Usability Metric Framework for Mobile Phone Application

Conference Paper · June 2009

CITATIONS
30

READS
4,285

2 authors:

Azham Hussain
Universiti Utara Malaysia
99 PUBLICATIONS **285** CITATIONS
SEE PROFILE

Maria Kutar
University of Salford
30 PUBLICATIONS **283** CITATIONS
SEE PROFILE

Usability Metric Framework for Mobile Phone Application

Azham Hussain
Informatics Research Institute
University of Salford
Greater Manchester
M5 4WT United Kingdom

Maria Kutar
Informatics Research Institute
University of Salford
Greater Manchester
M5 4WT United Kingdom

Abstract - **Measuring usability is an essential task to ensure the application is accurate, has sufficient in speed and to ensure the safety of the user from strain injury as well. The increasing number of mobile phone users means that many businesses have deployed mobile applications to gain competitive advantage. Literature on how to measure usability is limited in the area of mobile phone applications. Recent technology like GPS receiver embedded into mobile phone creates new challenges to researchers. This paper attempts to review existing metrics for desktop computing and subsequently develop a conceptual model to evaluate mobile phone application. The model will consist of usability metric to evaluate mobile application.**

I. INTRODUCTION

According to a report from International Telecommunications Union (ITU), mobile phone users worldwide soared to over 3.3 billion by the end of 2007. Mobile applications such as news alert, weather forecasting and entertainment become more popular and well accepted. The fast growth and high demand on mobile application have attracted researchers to extend the studies on any potential area in mobile applications.

Usability is commonly comprehended as a qualitative attribute that assesses quality-in-use or how easy applications are to use [1] and [2]. The word "usability" also refers to methods for improving ease-of-use during the early design process [3] and [4]. Focusing on usability and user experience is a key element in creating successful high-quality applications. Unfortunately, there are few clear guidelines on how various definitions of usability factor, rules and criteria are related and how to measure usability of mobile applications. Instead, some of the developers tend to employ usability methods that they are familiar with, whereas some of the methods may not be appropriate to apply to all applications.

The novelty of mobile applications and the unique features of mobile devices become the main challenges in usability measurement activity for mobile devices. A number of unique features of mobile devices that could be explored include limited bandwidth, unreliability of wireless networks, changing mobile context (e.g., location) and limited memory. Recent technology like GPS receiver embedded into mobile phone creates new challenges in HCI area. Many traditional usability metrics were purposely created for desktop applications, however the metrics may not be directly applicable to mobile applications [5]. Thus, there is a need for new systematic usability measurement for mobile applications as mentioned by Ahmed [6] and it is well supported by Brodkin [7] in his prediction report that revealed mobile phone will be the primary internet device by 2020.

This paper aims to review previous studies and current measurement models for usability evaluation through systematic literature review (SLR). The analysis of current model and previous study will result in a set of selected usability guidelines for mobile applications. The guidelines will be expended into measurement model consisting of metric for evaluation.

In the next section, a review of several usability measurement models will be presented and highlights the limitations and advantages of the various models. In section 3, we describe the approach undertaken to achieve the aims of this study and followed by the discussion of the result. Finally, the conclusion will take place.

II. RELATED STUDY

A number of models for usability measurement are available for reference; for instance, Quality in Use Integrated Measurement (QUIM) developed by Ahmed et al. [6]. QUIM is a consolidated model for usability measurement and metric; and also appropriate for users who have no or little knowledge of usability. The model consists of 10 factors which are subdivided into 26 criteria. For the measurement of the criteria, the model provides 127 metrics. The model is used to measure the actual use of working software and identifying the problem. However, the model is not optimal yet and needs to be validated.

On the other hand, Metrics for Usability Standards in Computing (MUSiC) develop by Bevan & MacLeod [8] is another project concerned on defining measures of software usability and was integrated into the original ISO 9241 standard. Examples of specific usability metrics in the MUSiC framework include user performance measures, such as task effectiveness, temporal efficiency, and length or proportion of productive period. However, a strictly performance-based view of usability cannot reflect other aspects of usability, such as user satisfaction or learnability. Software Usability Measurement Inventory (SUMI) develop by Kirakowski & Corbett [9] is a part of MUSiC project. SUMI was developed to provide measures of global satisfaction of five more specific usability areas, including

ISBN: 978-1-902560-22-9

effectiveness, efficiency, helpfulness, control, and learnability. Another MUSiC project related to software tool development is Diagnostic Recorder for Usability Measurement [10] developed by Macleod & Rengger [11]. This project concerns with the analysis of user-based evaluations and delivery of these data to the appropriate party, such as a usability engineer. The Log Processor component of DRUM is the tool concerned with metrics. It calculates several different performance-based usability metrics including 1) Task time 2) Snag, help, and search times 3) Effectiveness 4) Efficiency 5) Relative efficiency and 6) Productive period.

In addition, the Automated Interface Designer and Evaluator (AIDE) that was developed by Sears [12] concerns with evaluating static HTML pages according to a set of predetermined guidelines about Web page design. AIDE is a software tool that able to generate alternative interface layouts and evaluate some aspects of a design. Among things that are concerned in these guidelines include the placement and alignment of screen elements for example text, buttons, or links. There are two metrics to be evaluated in the design which are task-sensitive metric and task-independent metric. Task-sensitive metrics incorporate task information into the development process which may ensure that user tasks guide the semantics of interface design. Task-independent metrics tend to be based on principles of graphic design and help to ensure that the interface is aesthetically pleasing. AIDE tool can measure a total of five different usability metrics, including efficiency, alignment, horizontal balance, vertical balance, and designer-specified constraints.

Subsequently, other model that deal with the analysis of the quality of use for interactive devices was introduced which is The Skill Acquisition Network (SANe) by Macleod & Rengger [11]. This approach assumes a user interaction model that defines user tasks, the dynamics of the device, and procedures for executing user tasks. Specifically, a task model and a device model are simultaneously developed and subsequently linked. After that, user procedures are simulated within the linked task-device model. A total of 60 different metrics are described in this framework, of which 24 concerns with the quality measures. Scores from the latter are then combined to form a total of five composite quality measures including: Efficiency, Learning, Adaptiveness, Cognitive workload, Complexity and Effort for error correction.

A. International Organization for Standardization (ISO)

The International Organization for Standardization (ISO) is an international standard setting body composed of representatives from various national standards organizations. ISO has developed over 17000 International Standards on a variety of subjects and 1100 new ISO standards are published every year (source from ISO website). Most of literature in HCI employed ISO9241-11 for usability measurement [13]. Table 2.1 lists the ISO standard related to HCI.

ISO9241-11 specifically addresses the definition of usability measurement and thus it is chosen as foundation for framework in this study. In another study to analyze usability measurement dimensions by Constantinos & Dan [14], they found the highest characteristic in usability evaluation are effectiveness (62%), Efficiency (33%) and satisfaction (20%). These three characteristics reflect the ISO 9241 standard as they are the measure attributes for that standard.

TABLE 2.1
ISO Standard related to measurement

Usability in ISO Standard	Description
The ISO 9241-11 (1998)	Identify efficiency, effectiveness, and satisfaction as major attributes of usability.
ISO/IEC 9126-1 (2001)	Define the standard as a software quality attributes that can be decomposed into five different factors, including understandability, learnability, operability, attractiveness, and usability compliance.
ISO/IEC 9126-4 (2001)	Define the related concept of *quality in use* as a kind of higher-order software quality attribute.
The ISO/IEC 14598-1 (1999)	A model for measuring quality in use from the perspective of internal software quality attributes

III. RESEARCH APPROACH

This study consists of two phases in order to achieve the two objectives mentioned earlier. The first phase will explore the previous study on usability measurement and later come out with the guidelines to evaluate mobile application. In the second phase, the metric for usability evaluation will be develop by referring to the guidelines developed in first phase. Goal Question Metric (GQM) approach originated by Basili et al., [15] will be employed in this phase to develop usability metric for mobile phone application.

In GQM, the first step is identifying the goal. The guidelines created in the first phase will be the goal in this study, followed by the development of questions to assess each goal. The questions developed will constitute the basis for quantitative metrics definition. At the end, a set of metric will be produced to provide information to answer the questions developed previously.

B. Phase 1: Systematic literature Review

This phase describes the method for the review of usability measures employed in four studies from the HCI research literature. The total of 409 journals was reviewed based on keywords “usability”, “evaluation” and “metric”. Only 26 out of 409 journals selected for further review in obtaining the guidelines for mobile application development. Table 3.1 below describes the journal papers that were reviewed.

TABLE 3.1
Journal Paper Reviewed

Journal	Year	Candidate	Selected
TOCHI	2006-2008	54	8
HCI	2006-2008	36	2
IJHCI	2006-2008	97	5
IJHCS	2006-2008	222	11
Total		409	26

TOCHI : ACM Transactions on Computer-Human Interaction
HCI : Human-Computer Interaction
IJCHI : International Journal of Human-Computer Interaction
IJHCS : International Journal of Human-Computer Studies

The review is based on a conception of usability, similar to ISO 9241, part 11 [2] and Bevan & MacLeod [8]. This conception merely discusses studies related to usability evaluation instead of the broad concept of usability. We analyse the quality characteristic of each measures to ensure there are no duplication. We also refine the measures to simplify the guidelines and to ensure the model not too complex. Interestingly, we found most of studies employed *effectiveness, efficiency* and *satisfaction* as quality characteristics. Thus, we decide to make these three characteristics as a base of guidelines and others become sub guidelines. Table 3.2 below describes the most popular guidelines obtained from literature.

TABLE 3.2
Current practice on usability guidelines

No	Guidelines	Explanation
1	Completeness	The extent or completeness of users' solutions to tasks
2	Accurate	The accuracy with which users complete tasks
3	Less or no error	Errors made by the user during the process of completing a task
4	Ease to input the data	The data input process should be simple
5	Ease to use output	The output should be very simple and accurate.
6	Ease to install	Should be friendly while install the application.
7	Response time	The system must respond in an appropriate time
8	Simple	The application should be straightforward
9	Time	The duration of tasks or parts of tasks
10	Ease to learn	The user interface must be designed for user to learn easily
11	Application size	The space used by application should be appropriate.
12	Battery power used	The battery power use by the application
13	Wireless connectivity	The application should easily connect to network
14	Features available	Appropriate features available on application
15	Satisfy with interface	Measures satisfaction as the interface users prefer using
16	Provide support/help	The help information given by the application is useful
17	Safety	User should be saved and secured while using the application

C. *Phase 2: Metric Development*

This phase will first review the GQM model prior to developing the metric to measure usability of mobile phone application. GQM is a goal-driven method for developing and maintaining a meaningful metrics program that is based on three levels; Goals, Questions and Metrics. The approach brings success for the reason that it is adaptable to many different organizations and environments, as confirmed by a large number of companies that have employed it (e.g. Philips, Siemens, NASA) [16]. Even though GQM was originally used to define and evaluate goals for a particular project and environment, its use has been extended to larger perspectives such as quality improvement, progress measurement and project planning [15]. Hence, GQM approach could possibly be extended as well to measure the usability guidelines by providing metric for guidelines.

To apply the GQM approach and acquire the usability metric, the goal should be identified initially. All the guidelines obtained in first phase will be the goal in the GQM model. In second step, questions will be developed to assess each goal described in the first step. We carefully created the questions by refining the goal into several questions and ensure the questions we created are measurable. Finally, we have to develop a set of metric that provide the information to answer those questions. In this case, we will refine all the questions into metrics.

IV. RESULT AND DISCUSSION

Each guideline will be the goal for our GQM model similar to Gafni [17] in order to get usability metric. As seventeen guidelines are developed, it is possible to get a complex GQM model. Hence, we decided to simplify the guidelines by reviewing and combining the guidelines that will affect other guidelines when conducting usability test. We also remove the guidelines that are not relevant to HCI which are "application size", "battery power" and "wireless connectivity". Since this study focus on mobile application, we create our own guidelines that appropriate to mobile phone for instances; "touch screen facilities" and "safety while driving". We add a new guideline as well that is "automatic update" as the guideline did not appear in literature at all. As a result, 6 guidelines act as goals for GQM model as shown in table 4.1 below:

TABLE 4.1
Usability Guidelines

Quality Characteristic	**Goal**	**Guidelines**
Effectiveness	Simplicity	-Ease to input the data -Ease to use output -Ease to install -Ease to learn
	Accuracy	-Accurate -Should be no error -Successful
Efficiency	Time taken	-To response -To complete a task

	Features	-Support/help -Touch screen facilities -Voice guidance -System resources info. -Automatic update
Satisfaction	Safety	-While using the application -While driving
	Attractiveness	-User interface

From the goals above, we have created questions to assess each guideline as stated below:

Simplicity

1. Is it simple to key-in the data?
2. Does the application provide a virtual keypad?
3. Is the output easy to use?
4. How easy is it to install the application?
5. Is the application easy to learn?

Accuracy

1. Is the application accurate?
2. How many tasks are successful in the first attempt?
3. How many tasks are successful in a given time?

Time taken

1. How much time taken to complete a given task?
2. How much time taken by application to respond?
3. How much time taken by user to learn?

Features

1. Does the application provide appropriate help?
2. Does the application provide appropriate menu button for touch screen?
3. Does the application provide voice assistance?
4. How much information about system resources was displayed?
5. Does the application provide automatic update?

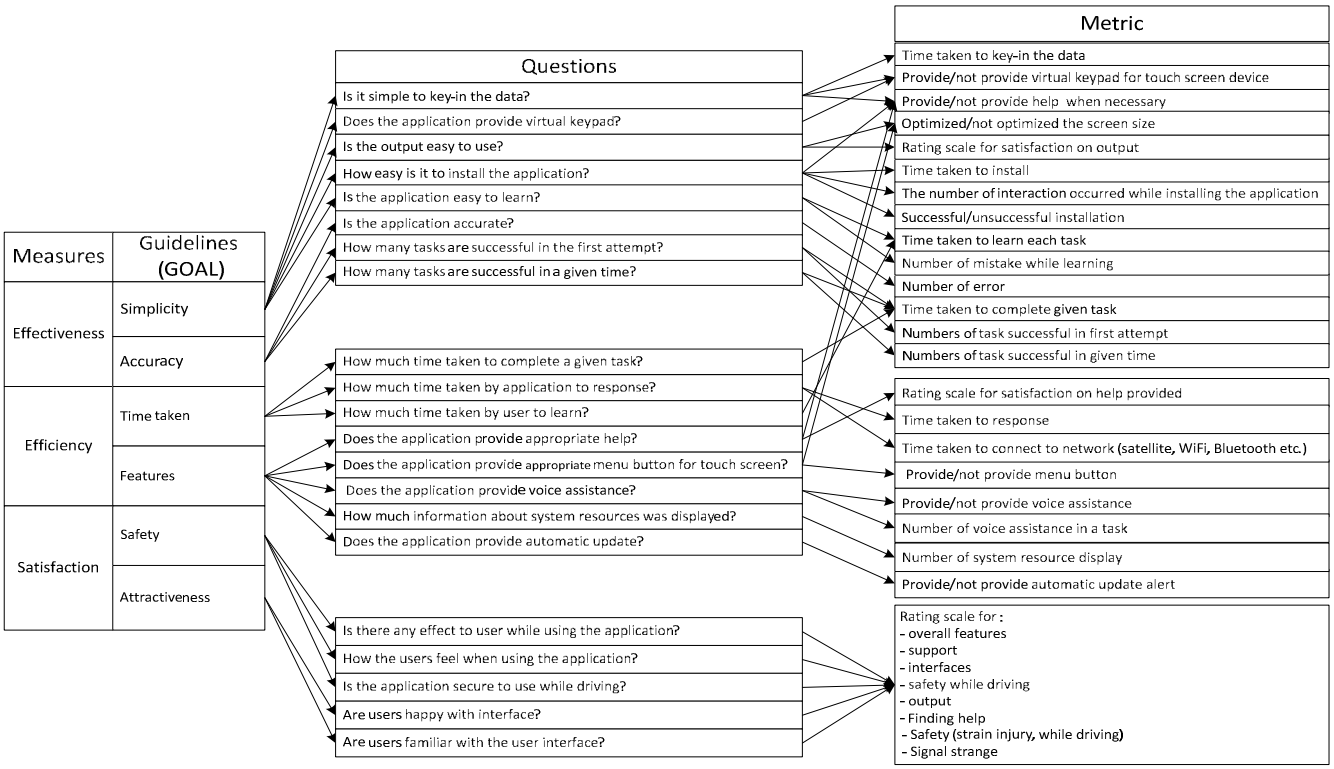

Figure 4.2: Complete GQM Model to Evaluate Mobile Usability

Safety

1. Is there any effect to user while using the application?
2. How the users feel when using the application?
3. Is the application secure to use while driving?

Attractive

1. Are users happy with interface?
2. Are users familiar with the user interface?

Next step is about examining how the questions should be answered, moving from the qualitative to a quantitative level. Once goals are refined into a list of questions, metrics need to

be defined that provide all the quantitative information to answer the questions in a satisfactory way. We carefully review all questions and found that not all questions can be answered objectively. Some measurements will be answered using questionnaire particularly to assess user satisfaction. As a result, figure 4.2 describes complete GQM model to evaluate mobile phone application. The model clearly describes the usability metric (derived after answering the question) and we also include quality characteristic from ISO 9241-11 in our model.

V. CONCLUSION

We have reviewed the current practice of measuring usability and come out with the guidelines to assist mobile application developer. We obtained the guidelines from previous research and simultaneously appended a few guidelines compatible with mobile application. The guidelines then become the goal of our GQM model for the reason that the guideline is "target to achieve" as well as a goal. Finally we have created a set of questions and metric to assess mobile application.

The GQM approach we employed in this study have shown how this approach can be used to develop usability metric. Existing literature suggests that choosing usability measures is difficult in current practice but our model provides an additional approach to choose. However, this model needs to be validated for future work to ensure all metric we created are applicable to mobile application as mobile computing is different from traditional computing. The next stage in this research will be to apply the model in the evaluation for a mobile application in order that we may validate the model.

VI. REFERENCES

[1] IEEE_Std_1061: 'IEEE standard for a software quality metrics methodology', IEEE Std 1061-1998, 1998

[2] ISO: 'INTERNATIONAL STANDARD: ISO 9241-11(Guidance on usability)', in Editor (Ed.)^(Eds.): 'Book INTERNATIONAL STANDARD: ISO 9241-11(Guidance on usability)' (1998, edn.), pp.

[3] Yonglei, T.: 'Work in progress - introducing usability concepts in early phases of software development'. Proc. Frontiers in Education, 2005. FIE '05. Proceedings 35th Annual Conference, Year pp. T4C-7-8

[4] Miller, J.: 'Usability Testing: A Journey, Not a Destination', Internet Computing, IEEE, 2006, 10, (6), pp. 80-83

[5] Zhang, D., and Adipat, B.: 'Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications', International Journal of Human-Computer Interaction, 2005, 18, (3), pp. 293 - 308

[6] Ahmed, S., Mohammad, D., Rex, B.K., and Harkirat, K.P.: 'Usability measurement and metrics: A consolidated model', Software Quality Control, 2006, 14, (2), pp. 159-178

[7] Brodkin, J., 'Mobile phones to be primary Internet device by 2020, experts predict', http://www.networkworld.com/news/2008/121508-pew-report.html, Access: 17 December 2008

[8] Bevan, N., and MacLeod, M.: 'Usability measurement in context', Behaviour & Information Technology, 1994, 13, (1), pp. 132 - 145

[9] Kirakowski, J., and Corbett, M.: 'SUMI: the Software Usability Measurement Inventory', British Journal of Educational Technology, 1993, 24, (3), pp. 210-212

[10] Drummond, J.S., and Themessl-Huber, M.: 'The cyclical process of action research: The contribution of Gilles Deleuze', Action Research, 2007, 5, (4), pp. 430-448

[11] Macleod, M., and Rengger, R., The development of DRUM: A software tool for video-assisted usability evaluation, http://www.nigelbevan.com/papers/drum93.pdf, Access: 29 January 2009

[12] Sears, A.: 'AIDE: a step toward metric-based interface development tools'. Proc. Proceedings of the 8th annual ACM symposium on User interface and software technology, Pittsburgh, Pennsylvania, United States1995 pp. Pages

[13] Hornbæk, K., and Law, E.L.-C.: 'Meta-analysis of correlations among usability measures'. Proc. Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA2007 pp. Pages

[14] Constantinos, K.C., and Dan, K.: 'A research agenda for mobile usability'. Proc. CHI '07 extended abstracts on Human factors in computing systems, San Jose, CA, USA2007 pp. Pages

[15] Basili, V., Caldeira, G., and Rombach, H.D.: 'The Goal Question Metric Approach', Encyclopedia of Software Engineering, 1994

[16] Solingen, R., and Berghout, E.: 'The Goal/Question/Metric Method: A Practical Guide for Quality Improvement of Software Development' (McGraw-Hill Publishing Company, 1999. 1999)

[17] Gafni, R.: 'Framework for Quality Metrics in Mobile-Wireless Information Systems', Interdisciplinary Journal of Information, Knowledge, and Management, 2008, 3, pp. 23 - 38