Design of the Usability Measurement Framework for Mobile Applications

Tapanee Treeratanapon

Abstract— Nowadays, mobile phones have become an essential tool for our everyday lives. People around the world use mobile phones extensively in different locations and situations. The increasing number of mobile users implies the importance of mobile application. However, there are obvious limitations for using mobile applications due to the characteristics of mobile phones including small screens, low display resolutions, non-traditional input methods, and difficult-to-use interface. Therefore, measuring mobile application usability is a more important issue for mobile phone, especially many mobile applications remain difficult to use. This paper proposes a methodology and framework to measure mobile application that are developed by free developers in various mobile software distribution platforms. We adopted the ISO 9241 standard and TAM model to establish the framework. The context of uses (user, task, equipment, and environment) and three main usability measures (effectiveness, efficiency, and satisfaction) from ISO 9241 standard that apply for general usability measurement are applied. We believe that ISO 9241 standard and TAM model are actually consistent and reflect the making a strong case for its use in related future study.

Keywords—Mobile Application; Usability Measurement; TAM Model; Context Of Use

I. INTRODUCTION

A global increase in mobile phone ownership and a rise in the use of mobile applications have caused necessary to provide a method for more user-friendly mobile applications on mobile phones. However, recently most mobile applications are still hard to use. One straightforward method and importance consideration is to measure mobile application usability before launch the mobile application. Usability measurement can help developers know the need for customer support decreases and importantly user satisfaction increases.

The most important goal of this paper is to apply our established framework to improve the usability measurement of mobile application. The characteristics of mobile applications and the unique features of mobile phones become the main challenges in usability measurement. The unique features of mobile devices that could be explored include small screens, low display resolutions, non-traditional input methods, limited bandwidth, unreliability of wireless networks, changing mobile context (e.g., location) and limited memory.

Moreover, the factors in the context of use include four main factors: user, task, equipment and environment that are significantly considered. We think that these factors highly and mainly impact to the usability of mobile application.

By adopting the framework, a developer can determine the degree of usability of mobile application according to how well designed the user interface is. The results of the analysis can be adopted as a guideline for mobile application design user interface fine-tuning in the future versions.

We found that many usability models and methods have been created in numerous areas. Among other models and methods that are previously developed such as the Metrics for Usability Standards in Computing (MUSiC) developed by MacLeod et al [1], the semi-Automated Interface Designer and Evaluator (AIDE) developed by Sears [2], and the Diagnostic Recorder for Usability Measurement (DRUM) and the Skill Acquisition Network (SANe). Two last models were developed by MacLeod & Rengger [3]. All the models and methods which aim to evaluate usability are useful but they still have some limitations. All of them are not focusing on mobile application.

Moreover, there are many generic usability questionnaires that have been developed for evaluating software usability, for example, the System Usability Scale (SUS) developed by Brooke [4], and the Software Usability Measurement Inventory (SUMI) developed by Kirakowski & Corbett [5], and the Post-Study System Usability Questionnaire (PSSUQ) [6]. Most of questionnaires are helpful but they have been deployed to evaluate the usability of specific software systems as well.

Usability helps mobile application easy to work but it cannot be directly measured [7]. It has been studied by measuring various different usability parameters and metrics. Bevan suggested that if we are going to select only one standard, this should be ISO 9241-210 as these provide the high level framework for usability work and provide tools for improving the usability capabilities of an organization [8]. Hornbaek noted that the ISO definition is commonly accepted [9]. The components of usability are both subjective and objective.

In this paper, we would like to publish works in relation to the mobile application and usability measurement that both objective and subjective measurement holistically comes up together. We established usability measurement framework based on ISO standard.
II. LITERATURE REVIEW

A. Mobile Applications

Mobile applications, also called mobile apps, are software applications, usually designed to run on smart phones and tablet computers [10]. The term "app" has become popular, and in 2010 was listed as "Word of the Year" by the American Dialect Society. Mobile applications are available through application distribution platforms, which are typically operated by the owner of the mobile operating system, such as the Apple App Store, Android Market, and BlackBerry App World. Some applications are free, and others have a price. Usually they are downloaded from the platform to a target device such as an iPhone, BlackBerry, or Android phone or downloaded by customers from various mobile software distribution platforms, or web applications.

Mobile applications were originally intended for productivity such as email, calendar, news alert, weather forecasting and contact databases, but public demand caused rapid expansion into other areas such as mobile games, factory automation, GPS and location-based services, banking, order-tracking, and ticket purchases [10]. Currently there are large numbers of new mobile applications happening but many mobile applications remain difficult to use.

Usability is one of the most important issues for mobile application. It can help mobile application easy to work because many mobile applications remain difficult to use, lack flexibility, and lack robustness.

B. Usability

Usability has been defined in various ways, but the concept of usability defined in ISO/IEC 9126 (1998) is widely accepted [9]. The International Organization for Standardization (ISO) 9241-11 [11] draft standard defines the term usability as "the extent to which a product can be used with effectiveness, efficiency and satisfaction in a specified context of use".

-Effectiveness is the accuracy and completeness with which specified users can achieve specified goals in a particular environment.

-Efficiency is the amount of resources expended in relation to the accuracy and completeness of the goal achieved.

-Satisfaction is the comfort and acceptability of the work system to its users and other people affected by its use.

The ISO9241 Usability Definition Framework is shown below in Figure 1.

A. Technology Acceptance Model

TAM is a theoretical model used in different contexts to help understand and explain the use of information technologies [12,13,14]. Technology Acceptance Model or TAM is an information system theory that models how users come to accept and use a technology. The model represents one of the explanatory models having most influenced the theories of human behavior [15]. TAM or Technology Acceptance Model is shown below in Figure 2.

III. RESEARCH DIRECTION

In this paper, the most important goal we will apply this framework to improve the usability of mobile application. We would like to use this analysis model to identify which factors, in the context of use, are significant and highly impact to the usability of mobile application. The results of the analysis can be adopted as a guideline for mobile application design user interface fine-tuning in the future versions.
From the ISO9241 usability definition framework, the main component that we considered firstly when we would like to measure usability is context of use.

A. Context of Use

One particular system placed in one context will probably display different usability measurement when placed in another context. Therefore context of use is important to consider when designing or measuring the system.

Based on the ISO9241 usability definition framework, the system’s context of use include users, tasks, equipment and environment.

- **Users** We think that there are many factors of users affect the mobile application. Those factors include user’s experience (expert, experience and novice users), demographics (age, gender, race, culture, knowledge, job specific roles, income, and location). After we reviewed many literatures, we found that the popular factors involving the users measured in usability are user’s experiences with mobile and user’s experiences with mobile applications as in [16,17,18] The culture, age group, gender and the level of interest in application may also be factors in determining mobile application but there are few literature measured in these factors as in [19].

- **Tasks:** We need to define a set of tasks for users to perform in order to measure mobile application. Task can create into two types: open (user defines outcome) vs. closed (tester pre-defines outcome). Task analysis may be manual tasks and analyzed as time and motion studies. Examples of tasks include working all steps as the way that mobile application should be, or using many functions of mobile application and recording the problems and encounters. In the set of tasks we created task according to TAM model that focuses on the concept of perceived usefulness. Therefore, the set of tasks should be actual, usable tasks, appropriated tasks and good representative of tasks from different difficulty levels ranged from hard to easy, and designed for different groups of users.

- **Equipment** Mobile variations include form factors such as input mode (e.g. pen, stylus, keypad, and button), processing power, battery life, screen size, resolution, and color depth which can have a significant effect on usability [20]. An example of using on-screen keyboards are today a standard of smart phone technology, it should be measured usability to understand the optimal design of on-screen keyboards according to target user groups and their characteristics as well. Furthermore mobile applications can run on smart phones and tablet computers, so we must set the category of devices that we want to measure usability first. Moreover, some mobile applications are available through application distribution platforms differently, which are typically operated by the owner of the mobile operating system. Then we must measure usability in different operating system that mobile application can run.

- **Environment:** During the evaluation process, the researcher needs to focus on all activities that the users perform. All responses and actions are important and should be recorded for the final outcome analysis. There are two situations to measuring mobile application usability: lab based method and field based method as in [21]. Both of them have difference advantages and disadvantages. If we used field based method, we will get the results in the actual learning environment. In additions, if we set the laboratory, we should consider light, changing mobile context (e.g., location), and user motion such as sitting or walking as well because these factors are the important situation should be considered. However, sometime we must add additional equipments in environment, such as video recorder or camera. Statistics gathering module should be developed and embedded into the platform to automate the user behavior tracking process, such as the total time used to complete the specified task.

After we finished setting the context of use of mobile application including users, tasks, equipment and environment, then we will measure the usability.

B. Usability Measure

We adopt the ISO 9241 standard to establish the framework. Thus the three main usability measures (effectiveness, efficiency, and satisfaction) from ISO 9241 standard were applied. The effectiveness and efficiency is measured by task experiment and observation. The satisfaction is measured by interviewing or using questionnaire.

We recommend the metric that use for efficiency is the total time to complete the task or how many times to get the right path of a task. The metric for the effectiveness is quality, and correctness of the solution. The satisfaction is measured by answering the questionnaires and by observing the users performing and other common behaviors such as eye movement, gesture, and face.

In the satisfaction measure, we adopted TAM evaluation model to develop. The satisfaction questions have been adapted from Davis’ original ones [22]. TAM model consists of three main constructions: 1. Perceived ease of use (PEOU); 2. Perceived usefulness (PU); 3. Intention to use (IU). The satisfaction questions in a PEOU, PU and IU construction show below respectively.

Questions in perceived ease of use (PEOU) show below.

Q1. I would find this application user-friendly.
Q2. Learning to operate and use this application would be easy for me.
Q3. I would easily find the information I am looking for using this application.
Q4. I would find the user interface of this application clear and intuitive.
Q5. I would find this application to be flexible to interact with.
Q6. To me it seems this application easily would do what I wanted it to.
The questions in a perceived usefulness (PU) show below.
Q1: Using this application would be useful for me in my everyday life.
Q2: Using this application would make my everyday life better.
Q3: Using this application would increase my efficiency in everyday life.

The questions in an intention to use (IU) show below.
Q1: When I have accessed to this application, I would like to intend to use it
Q2: Given that I have accessed to this application, I predict that I would use it.

Finally when we measured the three main usability measures (effectiveness, efficiency, and satisfaction), we got the data. Then the statistic model can be used to analyze the data.

C. Outcomes
There are many statistics which can be applied such as mean and SD of scores from each item, ANOVA, regression, t-test, frequency analysis (chi square) and correlation. The statistical models that we used depend on the experimental design. These outcomes should reflect the usability of the system directly.

IV. ADAPTATION USABILITY MEASUREMENT FRAMEWORK
In our study, for example, we would like to measure mobile application usability on two kinds of operating system (Android and IOS) on three groups of users (novice, experience, and expert). We set the experiment as shown below.

There are two variables interested that affect the usability: types of operating system and groups of uses. Thus before running in the experiment, participants are assessed their familiarity in mobile application experience through questionnaires or interviewing for studying and classifying groups of users that affect the usability.

A. Experimental Design
1) Participants: The experiences of participant were divided into three groups: novice, experience, and expert.
2) Dependent variables: We will measure usability: efficiency through the time, it takes to finish a task, effective through the correctness or error of the solution and satisfaction: it is assessed by asking participants in their satisfaction after they completed the tasks.
3) Independent variables: Groups of users, and types of operation system are two main independents variables. Users include 3 groups: novice, experience, and expert. Types of web are consisted of 2 operating system: Android and IOS.

4) Goal: An experiment was designed to measure mobile application usability on two kinds of operating system for three groups of users (novice, experience and expert).
5) Procedure: We installed mobile application on sites for experimenting that having same product but considering differential two kinds of operating system and three groups of users for experiment. Two factors factorial design (2 levels of operating system * 3 groups of user) was used. The participants were equally divided equal number of each group.

B. Method
Usability is the extent to which a website can be used with effectiveness, efficiency and satisfaction in a specified context of use.

There are many methods for measuring usability such as testing, inquiry, inspection and heuristic analysis. Each method has different advantages and disadvantages. Each method to employ depends on our limitations and the environment. For example the inspection method allows faster and more cost-effective problem identification but some problems can be identified less precisely.

In testing, users are requested to discover the answers to their questions by using the system. Testing is clearly an empirical method. It usually involves the evaluation of a prototype or the actual system through the observation of users working on relevant tasks. Test subjects are the population who will use the product and are asked to perform real tasks using the prototype and the task setting.

This measurement to support testing evaluation includes operation time and number of errors made. Operation time includes time taken to complete tasks, and the time interval of operation. We recommend measuring usability in a task context. The advantage of these measurements is that it is convenient to collect the data and is economic.

C. Data Analysis
There are two factors of data analysis in this study. The first dimension is two types of operating system. The other is groups of users (novice, experience expert). Outcome dependent variables are analyzed by using general statistical analysis method (e.g., ANOVA (two factors factorial design)).

Other factors in the system context of use can also be evaluated in the same way. With this analysis model, we can identify the degree of impact that the test factor has on the usability of the overall system. The analysis results can then be used to fine-tune the system user interface for specific user groups.

V. CONCLUSIONS
Our preliminary work has been to gather related information and design a usability measurement framework that is specific for use with the mobile application. The independent variables are the many factors or variables in context of uses. The dependent variables are usability metrics (completing time, correctness, and satisfaction). Statistical
model for mobile application usability will be derived for use as our analysis tools. The outcome of the experiment is to investigate which variables in the context of use affected the mobile usability metrics and how application should be designed in order to increase usability.

REFERENCES