Field Testing of Mobile Applications: An Effective Approach

Naveen Singh
Software Testing Engineer

Ravi Shanker Yadav
Research Associate

ABSTRACT

Field testing of software applications developed for mobile devices is a challenge due to unique features of mobile devices. Unreliability of wireless networks, limited bandwidth, changing context, phasing network disturbance or noises problems and different data transfer time for different applications. Methodologies used in testing of desktop applications cannot be directly applicable to a mobile context. The large number of users involved in field testing along with the variety of problems reported by them increases the complexity of managing the field testing process. The contribution of this paper is to propose a generic framework, detecting noise and estimation of time for different applications to conduct field tests for mobile applications with focus on field testing, detecting network disturbance and estimation of time for different mobile applications.

Keywords

Field, Testing, Laboratory, Usability, Mobile

1. INTRODUCTION

With the regular advancement in wireless technology and mobile devices such as cell phones, personal digital assistants (PDAs), palms, and pocket PCs, many innovative mobile applications are emerging, aiming to enhance wireless communication and provide users with ubiquitous access to information (Li & Liao,2000). Many businesses have deployed mobile applications to gain competitive advantage. Such applications developed specifically for small mobile devices include daily news alert services, classified mobile advertising, restaurant and entertainment listings, wireless Web portals, and mobile commerce (m-commerce) applications (Varshney & Vetter, 2002).The high demand and fast growth of mobile applications have attracted extensive research interests. Because developing mobile applications with an easy-to-use interface is critical for successful adoption and use of applications, one of the important research issues is regarding how to conduct an appropriate usability test using mobile devices in a wireless environment. Usability testing is an evaluation method used to measure how well users can use a specific software system.

An effective usability test has to be able to elicit feedback from users about whether they use an application without (or almost without) difficulty and how they like using the application, as well as evaluate levels of task performance achieved by users (Wichansky, 2000).There are various guidelines for usability testing of desktop applications. However, those established concepts, methodologies, and approaches commonly used in traditional human-computer interaction research are not always applicable to mobile applications(Jones et al., 1999) due to mobility and the distinct features of mobile devices and Wireless networks. Ideally, usability testing of mobile applications should be carefully designed to cover all or most possible situations of a mobile environment (Kim et al.,2002).

In reality, however, this poses many challenges. For example, it is difficult to Foresee the exact situations of the application use – users may be standing, walking, or sitting in a dark or bright environment while using an application. As a result, a usability test may have to concentrate only on certain aspects of a mobile application and sacrifice others. Furthermore, traditional research methodologies used in usability testing, including controlled laboratory experiments and field studies have various limitations in a mobile environment, such as ignoring the mobile context or lack of sufficient procedural control. Therefore, it is essential to develop guidelines for usability testing of mobile applications.

Usability testing is a common tool used to evaluate the usability of a mobile application in a development process. Usability tests are usually conducted using a think aloud protocol based on K. A. Ericsson and H. A. Simon’s work (1980, 1984). Users are given tasks in a test environment and encouraged to think aloud while trying to accomplish the tasks. This gives us, usability practitioners, information we need on how the user interface matches the natural human way of thinking and acting and highlights the features and processes to be improved. Severity of the usability problems is an important factor when defining the urgency of actions related to a problem. The most urgent actions are needed when the problem prevents completion of the task. Dumas and Redish (1993) use four point scale, where the first severity level represents the most severe problems and the last the least severe. Also Kallio et al. (2004) divide the severity of problems into categories; high (failure in task execution), medium (not so severe, task can be executed) and low (minor problems).

2. FIELD VERSUS LAB TEST

Field tests are traditionally conducted in test laboratories, consisting of e.g. a living room or office-like area connected to a monitoring area with a one-way mirror. The laboratory environment is a peaceful space, where a test user can concentrate on the given tasks.

Usability researchers and practitioners have been concerned that laboratory evaluations do not simulate the context where mobile phones are used (Johnson 1998) and lack the desired ecological validity. Interruptions, movement, noise, multitasking etc. (Tamminen et al. 2004) that could affect the users’ performance are not present in laboratory tests. The surrounding environment and mobility are assumed to set special requirements for mobile applications. Usability testing should take these requirements into account. Even if there seems to be a common concern about the adequateness of laboratory evaluations, field evaluations have been rather rare. A literature study by Kjeldskov and Graham (2003) revealed that most (71%) mobile device evaluations were done in laboratory settings. This may be due to data collection techniques such as think aloud, video recording or observations being difficult in the field.As mobile video recording systems, like small video cameras, have rapidly developed during past few years and conducting user tests in the field has become easier. It is now possible to attach a small camera to record the screen of the mobile device and collect that information for later review (Kjeldskov et al. 2004a, Roto et al. 2004). This development allows similar test setting in the field as in the laboratory; it is possible for test leader to follow what is happening on the screen and hear users' comments. This also allows the usage of think aloud protocol in usability test in the field. Despite the development of suitable tools testing in the field is still likely to be more time consuming (Kjeldskov et al. 2004a) than in laboratory setting. It may also require extra effort from test users and the test leader. Resources for application development are limited in the mobile industry, and usability activities such as user-cantered design and usability testing must be made very efficiently. The goal in a product development process is to find the biggest and most fatal usability problems within the strict limitations of project budgets and deadlines. The focus of the usability inspection is not on finding every possible detail. Decisions made by usability expert when planning usability tests are related to risk Management; how to optimize the effort and the outcome (Nielsen and Landauer 1993). Choosing the right evaluation method is important; scientifically validated information on suitable testing methods is valuable for usability practitioners. Kjeldskov, Skov and Stage (2004b) presented a good example of information practitioners' need when making decision on the method in their article Instant Data Analysis: Evaluating Usability in a Day. Hertzum (1999) compared role of three different methods (laboratory tests, workshops and field tests) in a product development cycle. His goal was also to increase the efficiency of the tests. In our study, the main question is to find out whether field tests are critical when evaluating mobile application usability or can the sufficient ecological validity be simulated in laboratory environment?

3. ISSUES IN TESTING OF MOBILE APPLICATIONS

The unique features of mobile devices and wireless networks pose a number of significant challenges for examining usability of mobile applications, including mobile context, multimodality, connectivity, small screen size, different display resolutions, limited processing capability and power, and restrictive data entry methods.

3.1 Display

The display capability of mobile devices supports less display resolution (normally 640*480 pixels or below) in comparison with desktops. Low resolution can degrade the quality of multimedia information displayed on the screen of a mobile device.

3.2 Small screen

Small screen size, can significantly affect the usability of mobile applications (Jones et al., 1999; L. Kim & Albers, 2001).Direct presentation of most WWW pages on small mobile devices can be aesthetically unpleasant, un-navigable, and in the worst case, completely illegible (Bickmore, 1997).

3.3 Connectivity

This problem largely affects data downloading time and quality of streaming media (e.g., video and audio streams). Strength of signals and data transfer speed in a wireless network may vary at different time and locations, compounded by user mobility (Sears & Jacko, 2000).

3.4 Constraint Power and Processing Capability

applications that require a large amount of memory for graphic support or fast processing speed, such as an application of 3D city maps for PDAs (Rakkolainen & Vainio, 2001), may not be practical for mobile devices. Because of limited processing capability of mobile devices, developers may have to disable some functions.

4. ENHANCED PARAMETERS

Field testing of mobile application requires Laboratory testing and field testing both but field testing is very essential because only laboratory testing cannot assure actual functionality of mobile applications. In this paper we focussed on field testing of mobile application, detecting noise during test and estimation of time for different applications. Mainly we have proposed given below points.

4.1 Route Identification

For performing field testing tester should have information regarding availability of cells in various locations (Mobile QA Zone) of the area where tester want to perform test. Drive route identification is the important aspect in field testing, it decides which operator area should be covered and what all are the test scenarios to be executed in the drive route etc.Lack of information of availability of cells for performing field testing may be more time consuming. If more time will spend than budget of project also will increase.

4.2 Location for field test

Testing performed on a specific device, handset or instrument in the defined region(FIELD).Now after having the information regarding availability of cells in various locations of area where tester wants perform test, tester can select location for testing applications as per his/her convenient and can save time .

4.3 Strength of signals

Tester should have information about availability of level of network. He/She should know that where signals are weak, where signals are strong and where there are no availability of signals. result of the testing depends on the strength of signals .from strong to weak signals performance of application would vary.

4.4 Measuring level of network disturbance or noises

Noise is one of the most common occupational health hazards. In heavy industrial and manufacturing environments, as well as in farms, cafeterias, permanent hearing loss is the main health concern. Annoyance, stress and interference with speech communication is the main concern in noisy offices, schools and computer rooms. To prevent adverse outcomes of noise exposure, noise levels should be reduced to acceptable levels. The best method of noise reduction is to use engineering modifications to the noise source itself, or to the workplace environment. Where technology cannot adequately control the problem, personal hearing protection (such as ear muffs or plugs) can be used. Personal protection, however, should be considered as an interim measure while other means of reducing workplace noise are being explored and implemented. As a first step in dealing with noise, workplaces need to identify areas or operations where excessive exposure to noise occurs. Noise can be continuous, variable, intermittent or impulsive depending on how it changes over time. Continuous noise is noise which remains constant and stable over a given time period. The noise of boilers in a power house is relatively constant and can therefore be classified as continuous. Most manufacturing noise is variable or intermittent. Different operations or different noise sources cause the sound changes over time. Noise is intermittent if there is a mix of relatively quiet periods and noisy. Impulse or impact noise is a very short burst of loud noise which lasts for less than one second. Gun fire or the noises produced by punch presses are examples of such noise. When we talk about noise, we are talking about unwanted signal. We 'break out' noise into three basic categories; that is background noise, modulated noise and interference noise. Background noise is the noise we always want to stay above, it is sometimes called the 'noise floor'. Signal can always be amplified above the noise floor but once it gets buried in background noise (falls into the signal floor) it cannot be retrieved; this is why the LNB amplifies the weak satellite signals as soon as they are received before passing them on into the cable and to your receiver. Modulated noise is undesirable signal that enters into a system and rides on a signal, using your system power, producing undesirable side effects in video quality. Interference is noise that comes in on the same frequency as signal and masks (overwhelms) parts or all of the desired signal.

External noise: There are many possible sources of external noise. There can be noise factors that are part of the photographic process, a smudge, or a bad spot on the film. Or something in the person's lung that is fine but just looks a bit like a tumor. All of these are to be examples of external noise. While the doctor makes every effort possible to reduce the external noise, there is little or nothing that they can do to reduce internal noise.

Internal noise: Internal noise refers to the fact that neural responses are noisy. To make this example really concrete, let's suppose that our doctor has a set of tumor detector neurons and that they monitor the response of one of these neurons to determine the likelihood that there is a tumor in the image (if we could find these neurons then perhaps we could publish and article entitled ``What the radiologist's eye tells the radiologist's brain"). These hypothetical tumor detectors will give noisy and variable responses. After one glance at a scan of a healthy lung, our hypothetical tumor detectors might fire 10 spikes per second. After a different glance at the same scan and under the same conditions, these neurons might fire 40 spikes per second.

Properties of noise that can be measured

- frequency
- sound pressure
- sound power
- time distribution

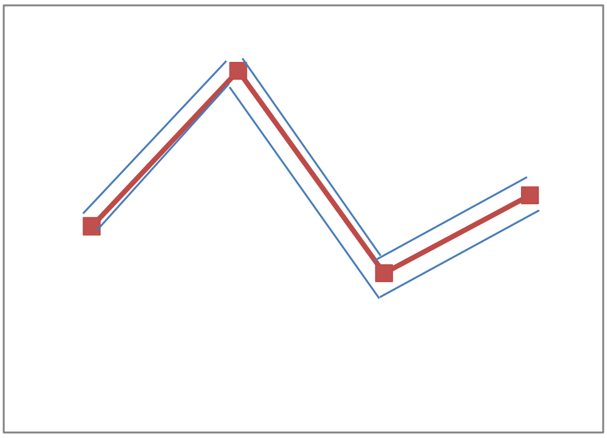

Fig.1 Noise modulated into signal.

Measurements of impulse or impact noise depend on the guidelines and standards in force. Before you measure impact or impulse noise, you must ensure that the equipment has the capacity to measure this kind of noise. Normally measurements of either peak noise levels together with the actual number of peaks, or percentage dose or equivalent sound levels are required. Where there is little background noise, as for example on an outdoor rifle range, the measuring of peak pressures may be most appropriate. In industrial settings, there is usually considerable background noise in addition to the impulse noise. In such cases, provided that a 3 dB (A) exchange rate is used, dosimeters or ISLMs which are sufficiently sensitive to respond well to peaks may be more appropriate. One can account for all of the noise, continuous and impulse, in the one measurement. The SLM consists of a microphone, electronic circuits and a readout display. The microphone detects the small air pressure variations associated with sound and changes them into electrical signals. These signals are then processed by the electronic circuitry of the instrument. The readout displays the sound level in decibels. The SLM takes the sound pressure level at one instant in a particular location. The integrating sound level meter (ISLM) is similar to the dosimeter. It determines equivalent sound levels over a measurement period. The major difference is that an ISLM does not provide personal exposures because it is hand-held like the SLM, and not worn. The ISLM determines equivalent sound levels at a particular location. It yields a single reading of a given noise, even if the actual sound level of the noise changes continually. It uses a pre-programmed exchange rate, with a time constant that is equivalent to the SLOW setting on the SLM. A noise dosimeter is a small, light device that clips to a person's belt with a small microphone that fastens to the person's collar, close to an ear. The dosimeter stores the noise level information and carries out an averaging process. It is useful in industry where noise usually varies in duration and intensity, and where the person changes locations.

From below given Table-1 instrument can be selected and can be embedded in mobile devices to detect network disturbance or noises.

Table-1 Guidelines for Instrument Selection

Guidelines for Instrument Selection		
Type of Measurement	**Appropriate Instruments (in order of preference)**	**Comments**
Personal noise exposure	1) Dosimeter	Most accurate for personal noise exposure.
	2) ISLM*	If the worker is mobile, it may be difficult to determine a personal exposure, unless work can be easily divided into defined activities.
	3) SLM**	If noise levels vary considerably, it is difficult to determine average exposure. Only useful when work can be easily divided into defined activities and noise levels are relatively stable all the time.
Noise levels generated by a particular source	1) SLM**	Measurement should be taken 1 to 3 metres from source (not directly at the source).
	2) ISLM**	Particularly useful if noise is highly variable; it can measure equivalent sound level over a short period of time (1 minute).
Noise survey	1) SLM	To produce noise map of an area; take measurements on a grid pattern.
	2) ISLM	For highly variable noise.
Impulse noise	1) Impulse SLM	To measure the peak of each impulse.
* SLM stands for Sound Level Meter ** ISLM stands for Integrating Sound Level Meter		

4.5 Calculation of data transfer time for different mobile applications

Some time it happens that at the same locations two different mobiles take different time to transfer data. reason behind this is that speed of data transfer for different type of files is different. we can improve the performance of mobile by estimating time for different files. Data transfer is the movement of digital information from place to place, whether over copper wires, fibre optic cable, or via radio waves. Bandwidth, the rate at which data can be transferred via these various media, is measured in units such as megabits per second. As recent new cell phone technologies have expanded data transfer capabilities, an explosion of new ways to use the added capacity of cell phone networks has taken place. As more devices take advantage of these technologies.

Factors affecting network load to consider for various mobile applications architectures are listed in the below table.

Table-2 Network Load Factors by Application Architecture

Application Architecture	**Network Load**	**Mbit Per Display**	**Comments**
Web Applications (Dynamic)	Medium	2	ADF based applications typically have a heavier network load then REST API applications.
Web Applications (Cached)	Medium	1	ADF based applications typically have a heavier network load then REST API applications.
Mobile Applications	Light	0.05	Only incremental updates are periodically sent.
Services (REST)	Medium	2	REST based services can be designed for lightweight loads.

4.5.1 Significant factors that can be adjusted that affects network load are

Internet browser caching policy - Some organizations may disable local caching, increasing network load.

Percent of new users - New users have no data cached on their local systems, increasing network load.

Image Size - The smaller the image displayed on the client, the less network traffic is required.

Image Format- JPEG is typically lightest for raster and PNG lightest for vector.

Blending Location - Applications can be designed to blend images at client or server-side.

5. VISUALIZATION OF THE RESULTS OF FIELD TEST

Software must endure rigorous field testing before market deployment. Field testing helps uncover problems. Field testing of software is necessary to find potential user problems before market deployment. The field testing helps the collection of field data in real-time without developer interference. The report also records the user who had the problem. Developers must analyze and resolve these reported problems. To verify the repair of a problem or to understand its peculiarity, developers must often replicate the scenarios which trigger the problem By understanding the characteristics of the users reporting a problem and its peculiarities, developers should be able to gain insight into ways of resolving it. To provide quality in mobile applications deeper analysis and tracking of field testing results is needed. As we have focussed in this paper on two parameters disturbance of network or noises and data transfer time for various mobile applications. Noises can be detected using given techniques. Why the data transfer time is more? factors affecting transfer rate is also explained here. From these two parameters and their reasons can solve the problems that occur in mobile applications.

6. CONCLUSION

In this paper, we present an effective field testing approach to reduce the problems of noises and delay in data transfer. With the advancement in mobile technology and applications, effective field testing becomes increasingly important for the deployment of successful mobile applications. Therefore, it is essential to develop and adopt appropriate research methodologies and tools to evaluate the field testing of mobile applications. The latest advance of mobile technology and increasing wireless network bandwidth makes it a reality for users to gain access to multimedia information (combines several communication media such as text, graphics, video, animation and sound) available on the Internet or other sources from mobile devices. However, the constraints of mobile devices and wireless communication pose a variety of challenges for mobile devices to handle mobile multimedia applications.

For example, low wireless network bandwidth may cause significant transmission delay, which can affect both user perception and performance while using mobile applications. Although many mobile multimedia applications have used audio or video compression techniques to compress multimedia content in order to reduce the file size and shorten the transmission delay, such data compression can result in reduced quality of multimedia content presented on mobile devices. So far, most usability studies of mobile applications deal with either traditional database data or textual documents. Few studies have focused on usability testing of multimedia applications.

7 ACKNOWLEDGMENTS

This work is supported by Mobile QA Zone and Prof. Dr. M. Rizwan Beg, Head of department of computer science and engineering, Integral University, Lucknow. They would also thank the anonymous reviewers for their significant and constructive critiques and suggestions, which substantially improved the quality of this paper.

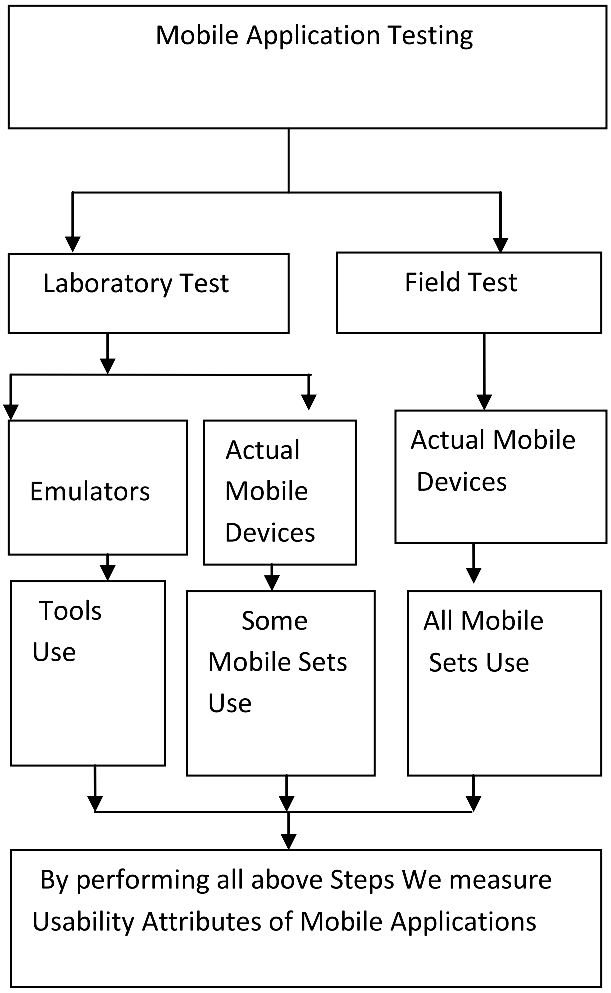

Fig.2 Framework for Implementing Field Testing

8. REFERENCES

[1] Bautsch - Vtense, H. S., Marmet, G. J., & Jacko, J. A. (2001). Investigating PDA web browsing through eye movement analysis. In Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality (pp. 6-10).

[2] Mahwah, NJ: Lawrence Erlbaum Associates.

[3] Beck, E. T., Christiansen, M. K., & Kjeldskov, J. (2003). Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in a Laboratory Setting. *Proceedings of OzCHI 2003*, Brisbane, Australia.

[4] Bederson, B. B., Clamage, A., Czerwinski, M. P., & Robertson, G. G. (2003, April 5-10).A fisheye calendar interface for PDAs: providing overviews for small displays. Proceedings of CHI'03 extended abstracts on Human Factors in computing systems,Ft. Lauderdale, FL.

[5] -Bederson, B. B., Czerwinski, M. P., & Robertson, G. G. (2002). A Fisheye Calendar Interface for PDAs: Providing Overviews for Small Displays. Technical report (No.#HCIL-2002-09): University of Maryland College Park.

[6] Björk, S., Holmquist, L. E., Redström, J., Bretan, I., Danielsson, R., Karlgren, J., et al.(1999). WEST: a Web browser for small terminals. Proceedings of the 12th AnnualACM Symposium on User Interface Software and Technology, Asheville, NorthCarolina.

[7] Björk, S., Redström, J., Ljungstrand, P., & Holmquist, L. E. (2000). POWERVIEW: Using information links and information views to navigate and visualize informationon small displays. Proceedings of Handheld and Ubiquitous Computing 2000 (HUC2K), Bristol, U.K

[8] Brachtl, M., Slajs, J., & Slavik, P. (2001). PDA based navigation system for a 3Denvironment. Computers & Graphics, 25, 627-634.

[9] Buchanan, G., Farrant, S., Jones, M., Thimbleby, H., & Pazzani, M. J. (2001, May 1-5).Improving Mobile Internet Usability. Proceedings of the 10th International Conference on World Wide Web, Hong Kong.

[10] Buyukkokten, O., Garcia-Molina, H., & Paepcke, A. (2001, May 1-5). Seeing the wholein parts: text summarization for web browsing on handheld devices. Proceedings of the tenth international conference on World Wide Web, Hong Kong.

[11] Buyukkokten, O., Garcia-Molina, H., Paepcke, A., & Winograd, T. (2002). Efficient WebBrowsing on Handheld Devices Using Page and Form Summarization. ACMTransaction on Information Systems, 20(1), 82-115.

[12] Christie, J., Klein, R. M., & Watters, C. (2004). A comparison of simple hierarchy and grid metaphors for option layouts on small-size screens. International Journal ofHuman-Computer Studies, 60(5-6), 564-584..

[13] Danesh, A., Inkpen, K., Lau, F., Shu, K., & Booth, K. (2001). Geney TM: Designing a Collaborative Activity for the Palm Handheld Computer. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, WA.

[14] Dey, A. K., Salber, D., & Abowd, G. D. (2001). A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. Human-Computer Interaction, 16, 2-4.

[15] Dumas, J. S., & Redish, J. C. (1999). A Practical Guide to Usability Testing. Intellect Book. Portland: Intellect.

[16] Ebling, M. R., & John, B. E. (2000). On the Contributions of Different Empirical Data inUsability Testing. Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods and Techniques, New York.

[17] Frokjaer, E., Hertzum, M., & Hornbaek, K. (2000, April 1-6). Measuring Usability: AreEffectiveness, Efficiency, and Satisfaction Really Correlated? Proceedings of theACM CHI 2000 Conference on Human Factors in Computing Systems, Hague,Netherlands.

[18] Garzotto, F., Mainetti, L., & Paolini, P. (1993). HDM - A Model Based Approach toHypermedia Application Design. ACM Transaction on Information Systems, 11(1), 1-26.

[19] Garzotto, F., Mainetti, L., & Paolini, P. (1994). Adding Multimedia Collections to theDexter Model. Proceedings of ACM Conference on Hypermedia Technology(ECHT'94), Edinburgh, UK.

[20] Dumas, J.R., Redish J.C (1993) A Practical Guide to Usability Testing. Ablex Publishing Corporation Norwood NJ.

[21] Ericsson, K. A., & Simon, H. A. (1980) Verbal Reports as Data. Psychological Review, 1980, 87, 215-251.

[22] Ericsson, K. A., & Simon, H. A. (1984) Protocol Analysis: Verbal Reports as Data. Cambridge, MA: Bradford Books / MIT Press.

[23] Faulkner, L. (2003) Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments and Computers 2003, 35 (3), 379- 383.

[24] Hertzum, M. (1999) User Testing in Industry: A case study of laboratory, Workshop and Field Tests. Proceedings of 5th ERCIM WORKSHOP ON "USER INTERFACES FOR ALL" http://ui4all.ics.forth.gr/UI4ALL-99/Hertzum.pdf

[25] Johnson, P. (1998) Usability and Mobility; Interactions on the move. In Proceedings of the First Workshop on Human-Computer Interaction with Mobile Devices, Glasgow, Scotland, GIST Technical Report G98-1.

[26] Kallio, T., Kekäläinen A (2004) Improving the Effectiveness of Mobile Application Design: User-Pairs Testing by Non-professionals. Proceedings MobileHCI 2004Conference, Glascow,UK Springer-Verlag 2004 315-319.

[27] Kjeldskov, J., and Graham, C. (2003) A Review of Mobile HCI Research methods. In Proceedings of the 5th International Mobile HCI 2003 Conference, Udine, Italy, Sringer-Verlag,

[28] Kjeldskov J., Skov M. B., Als B. S. and Høegh R. T. (2004a) Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field . In Proceedings MobileHCI 2004 conference, Glasgow, UK. Springer- Verlag, 2004. 61-73.

[29] Kjeldskov J., Skov M. B. and Stage J. (2004b) Instant Data Analysis: Evaluating Usability in a Day. In Proceedings of NordiCHI 2004, Tampere, Finland. ACM, 2004. 233-240.

[30] Kopomaa, T. (2000) City in your pocket. Birth of the Mobile Information Society. Gaudeamus, Helsinki 2000.

[31] Mäkelä, A., and Fulton Suri, J. (2001) Supporting Users' Creativity: Design to Induce Pleasurable Experiences. In M.G. Helander, H.M. Khalid, T. Ming Po (Eds.), Proceedings of International Conference on Affective Human Factors Design. ASEAN Academic Press, 2001. 387- 394.

[32] Nielsen J. and Landauer T.K. (1993) A Mathemetical Model of Finding Usability Problems. In proceedings of INTERCHI ’93. 206-213.

[33] Roto, V. Oulasvirta, A. Haikarainen, T. Lehmuskallio, H. & Nyyssönen, T. (2004) Examining mobile phone use in the wild with quasi-experimentation. HIIT Technical Report 2004-1, August 13.

[34] Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A. (2004). Understanding mobile contexts. Special Issue of Journal of Personal and Ubiquitous Computing. 8, 135-143.

[35] Wichansky, A. (2000). Usability test in 2000 and beyond. *Ergonomic, 43*(7), 998-1006.

[36] Kim, H., Kim, J., Lee, Y., Chae, M., & Choi, Y. (2002, January 07 - 10). An Empirical Study of the Use Contexts and Usability Problems in Mobile Internet. Proceedings of the 35th Hawaii International Conference on System Sciences (HICSS-35'02), Big Island, Hawaii.

[37] Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications**,** *by Zhang, D et al.*

[38] Visualizing the Results of Field Testing**,** *by Brian Chan et al*