Modeling Subjective Metrics for Mobile Evaluation

Azham Hussain[1]
e-mail: azham.h@uum.edu.my

Maria Kutar[2]
e-mail: m.kutar@salford.ac.uk

Arifin Abdul Mutalib[3]
e-mail: am.ariffin@uum.edu.my

Fazillah Md Kamal[4]
fazillah@uum.edu.my

Author(s) Contact Details:
[1,3,4] *Universiti Utara Malaysia, UUM Sintok, Kedah, Malaysia*
[2] *University of Salford, M5 4WT, Manchester, United Kingdom*

Abstract—There are several methods to evaluate mobile applications. The aim is to ensure that the applications are usable and satisfying users. Among the methods include lab test and field test. However, not many have focused on the quality characteristics should be used during the test. This study aims to develop subjective metrics for evaluation of mobile application using Goal Question Metric (GQM). To ensure the metrics we develop are reliable and useful, we conduct usability study that used the metrics we developed using GQM approach. The result indicates that the metrics can be used to evaluate mobile application in term of user satisfaction. Four test cases with different platforms employed in this study shows that the metrics are reliable to be used in mobile evaluation.

Keywords – mobile application; usability study; evaluation; goal question metric

1. INTRODUCTION

The increasing number of mobile phone users means that many businesses have deployed mobile applications to gain competitive advantage. Applications in a mobile phone such as news alert, weather forecasting, and entertainment have become increasingly popular and well accepted. The fast growth and high demand for mobile application has attracted researchers to extend their studies on various potential areas in mobile applications. One of the most popular research issues in Human Computer Interaction (HCI) is on how to measure usability [1]. In fact, measuring usability is an essential task to ensure an application is accurate and to ensure that users are safe [2].

Focusing on usability and user experience is a key element in creating successful high-quality applications. Unfortunately, there are too few clear guidelines on how various definitions of usability factor, rules, and criteria are related and how to measure usability of mobile applications. Consequently, developers tend to employ usability methods that they are familiar with, in which some of the methods may not be appropriate to apply to all applications.

Literatures on how to measure usability is very limited [1] and much less in the area of mobile phone application itself [3]. The novelty of mobile applications and the unique features of mobile devices become the main challenges in usability measurement activity. Some of the 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 technologies such as GPS receivers embedded into mobile phone create new challenges in HCI area. Many traditional usability metrics have been purposely created for desktop applications in the past, however the metric may not be directly applicable to mobile applications [4]. Thus, there is a need for new systematic usability measurement for mobile applications as mentioned by Ahmed [2], which is well supported by Brodkin [5] in his prediction report that revealed mobile phone will be the primary internet device by 2020.

2. USABILITY IN MOBILE APPLICATION

Mobile applications can be defined as software systems operating on mobile devices [4]. A wide range of mobile applications, for instance financial applications, marketing and advertising applications, education applications, and emergency applications creates many research opportunities to research communities. Based on the existing literatures, usability evaluation, usability evaluation methods, techniques for conducting usability test, usability standards, and guidelines are among popular research areas in mobile applications. Particularly, usability evaluation consists of methodologies for measuring the usability aspects of a system user interface and identifying specific problems [6]. Meanwhile, measuring usability is a part of usability evaluation process, in which effectiveness, efficiency, and satisfaction are three examples of the usability measures [7].

Usability testing is a part of the usability engineering process in which performance is measured to determine whether usability goals are achieved [8]. The two methods that are frequently used for usability testing are laboratory test and field

test. Both approaches have advantages and drawbacks, mainly in testing control, but the main objective of usability testing is to get accurate results while implementing the usability test. Further, selection of an appropriate method for a usability study depends on its objectives and usability attributes [4].

Usability standards and guidelines present rules and principles to guide and help developers in designing and developing the interface for their applications [9]. Standards are generally sold by the organizations that publish them, or by a national standards organization such as International Standards Organization (ISO). Literatures in HCI show many usability guidelines for desktop application have been developed but very little are suitable for use in mobile devices. Additionally, most of the guidelines were created by conducting usability testing and evaluating feedback from users; (see Alfredo, et al., [10], [11], [12] and [13]). Other researchers have produced guidelines based on the literatures on usability theoretical frameworks, constraints, and unique properties that are intrinsic to mobile computing (for example [14], [4] and [15]). However, the guidelines do not focus on interaction and are not validated.

The International Organization for Standardization (ISO) is an international-standard-setting body composed of representatives from various national standards organizations. Among all, most literatures in HCI employ ISO9241-11 for usability measurement [1]. Particularly, ISO9241-11 specifically addresses the definition of usability measurement and thus it is chosen as the foundation for the model in this study. Besides, Constantinos and Dan [3] found that the highest characteristics in usability evaluation are effectiveness (62%), Efficiency (33%), and satisfaction (20%). These three characteristics reflect the ISO 9241 standard as the characteristics are the measurement attributes for that standard. However, the guidelines are very general and need to be refined to put into practice [16] and [2].

Several methods are available to measure usability, such as usability testing, heuristic evaluation, focus group, and questionnaire. All the methods have advantages and disadvantages depending on the objective of the study. Among all, usability testing is the most recommended method in measuring usability [6]. Heuristic evaluation focuses on the interface and some criticisms of focus groups include bias by presence of dominant facilitator [17]. Meanwhile, QUIM, QUIS, SUMI, and AIDE are several examples of models to measure usability. Unfortunately, all the models do not focus on mobile applications and some of them are very difficult to apply [2]. In recent study by Gafni [33] on evaluation of mobile wireless system, a list of metrics for mobile evaluation were derived, however the study did not include the evaluation of interaction between human and mobile applications. Hence, this study will develop usability metrics to evaluate the interaction between human and mobile applications.

3. GOAL QUESTION METRIC (GQM)

Goal Question Metric (GQM), originated by Basili et al., [18] is a popular approach to measure quality of software project. It can be applied to various areas of measurement including effort, schedule, and process conformance [19]. It is a top-down approach to establish a goal-driven measurement system for software development. Some works employing GQM approach in developing measurement metric could be observed in [20], [21], and [22]. Besides, GQM has also been claimed as one of the most goal-focused and most widely used among the various measurement approaches [21].

Although GQM was originally used to define and evaluate goals for a particular project and environment, its purpose has been extended to a larger perspective including to improve quality, measure progress, and plan for project [18]. Hence, it is deduced that the GQM approach could also possibly be extended to measure the usability guidelines by providing metrics for usability guidelines [23]. In conjunction, Figure 1 describes the definition phase of the GQM paradigm, illustrating the outputs of the first three steps of Basili's GQM process, the hierarchy of goals, questions, and meaningful metrics.

Based on the discussions in the previous paragraphs, this study employs GQM, for the reason that the approach is the most widely adopted and well respected for developing measurement metric. In addition, the GQM plan could also be reused repetitively, which will reduce the cost [24].
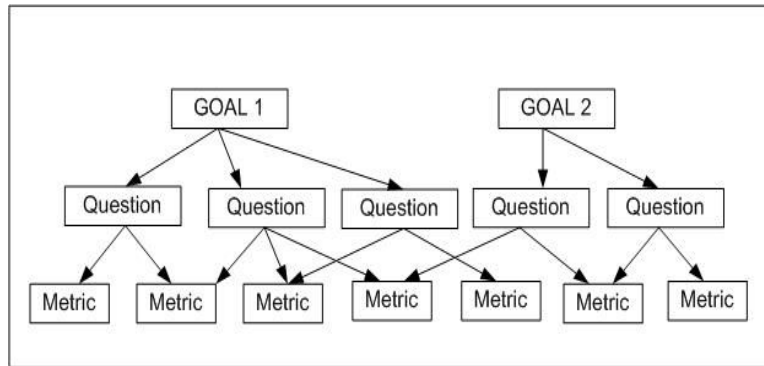
FIGURE 1: A GQM Model (adopted from Basili et al. [18])

4. RESEARCH APPROACH

Information system research is generally regarded as social research, which includes positivist, interpretive, and critical research [25] and [26]. Each of the philosophy has its strengths and weaknesses. In regards to that, the philosophy adopted in this study is based on how best it can serve to achieve research objectives and how to answer the research question. Thus, the appropriate philosophy adopted is broadly interpretive.

There are various research strategies available for selection. In conjunction, selecting appropriate methods for a particular research area requires a good understanding of each methodology as there is no prescribed approach to it [27]. In addition, the choice of research methodology also depends on a numbers of factors, such as budget, duration, and the nature of the research topics [28]. Those considerations are within the views of this study, which consists of three phases as illustrated in Figure 2. The figure explains that the first phase explores the previous models and frameworks on usability measurement. It is essential to come out with a theory in usability evaluation in HCI. A systematic literature review (SLR) method is used in this phase to ensure that previous studies on usability evaluation relevant to this study are considered. In the second phase, the metric for usability evaluation is developed by referring to the guidelines developed in phase 1. The GQM approach is employed in this phase to develop usability metric for mobile applications. Finally, the GQM model developed in the second phase is validated and refined through usability tests.
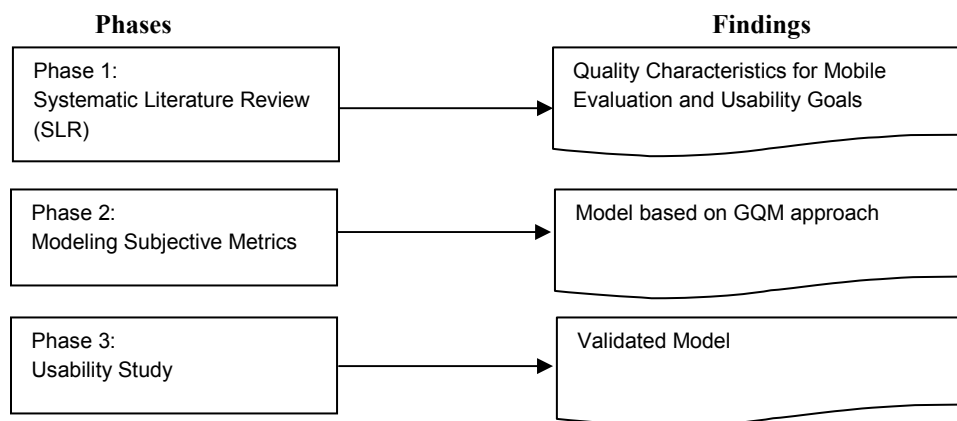

FIGURE 2: Research Design

A. *Systematic literature review (SLR)*

SLR is a mean of identifying, evaluating, and interpreting all available studies relevant to the research questions being attempted, or topic area, or phenomenon of interest [29]. This study follows the three main steps by Kitchenham in order to investigate the quality characteristics of mobile applications. The result from this phase is a summary of existing evidence in mobile evaluation; specifically the result is a usability goal for mobile application. An example of works employing SLR in their studies include Hornbæk [30], who studied the challenge on measuring usability. Besides, Kjeldskov et al. [31] developed a new technique for evaluating mobile devices.

A total of 1136 journal and proceeding papers were reviewed by "usability", "evaluation", and "metric" keywords. The method are similar to Bevan and MacLeod [32] and Hornbæk [30]. Having digested the contents, all papers were tabulated as outlined in Table 1. This conception merely outlines studies related to usability evaluation instead of the broad concept of usability. Accordingly, 56 journals were selected for further review to obtain the usability goals for mobile application.

TABLE 1: Journal Paper Reviewed

Year / Journal & Proceeding	2006	2007	2008	2009	2010	TOTAL
TOCHI	18	14	22	20	18	92
HCI	13	13	10	10	11	57
IJHCI	31	27	39	36	44	177
IJHCS	84	70	68	70	60	352
IJMHCI	-	-	-	22	15	37
MobileHCI	69	63	107	102	80	421
Total	215	187	246	260	228	1136

Analysis has been made to the quality characteristics of each measurement to ensure no duplication, focusing only on the interaction. Besides, only subjective goals are selected. Some examples of quality measures which have been removed are "easy to learn" and "easy to use". The term "easy to learn" is similar to "learnability" and "easy to use" is similar to "simplicity". Also, quality characteristics that do not focus on interaction have been removed such as "application size", "memory load" and "cache memory". Other factors included in the final quality measure are the metric related to mobility and interaction. New quality characteristic created are "touch screen facilities", "safety while driving" and "automatic update". Having refined the characteristics, the final list is exhibited in Table 2.

TABLE 2: Quality Characteristics, Goals and Measures

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

B. GQM Model

Five goals obtained from SLR method are the input into GQM approach. The measures shown in Table 2 support the derivation of the usability metric. The goals contain measurement goals regarding usability of mobile phone application, questions are a set of questions used to characterize how the goals should be attained, and the metric contains objective and subjective quantitative metrics that are collected in order to answer the questions.

To apply the GQM approach and acquire usability metrics, this generated a series of questions from the goals determined in Table 2. This study carefully created the questions by refining the goals into several questions and ensured the questions we created can be answered. To achieve the goal, the questions need to be answered by providing usability metrics. Figure 3 describes the GQM model, consists of the subjective usability metric.
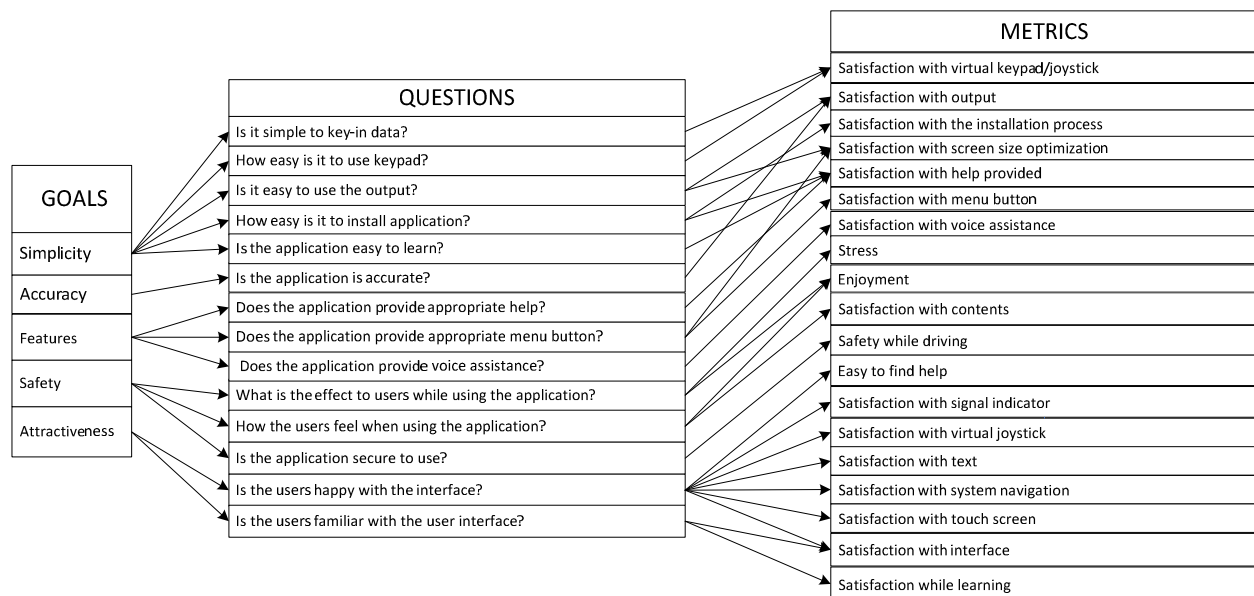

FIGURE 3: GQM Model

5. USABILITY STUDY

This section describes the implementation of usability study in validating the GQM Model. First, the participants experienced the selected mobile applications (on mobile phone). Then they were interviewed regarding their opinion and perception on using the application. The focal point of having interview session is to gather subjective measures. It was carried out informally to ensure participants can express their happiness or unhappiness during the test. 30 participants were involved in the usability study, were mixed between genders, expertise level, ages between 25 and 40 years old.

A. Usability Test Design

There are a huge amount of mobile applications installed on different mobile phones and operating systems. In order to ensure GQM model can be applied to many applications, four different applications were selected from four different categories, as suggested by Gafni [33]. This is essential in generalizing whether the GQM model can be applied to different types of applications. The four opted applications in this study include satellite navigation system, social network application, business application, and a game.

Having determined the types of application, four mobile phone applications were selected as test cases. These are CoPilot (SatNav), Facebook (social networking), Google apps (business apps), and TicTacToe (game). To determine whether different mobile phone platforms influence the metric, the usability tests were conducted using two different platforms; iPhone and O2 Orbit. There are many other platforms now becoming popular such as Android, but this study chooses Windows Mobile Professional 6 because it's available in many mobile phone devices and iOS iPhone is currently

at the top.In the test, each participant was required to experience two applications on two platforms before answering questions from the GQM. This makes the data complete for four sets, particularly 1) CoPilot on iPhone, 2) CoPilot on O2 Orbit, 3) Facebook on iPhone, and 4) Facebook on O2 Orbit.

B. *Interview Session*

The participants were interviewed to collect subjective data and their perception as suggested by Cairns and Cox [34] and [35]. Data from interviews can provide an understanding on participants' satisfaction on their interaction with the mobile phone applications. Particularly, this study made use of in-depth interviews; to ensure participants are able to talk freely about the mobile applications in an informal environment. This also provides an opportunity to discuss about the advantages and weaknesses of the applications. As a good guide, pre-prepared questions as listed in Table were prepared.

TABLE 3: Pre-prepared Questions for interview session

No	Questions	Rationale
1	How do you feel after using this application?	To determine whether participant enjoyed using the application.
2	Do you think the application was really helpful?	Satisfaction with the application contents.
3	Do you agree with the arrangement of the menu and submenu?	Satisfaction with the application layout.
4	Do you have any suggestion to improve the application?	Determine additional content required by participants, or any ideas to improve the layout.
5	Are you satisfied with the voice assistance provided?	Satisfaction with voice assistance.
6	What is your opinion on virtual keypad while inputting the data?	To obtain any comments on the virtual keypad.
7	If you had chance to use virtual joystick, did you enjoy using it?	To obtain any comments on the virtual joystick.
8	Do you think it is necessary to have this application in mobile phone?	To obtain feedbacks on contents.
9	What is your suggestion to make mobile phone user keen to use mobile application?	Discover the limitation of mobile phone application.

6. RESULT AND DISCUSSION

Data collected from the usability study were transcribed using NVIVO version 8. It was intended for data management and to classify the feedback from participants on mobile applications. This study also observed on how participants implement each task given to them to assess whether they have problems or any difficulties during usability test. All problems and difficulties faced by participants were recorded. The voice recordings of the interview sessions were fully transcribed and checked. A new project in NVIVO was created with nineteen internal nodes matching the number of subjective measures in the GQM model. The transcripts were analyzed and relevant statements were put into relevant nodes. The comments from participants were further distinguished between comments that refer to iPhone or O2 Orbit phone because NVIVO is able to. On top of that, this study separated the positive and negative feedbacks for both devices.

When the participants were asked about using applications on mobile devices in general, almost all participants agree that the applications on mobile phone are necessary and cannot be avoided. Among comments from the participants include:

"I think it's okay to have application on mobile phone. We have to utilise our mobile phone"

- Participant 2

"I think the demands are increase to use application on mobile phone and become ordinary to mobile phone user"

- Participant 6

"Well I think to be honest; I just use mobile phone to make a call and text only but I can see it quite useful. I don't have a time as well to play a game. I think the business device like iPhone is useful with many application and games"

- Participant 7

However, Participant 8 does not agree and he was suggesting that the mobile phone should be used to make a call or text only. He is the only one out of thirty participants does not agree that the applications on mobile devices were helpful. Participants also mentioned that they were happy with mobile applications because they can access much information anytime anywhere. However, some participants' comments on disadvantages of mobile applications as follows:

"Small screen, slow processor, unable to store huge data"

Participant 23

"Small text, we need to pay for internet connection"

Participant 29

In terms of mobile devices platforms, participants were very happy and they mentioned several time that the applications on iPhone are very impressive, attractive, and user friendly. In short they agreed that a good design of mobile application installed in a very good mobile device certainly provides the best experience to mobile users. However, some participants mentioned that the touch screen on iPhone is quite sensitive and they made some mistake while using the iPhone keypad. A few participants also mentioned that they are unable to use glove during CoPilot test on iPhone. Incidentally the tests were conducted in the winter season. Nevertheless, using applications on O2 orbit phone are also helpful. Although participants were not very happy with applications on O2 Orbit, they still considered that the applications on O2 phone were helpful as mentioned by two participants below:

"Well I'm happy and I think it's useful"

Participant 1 (During CoPilot test on O2 phone)

"This application is useful because it's mobile, we can access anytime anywhere, the device is small, make the user easy to carry everywhere"

Participant 2 (During Facebook test on O2 phone)

Participants were concerned about the battery consumed by CoPilot application. Some participants do not agree to use CoPilot on mobile phone and they prefer to use original SatNav device. They were also not very satisfied with satellite connection time particularly while using CoPilot on O2 Orbit phone.

On top of that, participants also commented on Facebook application in O2 phone. Particularly, it has a small virtual keypad and difficult to find the desired menu as mentioned by participant 4 below. Fonts in Facebook application are quite small particularly in O2 Orbit phone. Almost all participants mentioned about the small font on O2 phones in think aloud protocol during the usability test. Most of participants also refuse to use stylus as an input devices because they think stylus is a traditional input devices.

"Menu or submenu is quite difficult to find. Need more time to find it. Sometimes the menus are confusing"

Participant 4

Participants were very happy with the menu arrangement on Facebook in iPhone but not in O2 phone as mentioned by participant 30 who hates having to keep pushing the back button many times to go back to the main menu. Many users did

not realize that they can keep documents on Google Apps and they can access the document via mobile phone anytime anywhere. Almost all users are happy with the Google Docs services but some of them were happy to use Google Docs only on iPhone.

Participants also satisfied while playing games particularly the games on iPhone. Most of the comments from players are on the touch screen and they were very satisfied on iPhone's touch screen. Players also mentioned that the game on iPhone have a nice and pleasant design compared to the game on O2 phone with only green color and small text showing the score.

Additionally, this study also compares the overall feedbacks obtained through the interview. The feedbacks were separated into two which are feedback while using application on O2 Orbit and feedback while using applications on iPhone. The feedbacks were separated by creating another two nodes in NVIVO which are 'positive feedbacks' and 'negative feedback'. This study then summarized the number of positive and negative feedbacks about the platforms in Table 4. The feedbacks were obtained from 30 participants during the interview. The results clearly indicate that the participants were very happy with applications on iPhone compared to O2 Orbit phone. Based on the results, this study concludes that the GQM model has the ability to evaluate mobile applications. However this evaluation is limited only to subjective measures.

TABLE 4: Feedback based on Mobile Platforms

Platform	Applications	Positive feedback	Negative Feedback
O2 Orbit	CoPilot	70	24
	Facebook	48	50
	Google	45	53
	Game	36	33
iPhone	CoPilot	91	8
	Facebook	83	8
	Google	91	12
	Game	75	6

7. CONCLUSION

The research described in this paper has investigated subjective evaluation for mobile applications and provides a metric-based model to measure usability. The measures used to evaluate applications on mobile devices have been investigated. With regards to the efforts in this study, an evaluation model named GQM has been developed. The model contains a set of usability metrics to measure the usability of mobile applications. The model is able to help evaluators to measure usability of applications on mobile devices.

This model has been validated by carrying out a usability study. The model was used to evaluate four different mobile applications. The main purpose of carrying out the usability study is to determine whether the GQM is effective to collect appropriate data, analyzed, and provide results. The findings of this part of validation explain that the model is useful in evaluating the usability of mobile applications.

The used of SLR and GQM approach to develop the model seems very helpful. SLR is a method to search for literatures on particular research area. This systematic approach provides step by step guidance on how to search, evaluate, and interpret all available research relevant to a particular topic area or phenomenon of interest. Additionally, this study also shows that GQM approach is able to provide usability metrics for mobile applications. The metrics has also proven the capability to evaluate mobile applications during the usability study. GQM is a top-down approach that provides a direction to create the questions and the metrics. Although the approach has been widely used in software improvement, it can be expanded to a wider context. This research has shown that GQM can also be used to derive usability metrics for mobile applications. However, in order to use the model, evaluators need to design the test because the model did not describe on how to implement the test.

The usability metrics presented in this thesis will need to be updated in the future due to the rapid changes in mobile technology. Changes in mobile technology will affect the mobile applications. With GPS devices in mobile phones, many applications have been developed e.g. Mobile Millennium for traffic report and MyGeoRecorder to record personal location-based. Applications to edit HD video also have been developed due to high resolution camera embedded in mobile device. The number of mobile applications nowadays is increasing significantly. Applications in iTunes increase from more

than 3000 in 2008 [36] to 500,000 until October 2011 [37]. This increase will affect the model we have developed and need to be updated in future.

REFERENCES

[1] K. Hornbæk and E. L.-C. Law, "Meta-analysis of correlations among usability measures," in *Proceedings of the SIGCHI conference on Human factors in computing systems* San Jose, California, USA: ACM, 2007.

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

[3] K. C. Constantinos and K. Dan, "A research agenda for mobile usability," in *CHI '07 extended abstracts on Human factors in computing systems* San Jose, CA, USA: ACM, 2007.

[4] D. Zhang and B. Adipat, "Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications," *International Journal of Human-Computer Interaction,* vol. 18, pp. 293 - 308, 2005.

[5] J. Brodkin, "Mobile phones to be primary Internet device by 2020, experts predict." vol. 2008: Network World, 2008.

[6] J. Nielsen, *Usability Engineering*. London: Morgan Kaufmann, 1994.

[7] ISO, "INTERNATIONAL STANDARD: ISO 9241-11(Guidance on usability)," Geneva 1998.

[8] B.-L. D. Henry, C. B. T. Gerald, and C. Vivian Hsueh-hua, "Usability evaluation for mobile device: a comparison of laboratory and field tests," in *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, Helsinki, Finland, 2006.

[9] M. O. Leavitt and B. Shneiderman, "Research-Based Web Design & Usability Guidelines," Washington: U.S. Department of Health and Human Services, 2006.

[10] S. Alfredo, Oleg Starostenko, E. A. Castillo, and M. Gonzalez, "Generation of usable interfaces for mobile devices," in *Proceedings of the 2005 Latin American conference on Human-computer interaction* Cuernavaca, Mexico: ACM, 2005.

[11] Webcredible, "7 usability guidelines for websites on mobile devices," London: Webcredible Ltd., 2007.

[12] Nokia, "Top 10 Usability Guidelines for Enterprise Applications," Nokia Forum, Nokia Corporation, 2008.

[13] Openwave, "Top 10 Usability Guidelines for WAP Applications," Redwood City, USA: Openwave, 2008.

[14] G. Jun and P. Tarasewich, "Guidelines for handheld mobile device interface design.," in *Proceedings of DSI 2004 Annual Meeting*, Boston, 2004.

[15] S. K. S. Daniel, "Usability guidelines for designing mobile learning portals," in *Proceedings of the 3rd international conference on Mobile technology, applications and systems* Bangkok, Thailand: ACM, 2006.

[16] M. F. Bertoa, J. M. Troya, and A. Vallecillo, "Measuring the usability of software components," *Journal of Systems and Software,* vol. 79, pp. 427-439, 2006.

[17] S. Caplan, "Using focus group methodology for ergonomic design," *Ergonomics,* vol. 33, pp. 527 - 533, 1990.

[18] V. Basili, G. Caldeira, and H. D. Rombach, "The Goal Question Metric Approach," *Encyclopedia of Software Engineering,* 1994.

[19] R. Solingen and E. Berghout, *The Goal/Question/Metric Method: A Practical Guide for Quality Improvement of Software Development*. Maidenhead, England: McGraw-Hill Publishing Company, 1999.

[20] P. Botella, X. Burgués, J. P. Carvallo, X. Franch, G. Grau, J. Marco, and C. Quer, "ISO/IEC 9126 in practice: what do we need to know?," in *Software Measurement European Forum, (SMEF) 2004.*, Milan, 2004.

[21] A. A. A. Ghani, K. T. Wei, G. M. Muketha, and W. P. Wen, "Complexity Metrics for Measuring the Understandability and Maintainability of Business Process Models using Goal-Question-Metric (GQM)," *International Journal of Computer Science and Network Security,* vol. 8, 2008.

[22] A. Lerina, B. Thierry, C. Gerardo, and T. Maria, "A Framework for Measuring Business Processes Based on GQM," in *Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 1 - Volume 1*: IEEE Computer Society, 2004.

[23] H. Azham and E. Ferneley, "Usability metric for mobile application: a goal question metric (GQM) approach," in *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services* Linz, Austria: ACM, 2008.

[24] L. Luigi, "Providing Automated Support for the GQM Measurement Process," *IEEE Software,* vol. 17, pp. 56-62, 2000.

[25] R. Hirschheim, "Information systems epistemology: An historical perspective," in *Research Methods in Information Systems*, R. H. E. Mumford, R. Fitzberald et al., , Ed. Amsterdam: North-Holland Publishing Co., 1985, pp. 13-35.

[26] D. Silverman, "Qualitative research: meanings or practices?," *Information Systems Journal,* vol. 8, pp. 3-20, 1998.

[27] Galliers, "Choosing information systems research approaches," in *Information Systems Research: Issues, Methods, and Practical Guidelines*, R. D. Galliers, Ed. Oxford: Blackwell Scientific Publications, 1992.

[28] I. Benbasat, "An Analysis of Research Methodologies," in *The Information Systems Research Challenge*, F. W. McFarlan, Ed. Boston, Massachusett: Harvard Business School Press, 1984.

[29] B. Kitchenham, "Procedures for undertaking systematic reviews.," Technical Report TR/SE-0401, Department of Computer Science, Keele University and National ICT, Australia Ltd, 2004. Joint Technical Report. 2004.

[30] K. Hornbæk, "Current practice in measuring usability: Challenges to usability studies and research," *International Journal of Human-Computer Studies,* vol. 64, pp. 79-102, 2006.

[31] J. Kjeldskov, C. Graham, S. Pedell, F. Vetere, S. Howard, S. Balbo, and J. Davies, "Evaluating the usability of a mobile guide: the influence of location, participants and resources," *Behaviour and Information Technology,* vol. 24, pp. 51-65, 2005.

[32] N. Bevan and M. MacLeod, "Usability measurement in context," *Behaviour & Information Technology,* vol. 13, pp. 132 - 145, 1994.

[33] R. Gafni, "Framework for Quality Metrics in Mobile-Wireless Information Systems," *Interdisciplinary Journal of Information, Knowledge, and Management,* vol. 3, pp. 23 - 38, 2008.

[34] P. Cairns and A. L. Cox, *Research methods for human-computer interaction*. Cambridge: Cambridge University Press, 2008.

[35] C. McNamara, "General guidelines for conducting interviews," 2009.

[36] Apple, "App Store Downloads Top 100 Million Worldwide," 9 September 2008 2008.

[37] Apple, "The App Store," 2011.