is a client-server computer program that the client runs on a web browser on either mobile devices or computers. Plenty of websites have their own feedback channel, which can be used to collect user reviews. As shown in Figure 4 (a), on the pages of Tampere University [2019] the entrance of feedback channel can be seen on the right side of the screen with “Feedback” button. After entering the feedback channel, as presented in Figure 4 (b), users can write textual reviews and leave their e-mail address. Feedback channels on popular websites, observed by the author, are nearly the same design: the entrance is on the fixed position on each web page, with only textual reviews can be written on the floating window or linked web page. Some feedback channels ask the reviewer to leave contact information, and some only open to logged in users of their websites.
2.4.4 Review Channels on the Application Marketplaces

As mentioned above, Google Play is one of the most popular mobile application marketplaces for Android mobile operating system. Users can only give reviews of the applications they downloaded and installed from Google Play. The steps of posting a review can be described as follows:

a) The user reaches the target application from the searching box of Google Play. The target application can also be discovered from “My apps” column. And the entrance of review channel, which named “Rate this app”, is positioned in the middle of the application page.

b) As shown in Figure 5 (a), a user starts the review with rating the application with numerical stars from at least 1 star to at most 5 stars.

c) After the star rating, the user continues to answer three single choice questions relevant to experience about using the target application. For example, the question in Figure 5 (b) is “can you upload audio with this app”, which investigates if the audio uploading feature works normally. Answers to these questions are helpful for Google Play and application producers to understand the status of the application in operation and investigate specific issues of the application.

d) As shown in Figure 5 (c), the last step of reviewing is to write a short review in textual form, which is the most informative part of the whole reviewing process. The textual review will be analyzed and extracted by application producers and Google Play to generate keywords, acceptance of users and massive helpful
Reviewers are not required to complete precedent step before proceeding with the next step. Besides submitting the star rating, a user can skip the rest reviewing steps. After submitting the review, users are able to modify star ratings, their answers to the single choice question and textual reviews.

![Figure 5 Screenshots of Google Play review channel](https://via.placeholder.com/150)

**Figure 5** Screenshots of Google Play review channel

### 2.4.5 Built-in Review Channel

In order to better meet the needs of millions of users, many applications such as Facebook, YouTube and WeChat [2019] have their built-in channels for user review collection. Users can find the entrance of the review channel usually on the function list or on the page of settings. Several selection and page switches are needed before reaching the reviewing page, and the complexity of this process varies in different applications. Finally, users can write their feedback on the reviewing page and send them to the application producers.

YouTube is taken as an example to illustrate how to give a review through its built-in review channel.

a) As shown in Figure 6 (a), the button named “Help & feedback” can be found on the function list on the home page of YouTube.

b) Selecting “Send feedback” on Figure 6 (b), reviewers will be led to the feedback page of YouTube. This requires extra time from users to make selections and wait for the page jumps.

c) As shown in Figure 6 (c), the optional contact information, compulsory textual
review and an optional screenshot are comprised of the feedback to YouTube. The screenshots on feedback page of YouTube has been generated automatically, while some applications ask reviewers to upload pictures or screenshots by themselves.

Figure 6 Screenshots of YouTube review channel

However, once the feedback has been sent to the application producers, the reviewers have no chance to modify their feedback anymore. If a reviewer intends to make modification, he or she needs to repeat the whole reviewing process and send new feedback. Moreover, the process of the step b) varies from different applications. WeChat asks reviewers to choose the type of issues and scenarios many times before reaching the feedback page. In addition, YouTube does not give any notification when taking the screenshot. Even though the screenshot is enclosed within the feedback automatically, it cannot be seen by the reviewer until reaching the feedback page. It is high possible that the reviewer does not want it.

2.5 Summary

A mobile application user review is the opinion given by a user to a specific mobile application. A review includes the textual description, screenshot, and star rating. In addition, the contact information and other automatically generated information such as date and time are optional. User reviews can be classified into bug reporting, feature requests and user experience. Different information expressed in different types of reviews. Online feedback forum, bug tracking system, website feedback channel, Google
Play review channel and built-in review channel are the five commonly used user review collection methods. Google Play review channel and built-in review channel are the most common methods for collecting user reviews of mobile applications.
3. Consideration of Usability in User Review Collection Methods

For the purpose of improving the usability of mobile application user review collection, this chapter discusses usability, which is an important quality attribute of the software applications. The usability attributes of the existing methods will be analyzed and discussed in this chapter.

3.1 Usability

ISO 9241 [1997] defines the usability as:

“The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.”

In this definition, effectiveness, efficiency and satisfaction are the main properties of usability. Effectiveness means the extent of completeness to which a goal or task is accomplished. Jordan [1998] gives an example of the machine operator which produces 100 components per day. If it is able to produce 80 components per day, the effectiveness level of the operator is 80%. Efficiency refers to the amount of effort required to achieve a goal. The less effort required, the higher the efficiency is [Jordan 1998]. Efficiency can also be regarded as resources expended to achieve the goal [Harrison et al. 2013]. Satisfaction is a more subjective aspect of usability. It is the level of acceptance and comfort that users feel when using a product [Jordan 1998]. In addition, the user, goal and context of use are regarded as three factors of usability in this definition, and it clarifies that usability is a property of interaction among a product, a user and the task.

Nielsen [1994] proposed a model of usability including five attributes, including learnability, efficiency, memorability, errors covering and satisfaction. Learnability means that it should be easy for user to know how to use the system and get tasks done with it rapidly. In addition, as the first experience with a new system when learning to use it, learnability is the most fundamental usability attribute. Efficiency means once the user has learned about the software, they can quickly perform tasks. The systems with good efficiency enables a high level of productivity. Memorability means the user should not have to spend long time to learn the design again even when not using the software for some time. The system should become sufficiently memorable to users once they have learned it. Errors mean the system should be able to recover from most errors resulted from user’s error operating. The users should make few errors during the use of the system, and once the error occurs, it could be returned by the operations. Satisfaction means that the system should make the user feel pleasant when using it. The system needs to win the subjective preference of users through a good design. At the same time, utility is taken as
a separate attribute. Nielsen [1994] believed that usability is associated with utility, and they should be considered together during the software design.

Currently, users have a high demand for mobile applications with the increasing development of mobile devices. Considerable differences influence the design and function of mobile applications that set it apart from the desktop environment [Wisniewski 2011]. Therefore, designing the application with high quality is the key to succeeding in business competition. Usability is a quality attribute of mobile applications, when improving it is one aspect of quality enhancement [Kabir et al., 2017]. However, the definitions of usability above are derived from traditional software design. Mobile application design is quite different from a traditional desktop application. Compared to previous desktop software and web, development and evolution of smartphones have brought a new carrier with special characteristics [Rahmat et al. 2015]. Some issues from the advent of mobile devices are summarized as follows:

- Using the mobile application in different contexts, which contains the information to characterize the situation of use [Li and Zhang 2015], such as location, people and objects nearby, network status such as the strength of signal and speed of data transfer, and other environmental elements, may distract user’s attention and largely affect the performance of mobile application [Zhang and Adipat, 2005].

- A physical constraint of the small screen is not able to load the large body of content with information, site links, pictures and long text [Adipat et al. 2011]. Not all the information and services offered by designers and developers are core, some are ancillary. Mobile device users will neither need nor intend to read all the detailed information [Wisniewski 2011].

- The content for mobile applications is identically replicated from the web-based application, so the unique characteristics of the mobile device have not been taken into consideration [Hoehle and Venkatesh 2015]. When users are immersed in the mobile device or constantly devote their mind to the state of attention, challenges about interacting with small screens, which is quite different with larger screens, emerge at the same time [Wisniewski, 2011]. It cannot be simplified saying that mobile is a minified version of the desktop.

- The emergence of the multimodal mobile application provides new interaction methods for users, and also brings challenges for usability [Zhang and Adipat, 2005].

The usability of mobile applications has been discussed in a range of studies, which employed a variety of attributes for it. Hoehle and Venkatesh [2015] present a summary of previous works using a table including the content such as usability attributes, conceptualization, evaluation techniques and related studies. Some representative studies
are shown in Table 1, and only usability attributes and study columns are listed. In this thesis, the graphical design of user interface is not the point of discussion, and the focus of the research is to study the functional design of mobile application user review collection. In addition, as mentioned before, satisfaction is a very subjective attribute which can be different among users. It is difficult to have a unified standard about satisfaction. Therefore, combining with the repeatability among previous studies, attributes of usability can be summarized as: effectiveness, efficiency, learnability and errors covering.

<table>
<thead>
<tr>
<th>Usability Attributes</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness, efficiency and satisfaction</td>
<td>Hyvarinen et al. [2005]</td>
</tr>
<tr>
<td>Goals (hedonic/utilitarian), emotion, hands, leg, visual, auditory, colocation interaction</td>
<td>Kim et al. [2005]</td>
</tr>
<tr>
<td>General, users were asked to use mobile Internet and complete a usability diary based on the experiences</td>
<td>Zhang and Adipat [2005]</td>
</tr>
<tr>
<td>Learnability, efficiency, memorability, error, satisfaction, effectiveness, simplicity, comprehensibility and learning performance</td>
<td>Nah et al. [2005]</td>
</tr>
<tr>
<td>Satisfaction, effectiveness and efficiency</td>
<td>Cyr at al. [2006]</td>
</tr>
<tr>
<td>Effectiveness, contextual awareness, task hierarchy, visual attention, hand manipulation and mobility</td>
<td>Duh et al. [2006]</td>
</tr>
<tr>
<td>Effectiveness, efficiency and satisfaction</td>
<td>Nielsen et al. [2006]</td>
</tr>
<tr>
<td>Satisfaction, ease of use, ease of learn</td>
<td>Urbaczewski and Koivisto [2007]</td>
</tr>
<tr>
<td>Effectiveness, efficiency, satisfaction, learnability, security</td>
<td>Min et al. [2009]</td>
</tr>
<tr>
<td>Usefulness, enjoyment, ease of use</td>
<td>Kim et al. [2010]</td>
</tr>
<tr>
<td>Design, customer needs, satisfaction, innovativeness, feedback and efficiency</td>
<td>Kim et al. [2012]</td>
</tr>
</tbody>
</table>

Table 1 Prior attributes for mobile application usability [Hoehle and Venkatesh 2015, pp. 438-439]

The effectiveness means the reviews collected through the collection tool should be valid for users and available for application producers. First, the information included should have in a variety of forms such as images, texts and systematical information. These diverse forms are valid for user to fully express their thinking, when the contextual information is collected. On the other hand, the collected reviews contain the combination of various information, which are available for developers to easily analyze and extract.

The efficiency means the whole reviewing process should be brief and convenient, with each step having low complexity. Users will not spend a long time to give reviews, and most likely prefer easy ways to achieve their goals. If the reviewing process is complicated with a lot of page jumps, each jump may lead users to quit the reviewing
task. In addition, the design of the reviewing page should be clear and easy to use, so that the user could review quickly.

Learnability means users should be familiar with the reviewing process relatively quickly. If the review collection tool is difficult to learn, or the entrance not highlighted to be found, the user will stop giving reviews at the beginning.

Errors covering means that issues resulting from erroneous operations should be covered as much as possible. The erroneous operations such as clicking the wrong button by mistake happen commonly during the use of any applications. It will be better if there are tips reminded when the erroneous operation occurs.

3.2 Discussion of Usability in User Review Collection Methods

This thesis focuses on the Google Play review channel and built-in functionality of user review collection to analyze the usability issues identified in these channels.

3.2.1 Effectiveness

Textual reviews alone are not able to fully express the opinion of users. In general, an application is comprised of one or more pages, with every page consisting of various entities, such as buttons, images, texts, progress bars, input boxes and many elements. A feature is visually composed of entities also. For example, the feature of music playing consists of a group of entities: showing the progress of playing; starting, pausing or jumping to the next song; etc.. Under such a scenario, when a user intends to report bugs, request new features or just give comments on a specific page of the mobile application, he or she has to specify the target entities in text. Furthermore, the application producers, who have received the textual reviews, still need to comprehend those descriptions and relocate the according entities. The information transformation of entity related issues can be summarized in Figure 7. Effort is wasted on explaining and comprehending the entities with which a reported issue in the review is related. Practically, most reviews are related to specific entities of applications. It is common that reviewers give praises or complaints to one or more entities of the applications. As mentioned in Section 2.4.1, most built-in review channels allow screenshots to be sent together with textual reviews. This makes it easy to locate the specific pages of issues for application producers. Therefore, in terms of effectiveness, built-in review channels perform well in helping both the users to express their thoughts and the application producers to obtain information. Furthermore, it is more beneficial if the specific entities on the screenshots can also be marked and recorded.
Efficiency

The path to reach the actual review page of Google Play review channel and built-in review channels have been already introduced in Section 2.4.1. When users intend to write reviews through the Google Play review channel, they need to open Google Play and reach the page of target applications first. The similar steps have to be taken before reviewing when using a built-in review channel. Users reach the feedback page through a series of selections and clicks. These preparation steps are really inconvenient when it can result in users’ giving up the reviewing actions. Therefore, in terms of efficiency, both review collection methods require improvement.

Learnability

Google Play review channel performs well in learnability because the entrance of the channel is highlighted which is easy for users to understand how to give the reviews even without guidance. Even the entrance buttons of built-in review channels are not always highlighted. Once users are told where they are, it is also easy to give reviews by following the process. Both review collection methods have good learnability.

Errors Covering

Finally, the ability to cover errors differs significantly with Google Play review channel and built-in review channels. When using the Google Play review channel, reviewers are able to make changes to their reviews after submitting, and the submission history is accessible. However, when using the built-in review channels, reviewers cannot modify or even check the reviews they have given. If reviewers posted an improper review, they cannot trace the review back and have to post a new one which correctly expresses the opinion.

Summary

To better evaluate the usability of user review collection functionality, effectiveness, efficiency, learnability and errors covering are seen as the attributes of usability. The
performances of the two most popular mobile application user review collection methods about those usability attributes are quite different. Google Play review channel has a good ability of errors covering, and performs badly on effectiveness. However, built-in review channels perform poorly in errors covering but better on effectiveness. Both of them have good learnability, when their efficiency should be improved.
4. Design the Reviewer app

In order to make the improvement on the usability of review collection, the functionality of review collection should be redesign. This chapter starts with the background of Reviewer app, introducing the steps of nature design, software design and User-centered design (UCD). This chapter also introduces the design process including the implementation of UCD and the working flow of Reviewer app.

4.1 Background of the Reviewer app

As presented in section 2, the Google Play review channel and built-in review channels are currently the most popular methods of mobile application user review collection. Google Play review channel is operated by the application marketplace, while the built-in review channels are operated by application producers. It is impossible for the market owners and application producers to give permission to the thirty-party developers to make changes on their review channels. In addition, all the collected reviews from built-in review channels are not open to the public; thus it is difficult for elicitation and further analysis. Therefore, in order to concentrate on the user review collection functionality itself and facilitate the research, a demo application is designed. The application is called Reviewer app, which contains the functionality of collecting user reviews. It also can be embedded into different applications and marketplaces as a review channel. The application is developed to run on the Android operating system above version of Android 2.3. The design process of Reviewer app will be introduced in the following sections.

4.2 User-centered Design

Jones [1981] defined design as “a creative activity that involves bringing into being something new and useful that has not existed previously”. Software design is one of the most important procedures in the software development process. Bourque and Fairley [2014] define the software design as “the process of defining the architecture, components, interfaces, and other characteristics of a system or component”. The design is not only the activity after requirements specification and before implementation, but also contains all the activities involved in conceptualizing, framing, implementing, commissioning and ultimately modifying complex systems [Freeman and Hart 2004].

UCD was originally coined by Dr. Donald Norman [1986] at the University of California San Diego. UCD describes a method that has been proposed for decades in different names, such as human factors engineering, ergonomics, and usability engineering [Rubin and Chisnell 2008]. The UCD is also called the human-centered design process, ISO 13407 [1999] states:
“Human-centered design is an approach to interactive system development that focuses specifically on making systems usable. It is a multi-disciplinary activity.”

Norman [1988] built further on the UCD concept and suggested the user should be placed in the center place of the design. The designer should play the role of facilitator to ensure that the user can make use of the product as intended without extra effort on learning how to use it [Abras et al. 2004]. In addition, Rubin and Chisnell [2008] suggest that

“UCD seeks to support how target users actually work, rather than forcing users to change what they do to use something.”

Instead of making users to adapt products, UCD emphasizes the needs and feelings of users in product design, development, maintenance and around user-centered in product design, development and maintenance [Wei and Xing 2010]. Designers must know who the users will be and understand what tasks they will do, and it requires direct communication between designers and users [Shackel 1991].

4.2.1 Principles of User-centered Design

Any system, which contains compulsory functions for people’s work, should be easy to learn and pleasant to use [Gould and Lewis, 1985]. Theoretically, Gould and Lewis [1985] also present the three principles, which are prerequisites of producing a useful and easy to use application, and also regarded as the basic principles of user-centered design:

- **Early focus on users and tasks**: Designers must understand potential users by studying their cognitive, behaviors and even attitudes. The systematic and structured approach is required to collect information from users.
- **Empirical measurement**: Designers should record and analyze the reactions of users toward the products started from the prototypes, which should be insisted throughout the entire development and testing process.
- **Iterative design**: Designing, testing, measuring, and the redesigning should be repeated as a cycle during the process of production. Through early testing of conceptual models and design ideas, the true iterative design allows for the overhaul and rethinking of a design, which is the process of shaping the product [Rubin and Chisnell 2008].

To facilitate design tasks and to guide the designers, Norman [1988] also proposes several advices about user-centered design. First, the designers should use both the knowledge in the world and in the head. The designer could build conceptual models by writing manuals, which are easily understood and should be written before the design is
implemented. Moreover, the structure of tasks should be simplified. Practically, the user is able to remember five things at a time on average. Therefore, the short-term or long-term memory of the user should not be overload, while mental aids shall be provided for easy retrieval of information. In addition, designers should make things visible so that users can figure out the use of an object by seeing the right buttons or devices for executing an operation. Using graphics is an effective way to make things understandable. Furthermore, the design for errors is needed. Planning for any possible errors that can be made, when users’ option of recovery from the error they made will be allowed.

4.2.2 User-centered Design Process

As presented in Figure 8, ISO 13407 [1999] briefly described the activities of user-centered design (also known as human-centered design) as follow:

- **Identify the need for human-centered design**: Understand why the user is going to use the application, and what the user intends to acquire by the use.
- **Understand and specify the context of use**: The user, environment and task of use should be clear, and the contextual information should be collected.
- **Specify the user and organizational requirements**: The success criteria of usability for the product in terms of user tasks and the constraints of design should be determined. In addition, the unique requirements should be specified in advance.
- **Produce design solutions**: Incorporate visual design, interaction design, usability and several important elements into design solutions.
- **Evaluate designs against requirements**: The usability of design is evaluated against user tasks.
- **Satisfy specified user and organizational requirements or start another iteration**: After the evaluation process, it is shown in the result whether the design satisfies the original requirements or not. If all the requirements are satisfied, the design is approved and the user-centered design process can be finished. If not, another iteration of the design process should be activated from the step of understanding the context of use.
Figure 8 Activities of user-centered design [Jokela et al. 2003]

The most important point in UCD is involving users to the design process as well as the development process. Products can be refined through involving users in an interactive iterative process [Abras et al. 2004]. The issues related to the attributes of usability can be addressed by the subjective criteria of users. As shown in Table 2, Preece et al. [2002] suggest ways of involving users in the design and development of products.

Different methods are available in interacting with users on any stage in the design cycle. Once the users have been identified and a thorough investigation of their needs has been conducted by performing tasks and needs analysis, designers can develop alternative design prototypes to be evaluated by the users [Abras et al. 2004]. Collecting the thinking and suggestion of users to the prototypes can help designers discover the advantages and disadvantages of their design and examine whether requirements from users are truly satisfied. In addition, suggestions from users are essential to be considered.
Table 2 Involving users in the design process [Preece et al. 2002]

User involvement can make designers better understand the requirements and produce more a creative design. However, the iteration of UCD is time-consuming and costly, which requires extra work from designers. Therefore, instead of applying UCD methods irrationally, designers can select suitable methods depending on their own situations.

4.3 Design of the Reviewer app

As mentioned in Section 4.1, a demo application named Reviewer app, as a new design of mobile application user review collection, is designed and developed in this research. The UCD approach is applied in the design process.

4.3.1 Implementation of User-centered Design

The overall design process is shown in Figure 9. It is under the guidance of UCD principles and activities mentioned in the previous sections. Taking the time and cost into consideration, two iterations of UCD are applied in the design process. As the usability testing about review channels and the Reviewer app is planned in this thesis, besides the step of usability evaluation, acquiring feedback from users is the connector between the iterations.
Figure 9 Design process of the Reviewer app

As the preparation work, the problems of Google Play review channel and built-in review channel needs to be identified. As mentioned in 3.2, Google Play review channel has a good capability of errors covering but performs badly on effectiveness. However, built-in review channels have a poor capability of errors covering but perform better on effectiveness. Both of them have good learnability, when their efficiency should be improved. Therefore, the Reviewer app should extend the advantages of each review channel and has a better design about efficiency.

In the first iteration, the scenario of use should be defined. To provide users with the real usage scenario is the most important aspect of the design process [Kangas and Kinnunen 2005]. The users are more likely to report a bug or request new features at the first moment encountering the issue, while they are also more willing to share their experience then. Therefore, the scenario is that users give reviews when using the target applications. The second step is to design a way in which the user reviews can be collected. In Google Play review channels, users have to open the Google Play and search the target applications, then reach the reviewing page. In the built-in review channel, users have to reach the reviewing page through the entrance button and page jumps. Obviously, both
of them completely interrupt users’ using the target applications. In addition, as mentioned previously, the review with screenshots can help application producers to locate issues easily. Hence, the screenshot is used as an entry of the reviewing page. When a user encounters an issue and intends to leave a review, he or she can take a screenshot of the current page and share it to the Reviewer app, which lead the user to reach the reviewing page with the screenshot directly. This design shall significantly reduce the steps of review collection. As shown in Figure 10, the main body of the user review, such as the preview of screenshot, rating bar, application name and textual review, are all included in the reviewing page. Moreover, when saving the review, the Reviewer app can also automatically collect the system information including the mobile device brand and model, network status and remaining battery, etc. The analysis and design steps in the iteration one finish with the prototype of Reviewer app.

![Prototype of Reviewer app in iteration one](image)

Thereafter, several users are invited to a workshop where they are asked to give feedback about the prototype of Reviewer app designed in iteration one. Two useful ideas and suggestions are proposed in the workshop:

- This design can only handle one scenario that the user gives review during use. If a user intends to give a review after use, an entrance to upload a screenshot in the reviewing page is needed.
The screenshot can help producers to locate the problematic page, but it cannot help to identify the specific entity on the page. It will be better if the screenshot can be processed that the target entity can be highlighted.

With these two suggestions, the design proceeds to the iteration two. First, the scenario of giving reviews after using the target application is to be specified. Users should be able to upload screenshots by themselves. Thus, as shown in Figure 11 (a), an upload button is added to the main page of Reviewer app. Moreover, following the suggestion about highlighting the entities on screenshots, a button of adding markers is also added to the main page and a page of processing screenshot is added. As shown in Figure 11 (b), in addition to putting markers on the screenshot, users are also able to write textual reviews to each marked entity. In this way, users can give reviews to the specific entities without describing the entities in text.

Figure 11 Prototype of Reviewer app in iteration two

4.3.2 Design of the Reviewer app

The flow of giving review through Reviewer app is given in Figure 12.
There are two entrances of the Reviewer app in different scenarios of use:

- **Scenario one:** The user intends to give reviews when using the target application. As shown in Figure 13 (a), the user should just take a screenshot of the current application, and the sharing button is notified on the top of the screen by Android operating system. The position of the sharing button could be different depending on the different models of mobile devices. Next, as
shown in Figure 13 (b), sharing the screenshot to the Reviewer app, and reaching the main page which is shown in Figure 13 (c).

- Scenario two: The user intends to give reviews with the existing screenshot after using the target application. The user should open the Reviewer app and reach the main page of it directly. Next, upload the screenshot by clicking the “Upload” button which is shown in the middle of Figure 13 (c), and choosing the screenshot from the memory of mobile device.
The results of these two entrances are the same that one screenshot has been recognized by Reviewer app and shown in the main page. The following steps have no difference.

If the user intends to put one or more markers on the screenshot, he or she could click the “Add markers” button and jump to the marker page. As shown in Figure 13 (d), the user adds markers to the screenshot by touching any position of it. The pattern of the marker is a yellow star which is shown on the bottom right corner of Figure 13 (d). At the same time, the user can also write textual reviews to the marking points. The return button on the top right corner enables the user to erase the existing markers. The button with a tick pattern, which is shown on the top left corner of Figure 13 (d), can save the markers and lead back to the main page. If the user does not intend to add any marker to the screenshot, he or she can move to the next step straightly.

The next step is to give the star rating (1 - 5) to the target application. In the last part, the application name and overall textual review should be entered respectively. As shown in Figure 13 (e), after finishing filling on the review page, the user save the screenshot, markers, star ratings, application name and overall review using the “Save” button. At the same time, various contextual information such as the mobile device brand and model, network status, remaining battery time and the length of use on the application, are collected and saved automatically.

The saved review in the database contains ten columns, which are 1) application name, 2) rating, 3) overall review, 4) screenshot, 5) markers, 6) time, 7) battery, 8) network, 9) brand and model, and 10) length of use. First, the columns of application name, rating and overall review are directly saved from the input of users. Moreover, the screenshot is saved using the encoding scheme base64 [Josefsson 2003], which is the format for picture storage and can be decoded to the picture easily. In addition, markers are saved as pairs of coordinates which refers to the specific location on the screenshot, and the reviews to the markers are saved together. Furthermore, the rest columns are the contextual information collected and saved by Reviewer app automatically. The time is the systematic timestamp of the review submission. The percentage of remaining battery and the strength of signal are saved in the columns of battery and network respectively. The brand and model information of the mobile device are saved in the column of brand and model. Finally, the number of minutes used about the reviewed application is saved in the column of length of use.
5. Usability Evaluation

This chapter begins with the introduction of the usability evaluation methodology, several methods are listed. It presents the objectives of usability evaluation tasks and procedures including tasks, questionnaire and interview. The Reviewer app, the Google Play review channel and a built-in review channel are compared through the evaluation of different usability attributes.

5.1 Objectives of Usability Evaluation

The overall goal of this research is to improve the usability of the user review collection tool. Therefore, the objective of the evaluation is to examine whether or not the user review collection prototype provides improved effectiveness, efficiency, learnability and errors covering. The questions to study on each aspect are given below.

- **Effectiveness**: Can the Reviewer App improve the accuracy of information presented in the users’ feedback? Can the useful information be extracted more easily from reviews collected by Reviewer app?
- **Efficiency**: Can the Reviewer app improve the agility of the reviewing process?
- **Learnability**: Can the user learn how to use the Reviewer app easily?
- **Error covering**: Can the Reviewer app cover the most issues resulting from erroneous operation by the users?

5.2 What is the Usability Evaluation?

Rubin and Chisnell [2008] describes the usability evaluation as “a process that employs people as testing participants who are representative of the target audience to evaluate the degree to which a product meets specific usability criteria.” The participants do tasks, the researcher observes and record what participants do and say, and analyzes the data to diagnose issues to study [Dumas and Redish 1999]. The result of the evaluation can provide the information on how the user interface matches the natural human way of thinking and acting as well as highlights the features and processes to be improved [Kaikkonen et al. 2005].

The range of the evaluation process varies from true classical experiments with large sample sizes and complex test designs to very informal qualitative studies with only a single participant [Rubin and Chisnell 2008]. There are currently three types of methods for mobile usability evaluation [Nayebi et al. 2012]:

- **Laboratory experiments**: Under the controlling and observing setting of the laboratory, participants are required to perform specific tasks with specific mobile application. Different aspects can be considered and designed in the lab
and it is possible to record the experiment data in time. The testing environment in the laboratory is isolated and differs from real-world context.

- **Field studies**: Observation and interview are included in this method. Participant is being observed when getting involved in an activity. Researchers take notes and ask questions at the same time. The reality of testing environment is highly closed to the real-world context. On the other hand, insufficient control over users when taking the field study would occur in some case.

- **Hands-on measurements**: This method is to define specific measures for the case in different aspects, and then, to evaluate the usability of the application through those measures directly.

Each approach has different objectives, time and resource requirements [Rubin and Chisnell 2008]. Therefore, the tester should select a suitable method according to the goal, time-consuming and cost of the research.

Several techniques can be employed in usability evaluation [Abras et al. 2004]:

- **Think aloud**: The user is encouraged to articulate all the steps of his or her actions when doing the given tasks.

- **Videotaping**: The whole evaluation process is recorded by a camera, and it is valuable for the tester to review what the participants did, and to discover what the problems are.

- **Interview and user satisfaction questionnaire**: These techniques enable the tester to evaluate the users likes and dislikes about the design and to gain a deeper understanding of any problems.

In addition, the system usability scale (SUS) is a simple survey scale, giving a global view of subjective assessments of usability [Brooke 1996]. This would allow the usability practitioner to assess the usability of the application quickly and easily [Bangor et al. 2008]. The combination of the SUS questionnaire and usability testing technique can make us better understand the usability attributes [Prata et al. 2013]. For example, Kurkovsky and Meesangnil [2012] used a generic questionnaire in the evaluation along with usability testing technique to measure mobile application usability. Furthermore, the customized questionnaire is not only composed of SUS items, but also includes questions focusing on specific aspects of evaluation like design and usage of the mobile, is also a good way to evaluate the usability [Reynoldson et al. 2014].

### 5.3 Participants of Usability Evaluation

The of user review collection tool can be used by any user who has downloaded the application and would like to submit his/her opinion on this application. Considering the factors of time and cost, two participants are invited into the usability evaluation. Both of
them are degree students in the MSc degree programme in software development at University of Tampere. They have taken the course related to the usability study, and have experience on posting reviews of mobile applications through either the application marketplaces channel or the built-in review channel. Therefore, compared to the evaluation from random users, the more professional evaluation can be obtained in this research.

Before the usability evaluation process, all the participants were informed of the propose and procedures of this research and signed a consent form which is presented in Appendix 1. Their privacy is protected, and the data collected are used only in this research.

5.4 Procedures of Usability Evaluation

The evaluation method of this research is designed by adapting the related work on usability evaluation methods with cost taken into account. In order to obtain the objective data from the participants, the questionnaire and interview are applied after the task performing. The questionnaire lets participants evaluate the usability attributes towards different review collection methods. Therefore, the data collected from the questionnaire is easy to be analyzed, which also ensures the result of the usability evaluation is convincing.

As shown in Figure 14, the usability evaluation process contains four steps, including techniques of tasks, questionnaire and interview. To compare the usability of the Google Play review channel, the built-in review channel and the Reviewer app, each participant is required to do same review writing tasks by using the three methods separately. The same questionnaire is given each time after the participate finishes a review writing task using one method. After finishing the tasks by using the provided three methods, the participant is interviewed in step four.
5.4.1 Tasks

In tasks completion, each participant is asked to use the required review collection method to complete a set of review tasks. The reviews collected by Reviewer app will be saved into the memory of the smartphone which will be checked by exporting the data. For the purpose of data safety and convenience, all the participants are required to use the testing smartphone, into which the Reviewer app has been installed in advance. As the Google Play review channel, the built-in review channel and the Reviewer app are being used in tasks respectively. The target application, which will be reviewed during the test, should be able to be reviewed on Google Play as well as have its own built-in review channel. Therefore, YouTube is chosen as the target application.

Before the task, the participants are introduced how to use three review collection methods briefly. The participants are asked to consider what reviews he or she wants to offer to the target application. As mentioned in Section 2.3, users’ reviews are classified into three types in this thesis. In order to make a comparison for different types of user reviews, the reviews given by the participants should only include bug reporting, feature request and user experience. In other words, three pieces of reviews in three types should be created by each participant. The following questions could lead the participants to prepare for their feedback:
• What do you like most in the application?
• What do you dislike most in the application?
• Would you like to have a feature or user interface improved in the application? If yes, what are they?
• Would you like to have other features which are not available in the application? If yes, what are they?
• Is there any problem when you use the application? If yes, what are they?
• What design makes you feeling comfortable when you use the application?
• What design makes you feeling uncomfortable when you use the application?

As these reviews created by each participant will be posted three times through Google Play review channel, built-in review channel and the Reviewer app, the tester should record the reviews and ensure the participant to give the same ones for different review collection methods. In addition, as mentioned in Section 4.4, there are two scenarios of use specified, giving reviews when using the target application and giving reviews after using the target application. Therefore, both of the scenarios should be taken into account.

According to the review types and scenarios of use, four tasks are asked to complete. Tasks 1-3 are designed for different types of reviews in the scenario of giving reviews when using the target application. Task 4 is designed to imitate another scenario that giving a review after using the target application. The type of review in Task 4 can be any of bug reporting, feature request and user experience.

Task 1: The participant is asked to give a review of bug reporting when using the target application. First, the participant should reach the problematic page or entity of the target application. Next, the participant should enter the reviewing page of the review collection method and give the review of bug reporting.

Task 2: The participant is asked to give a review of feature request when using the target application. First, the participant should reach the problematic page or entity of the target application. Next, the participant should enter the reviewing page of the review collection method and give the review of feature request.

Task 3: The participant is asked to give a review of user experience when using the target application. First, the participant should reach the problematic page or entity of the target application. Next, the participant should enter the reviewing page of the review collection method and give the review of user experience.

Task 4: Participant is asked to give a review of any type after using the target application. First, the participant should reach the problematic page or entity of the target application and take a screenshot. Next, the participant should close the target application.
Finally, the participant should enter the reviewing page of the review collection method and give the review.

5.4.2 Questionnaire

After the tasks have been finished, the participants are asked to fill in the questionnaire about usability evaluation. System usability scale is regarded as a reference, and the overall design of the questionnaire in this research is based on forced-choice questions, as shown in Figure 15, where a statement is made and the respondent then indicates the degree of agreement or disagreement with the statement on a 5 point scale.

![Figure 15 Example of a 5 point scale [Brooke, 1996]](image)

The statements in the questionnaire are corresponding to the usability attributes of effectiveness, efficiency, learnability and errors covering. The statements and their related usability attributes are shown in Table 3. The questionnaire used in the evaluation can be seen in Appendix 2.
Table 3 Statements in the questionnaire

<table>
<thead>
<tr>
<th>Usability attribute</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity</td>
<td>I fully expressed my thinking in the review of bug reporting.</td>
</tr>
<tr>
<td></td>
<td>I fully expressed my thinking in the review of feature request.</td>
</tr>
<tr>
<td></td>
<td>I fully expressed my thinking in the review of user experience.</td>
</tr>
<tr>
<td>Availability</td>
<td>The review of bug reporting I gave is concrete and precise.</td>
</tr>
<tr>
<td></td>
<td>The review of feature request I gave is concrete and precise.</td>
</tr>
<tr>
<td></td>
<td>The review of user experience I gave is concrete and precise.</td>
</tr>
<tr>
<td>Convenience</td>
<td>I can reach the reviewing page quickly and easily.</td>
</tr>
<tr>
<td>Complexity</td>
<td>I think this reviewing method has simple steps and is easy to use.</td>
</tr>
<tr>
<td>Scenario one</td>
<td>I can give reviews without interrupting the current use of target application.</td>
</tr>
<tr>
<td>Scenario two</td>
<td>I can give reviews quickly after using the target application.</td>
</tr>
<tr>
<td>Learnability</td>
<td>I can easily figure out how to give review by this method during the brief introduction.</td>
</tr>
<tr>
<td>Errors covering</td>
<td>If I click other buttons by mistake, I can easily return to the reviewing page and continue the reviewing process.</td>
</tr>
<tr>
<td></td>
<td>If I submit the review by mistake, I can edit it and submit the correct one.</td>
</tr>
</tbody>
</table>

5.4.3 Interview

After finishing the tasks and questionnaire, the participants also need to attend an interview. The participants are asked pre-designed questions. The questions are shown in Table 4. Q1 is designed to collect the satisfaction of those three review collection methods. Q2, Q3 and Q4 aim to collect the evaluation and comments on the Reviewer app.
<table>
<thead>
<tr>
<th>NO.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Which review method will you prefer to use in future? Why?</td>
</tr>
<tr>
<td>Q2</td>
<td>Which type of review do you think can be best supported by the “Reviewer app”? Why?</td>
</tr>
<tr>
<td>Q3</td>
<td>Which design of the “Reviewer app” do you think is the best? Why?</td>
</tr>
<tr>
<td>Q4</td>
<td>Do you have any suggestion to the “Reviewer app”?</td>
</tr>
</tbody>
</table>

Table 4 Interview questions
6. Discussions

This chapter begins with presenting the result of the usability evaluation questionnaire and interview. In addition, a comprehensive discussion according to the result is followed, including the analysis and potential improvements.

6.1 Results

The result of usability evaluation consists of the analysis of answers to the questionnaire and the interview questions. These two aspects are presented as follow.

6.1.1 Result of Questionnaire

The result of questionnaire towards the Google Play review channel, the built-in review channel of YouTube and the Reviewer app according to two participants are shown in Table 5. The blue, yellow, green and red rows correspond to usability attributes of effectiveness, efficiency, learnability and errors covering respectively.

<table>
<thead>
<tr>
<th>NO.</th>
<th>Statement</th>
<th>Google Play review channel</th>
<th>Built-in review channel</th>
<th>The &quot;Reviewer app&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Participant 1</td>
<td>Participant 2</td>
<td>Participant 1</td>
</tr>
<tr>
<td>S1</td>
<td>I fully expressed my thinking in the review of bug reporting.</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>S2</td>
<td>I fully expressed my thinking in the review of feature request.</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>S3</td>
<td>I fully expressed my thinking in the review of user experience.</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S4</td>
<td>The review of bug reporting I gave is concrete and precise.</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>S5</td>
<td>The review of feature request I gave is concrete and precise.</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>S6</td>
<td>The review of user experience I gave is concrete and precise.</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>S7</td>
<td>I can reach the reviewing page quickly and easily.</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>S8</td>
<td>I think this reviewing method has simple steps and is easy to use.</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S9</td>
<td>I can give reviews without interrupting the current use of target application.</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>S10</td>
<td>I can give reviews quickly after using the target application.</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>S11</td>
<td>I can easily figure out how to give review by this method during the brief introduction.</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>S12</td>
<td>If I click other buttons by mistake, I can easily return to the reviewing page and continue the reviewing process.</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S13</td>
<td>If I submit the review by mistake, I can edit it and submit the correct one.</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5 Result of questionnaire
6.1.2 Answers to Interview Questions

Q1 Which review method will you prefer to use in future? Why?

Both participants choose the Reviewer app as the preferred way of submitting reviews in the future. The reason can be summarized as follow:

- The Reviewer app offers the functionality that markers can be added to the screenshot, and it is helpful for the reviewer to describe the entities of the target application.
- The Reviewer app offers a way that the reviewer can give a review immediately when needed. Users can give reviews quickly when using the target application.
- The entrance of Reviewer app from taking a screenshot is rapid and convenient.

Q2 Which type of review do you think can be best supported by the Reviewer app? Why?

Both participants think the feature request can be easily explained by the Reviewer app. They think that feature requests can be expressed more specific with the screenshots and markers. Instead of describing the feature related entities, they can directly add markers to the screenshots of requested pages. Furthermore, they can point out location on the page where new features could be added by adding markers. However, there are bugs such as crash and server error which cannot be shown by screenshots. In addition, the user experience is usually a type of general reviews. Unless including comments to specific features or entities, there is no need to enclose screenshots with reviews. Therefore, the Reviewer app is more supportive to the reviews of feature request than bug reporting and user experience.

Q3 Which design of the Reviewer app do you think is the best? Why?

One participant thinks that marker adding is the best design of the Reviewer app, and he insists that users are able to offer more concrete and precise information in their reviews. On the other hand, the other participant thinks that taking a screenshot as the entrance to the Reviewer app is the best design, and he insists that only an efficient way with simple steps can make more users to offer reviews.

Q4 Do you have any suggestion to the Reviewer app?

Several useful and constructive suggestions collected from participants are listed below:

- The system log about the target application could be collected and saved with the review together.
- The protection to the privacy of users should be enhanced, some privacy information such as the contact information on the screenshot should be removed.
6.2 Results Analysis

6.2.1 Analysis of Results

The average scale of the questionnaire by different usability attributes is shown in Table 6.

<table>
<thead>
<tr>
<th>Usability attribute</th>
<th>Google Play review channel</th>
<th>Built-in review channel</th>
<th>“Reviewer app”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validity</td>
<td>2.17</td>
<td>2.67</td>
<td>4.33</td>
</tr>
<tr>
<td>Availability</td>
<td>2.17</td>
<td>2.5</td>
<td>4.33</td>
</tr>
<tr>
<td>Efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience</td>
<td>2.5</td>
<td>2.5</td>
<td>5</td>
</tr>
<tr>
<td>Complexity</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>Scenario one</td>
<td>1</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>Scenario two</td>
<td>4</td>
<td>1.5</td>
<td>3.5</td>
</tr>
<tr>
<td>Learnability</td>
<td>4</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>Errors covering</td>
<td>4.25</td>
<td>2</td>
<td>1.25</td>
</tr>
</tbody>
</table>

Table 6 Average scale for usability attributes

As shown in Table 6, compared with the Google Play review channel and the built-in review channel of YouTube, the effectiveness and efficiency of the Reviewer app receives a higher point. In addition, from the answers of the interview questions, it can be easily discovered that the Reviewer app has an obvious advantage on validity, availability, convenience and complexity. Furthermore, the result of scenario one shows the Reviewer app can help users initiate the review action swiftly. The entrance of taking a screenshot can lead users to the reviewing page directly and make users give reviews without interrupting the use of target application. However, the evaluation result of scenario two shows that the Reviewer app receives a lower point than the Google Play review channel. The reason is the participants sometimes forgot to take a screenshot when using the target application. When they opened the Reviewer app, they had to go back to the target application and retake the screenshot, which made them feel inconvenient and give a lower point.

Additionally, the evaluation result of the Reviewer app about learnability is as good as the other two methods. As the evaluation is only for the single function of review collection, it is easy for users to be familiar with it. Although the entrance of the Reviewer app has been changed and a new function of adding markers has been added, reviewers still believe they can understand how to use it in a short time.

As presented in Table 6, the Reviewer app performs poorly at errors covering. First, compared with the other two review collection methods, there is no alert window notified when moving back from the markers adding page to the main page. In this case, when adding a marker, if a user clicks the back button by mistake, he or she cannot continue to give the review when return to the markers adding page. In contrast, when giving reviews through Google Play review channel and built-in review channel of YouTube, if users...
click the back button or try to close the reviewing page, there are alert windows warning about the current operations. In addition, once the reviews have submitted from the Reviewer app and built-in review channel of YouTube, users cannot check, edit or delete them. If the user intends to modify the previous review in the Reviewer app, he or she should enter the reviewing page, upload the screenshot and give a new review again. However, when using Google Play review channel, users are able to check the reviews given by either themselves or other users. Further, users can edit their reviews after submission, and even the history of review editing can also be checked.

6.2.2 Potential Improvements

Reviewer app should be improved regarding error covering. First, the alert windows should be added when the user intends to quit. A confirmation window should also be added before the review submission. Next, the Reviewer app should add the function by which users are able to check and edit their previous reviews. Thus, if the user submits the review by mistake, he or she can edit the review and submit the correct one without uploading the screenshot and writing the review again.

In addition, the suggestions given by the participants can also be considered. The system log of the target application could be collected together with the review automatically, which can help developers better solve the bugs. For example, when a user intends to report the bug of application no respond (ANR), it is difficult to describe it by text or even the screenshot. The end users might not know how to describe the bug explicitly. In this case, the information of system log can help the application developers better analyze and solve the problem. Moreover, to avoid exposing user privacy, a modification tool is needed that can add mosaic to the screenshot in the markers adding page.

6.3 Limitations

First, as mentioned in the previous section, two aspects should be considered regarding effectiveness: 1) users can express their thinking precisely and perfectly; 2) the useful information can be extracted easily from collected reviews. However, only the first half of effectiveness has been evaluated. If the staff from real application producers join the evaluation, and collected reviews are analyzed from the perspective of them, the second half of effectiveness can also be assessed. Moreover, the number of participants was limited for obtaining a more convincing result of evaluation. Third, only two scenarios of use are considered in the evaluation. Considering more various scenarios and user interface also in the design and evaluation process can certainly improve the research.
7. Conclusion

The objective of the research is to provide an innovative way of collecting mobile app user reviews and to improve the usability of review collection on mobile apps. The existing review collection methods are not convenient to use for users. Meanwhile, users express their feedback in text on the heterogeneous level of quality, which brings difficulty for application producers to extract needed information for software maintenance and evolution. The Reviewer app, a new design of review collection, has been proposed in this research by applying user-centered design method with the effectiveness and efficiency taken into account. This design should ensure the validity and the availability of collected reviews, which makes the process of user review collection more efficient. The usability of the Reviewer app and two existing review collection methods are evaluated by the techniques of tasks, questionnaire and interview. Based on the result of the evaluation, the effectiveness and efficiency of the Reviewer app are much better than the others. The Reviewer app can make users express their thinking precisely about any of bug reporting, feature requests and user experience. It also makes the reviewing process conveniently. This design can be embedded into different applications and marketplaces as a review channel.
References


Appendix 1: Consent Form

Consent form

Please read and sign this form.

You have been invited to participate in a usability evaluation which is part of my master thesis work at Tampere University. By participating in the evaluation, you will help to evaluate and improve the usability of mobile application user review collection.

The evaluation process includes three parts:

1) In the tasks, you will be asked to perform a set of tasks in different scenarios;
2) In the questionnaire, you will give scales to a series of statements;
3) In the interview, you will answer the questions given by the tester.

Participation in this interview is voluntary. All information will remain strictly confidential. The findings might be used to analyze the proposed design in this research. However, at no time will your name or any other identification be used.

You can withdraw your consent to the interview and stop participation at any time. If you have any questions, please contact Ruijia Zha at zha.ruijia.x@student.uta.fi.

I have read and understood the information on this form.

Data and Place: ____________________ Signature: ________________
### Appendix 2: Questionnaire

#### Effectiveness

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1) I fully expressed my thinking in the review of bug reporting.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S2) I fully expressed my thinking in the review of feature request.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S3) I fully expressed my thinking in the review of user experience.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S4) The review of bug reporting I gave is concrete and precise.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S5) The review of feature request I gave is concrete and precise.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S6) The review of user experience I gave is concrete and precise.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Efficiency

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>S7) I can reach the reviewing page quickly and easily.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S8) I think this reviewing method has simple steps and is easy to use.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S9) I can give reviews without interrupting the current use of target application.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S10) I can give reviews quickly after using the target application.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Learnability

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>S11) I can easily understand how to give review by this method during the brief introduction.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Errors covering

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>S13) If I click other buttons by mistake, I can easily return to the reviewing page and continue the reviewing process.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>S14) If I submit the review by mistake, I can edit it and submit the correct one.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>