A Quality Model for Human-Computer Interaction Evaluation in Ubiquitous Systems*

Rainara M. Santos[1,**], Káthia M. de Oliveira[2], Rossana M.C. Andrade[1,***], Ismayle S. Santos[1,†], and Edmilson R. Lima[1,‡]

[1] Group of Computer Networks, Software Engineering and Systems (GREat)
Department of Computer Science, Federal University of Ceará. Fortaleza, Brazil
{rainarasantos,rossana,ismaylesantos,edmilsonrocha}@great.ufc.br
[2] University of Valenciennes, LAMIH, CNRS UMR 8201 Valenciennes, France
kathia.oliveira@univ-valenciennes.fr

Abstract. The improvement in computational device miniaturization and in wireless communication has moved forward relevant advances in ubiquitous systems development. Such systems are capable of monitoring environments and users in order to provide services as naturally as possible. These systems offer new types of interactions, such as more implicit and transparent exchanges with users. Thus, the ubiquitous systems present new challenges in quality evaluation of human-computer interaction, as any assessment of quality should take into account the peculiarities of these new types of interactions. This paper proposes a quality model composed of specific characteristics and measures to human-computer interaction quality evaluation in ubiquitous systems. It also reports results obtained from a case study conducted to evaluate an application based on this model.

Keywords: Ubiquitous Systems, HCI Evaluation, Quality Model.

1 Introduction

Ubiquitous Computing is a new computing paradigm that proposes the adoption of computational devices in various sizes, shapes and functions to support users daily activities. These systems will be everywhere around users, connected to each other, and providing services which have to be as natural as possible. To achieve this, the applications are embedded in everyday objects and capable of monitoring user behaviour and environment [1]. To that end, the interaction between the user and the system is of utmost importance and the quality of this interaction has a direct impact on the use and adoption of the system. In this

* This work is a result of Maximum project supported by FUNCAP and CNRS under grant number INC-0064-00012.01.00/12.
** Master Scholarship (MDCC/DC/UFC) sponsored by CAPES.
*** Researcher scholarship - DT Level 2, sponsored by CNPq.
† PhD Scholarship (MDCC/DC/UFC) sponsored by CAPES.
‡ Undergraduate Scholarship sponsored by PIBIC/CNPq.

C. Collazos, A. Liborio, and C. Rusu (Eds.): CLIHC 2013, LNCS 8278, pp. 63–70, 2013.
© Springer International Publishing Switzerland 2013

scenario, it is necessary to assure that these systems support user activities in a transparent way with little or no need for attention or input from a user.

Considering then the software quality evaluation, it is usually supported by a quality model that defines a set of characteristics (usually known as abilities of a system, such as usability and maintainability). Such characteristics are often organized in a hierarchical tree that starts with a generic definition and develops into measures that allow for the product assessment. The most commonly used quality model is the ISO 9126 Standard [2]. This standard specifies both the usability characteristic and measures for evaluating the Human-Computer Interaction (HCI). However, the nature of ubiquitous systems suggests that new quality characteristics should be taken into account. For example, an evaluation of ubiquitous systems should value an implicit and transparent user interaction over an interaction that requires direct input from the user.

This paper proposes a quality model to support HCI evaluation in ubiquitous systems. This model consists of characteristics and sub-characteristics that have impacts on user interaction quality, and measures capable of evaluating them for a particular system. It is important to mention that is not our goal to propose a complete model with all possible characteristics and their sub-characteristics, but rather to define a model with primary characteristics necessary for evaluating the HCI in ubiquitous systems. We called this model TRUU Quality Model, which means Trustability, Resource-limitedness, Usability and Ubiquity. We also present results obtained from a case study using this quality model.

2 The Quality Evaluation of HCI in Ubiquitous Systems

Mark Weiser's vision of ubiquitous computing is well expressed in his famous quote: "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" [3]. This paradigm includes services and information provision from a variety of computers that support users in everyday tasks. This support should be executed without users needing to be aware that they are interacting with various computer technologies. To achieve this goal, ubiquitous systems have to comply with challenging requirements such as autonomy, heterogeneity, coordination of activities, mobility and context-awareness [1].

To support these new systems, it is evident that the quality evaluation for ubiquitous systems should take into account new characteristics to evaluate the interaction between the user and the system. With the goal of identifying those characteristics, we performed a large literature review using a systematic mapping (SM) study. SM is an empirical methodology that provides a wide overview of a research area to establish if research evidence exists on a topic [6]. Using this methodology[1], we found 369 papers pertinent to the subject. By reading all abstracts, we selected 60 papers. With a deep reading of all papers, we selected only 18 of these papers ([4][5][9][10][11][12][13][14][15][16][17][18] [19][20][21][22][23]

[1] Details of this SM is out of scope of this paper, but it can be found in `http://www.great.ufc.br/maximum/images/arquivos/protocolo.pdf`

[24]) that discussed some quality characteristics specific to the HCI in ubiquitous systems. Then, we found that: *a*) Only two papers propose a model [5][4], but they do not present all important ubiquitous features (for example, [5] does not consider availability and [4] does not consider transparency); *b*) Some of them [17] [16], not organized in a model structure, focus in only specific aspects ignoring important aspects from ubiquitous systems (for example, calmness [21] and transparency [5]); and *c*) The papers present incomplete measures to assess the characteristics, with no formulas, interpretation or collection methods.

As a result of this analysis, we concluded they could be complementary, because some of them have important characteristic that the others do not have. Thus, it is important to organize a more complete quality model focused in all aspects mentioned by those authors and that should be considered in an evaluation of the HCI in ubiquitous systems.

3 The TRUU Quality Model

To define a quality model to evaluate HCI in ubiquitous systems, we decide to analyze and synthesize the 18 papers found in the literature (see in the previous section) concerning the subject. Based on these findings, we joined together the existing propositions in a single quality model, specific for HCI evaluation and called TRUU as presented in Table 1.

Table 1. The TRUU Quality Model

Characteristics	Sub-characteristics	References
Trustability	Security	[4] [9] [10] [11] [12] [13]
	Privacy	[10] [13] [5] [14] [15]
	Control	[13] [5]
	Awareness	[5]
Resource-limitedness	Device Capability	[12] [5] [16]
	Network Capability	[16]
Usability	Satisfaction	[11] [5] [16] [17] [18] [19]
	Ease of Use	[4] [9] [13] [17] [20] [21]
	Efficiency	[11] [5] [13] [17]
	Effectiveness	[5] [13]
	Familiarity	[5]
Ubiquity	Context-Awareness	[4] [11] [14] [9] [21] [22] [23]
	Transparency	[5] [23] [18]
	Availability	[4] [9] [13] [23]
	Focus	[13] [12] [5]
	Calmness	[13] [18] [21]

As shown in Table 1, we found characteristics that can evaluate any type of system (for example, satisfaction) and others specific to ubiquitous systems evaluation (for example, transparency). Then, we propose to group these specifics characteristics (called sub-characteristics in the model) into the TRUU characteristics. For space reasons, we have only detailed here the sub-characteristics from Ubiquity (that has aspects related to ubiquitous systems), as follows:

- Context-Awareness is the ability of the system to use context to provide relevant services to user, where relevancy depends on the user's task [7];
- Transparency is "the extension of the system which consists of hidden components in the physical space and interaction is performed through natural interfaces" [5];
- Availability is the system's capability to provide continuous access to information resources anywhere and anytime [23];
- Focus is the system's capability to maintain the user's focus on the main task [5]; and
- Calmness prevents humans from feeling overwhelmed by information [18].

To complete the TRUU Quality Model, measures were defined for evaluating each sub-characteristic. In this paper, we focus on the measures for context-awareness that were used in the case study presented in the next section. To define the measures we used the Goal-Question-Metric method [8] and considered suggestions presented in [23], [11], [4] and [25]. Using GQM, we define our goal as to "**analyse** ubiquitous systems **for the purpose of** evaluating quality **with respect to** context-awareness **from the point of view of** the user.

To define the questions and measures, we had to analyse the meaning of being context-aware. [7] defines context-awareness when a system uses context information to provide relevant services to the user. This context can include any information used to characterize the situation of an entity, which is a person, place, or object considered relevant to the interaction, including the user and applications themselves. Some aspects of context-awareness directly impact HCI quality and, therefore, they need to be evaluated. One aspect is the adaptation correctness, which means the system adapts in a correct way, providing services and information correctly. Some factors that can influence this correctness are: the context correctness, as if the context is wrong, the adaptation will be likely be wrong too; and the context changing frequency, as if the changes occur quite often, the adaptation may not take place before another change occurs [25].

Another aspect is the time taken to adapt, since the information and/or services should be delivered in a reasonable time to the user. Based on this analysis, we defined the questions and measures presented in Table 2. We note that to calculate the measure *Adaptation Correctness* and *Context Correctness*, we should identify which adaptations the ubiquitous system proposes to do (for example, adaptation for different devices) and also which context information they use for these adaptations (for example, the screen resolution context information to adapt the application behaviour in different devices). The resulting measures will be the average from the individual adaptation and context correctness. The interpretation values are an initial proposition, based on our own experiences and [25]. They will be refined after concluding more case studies.

Table 2. Questions and Measures

Questions	Measures		
	Name	**Formula**	**Interpretation**
What is the adaptation correctness degree?	Adaptation Correctness	$\frac{\sum_{i=1}^{N}(Ai/Bi)*100}{N}$ N=Number of adaptations Ai=Number of correctly performed adaptations i Bi=Number of performed adaptations i	Low - 0 to 25% Medium - 26 to 80% High - 81 to 100%
	Context Correctness	$\frac{\sum_{j=1}^{N}(Aj/Bj)*100}{N}$ M=Number of different context information Aj=Number of corrects collected context information j Bj=Number of collected context information j	Low - 0 to 25% Medium - 26 to 80% High - 81 to 100%
	Context Frequency	F=Frequency of changing	Low-minutes Medium-seconds High-milliseconds
What is the adaptation average time?	Adaptation Time	T=The time taken to adapt	Short - milliseconds Medium - seconds High - minutes

4 Case Study

The case study was performed using a mobile and context-aware application (MCAA) called GREat Tour [26]. This application is a tour guide for a large laboratory from Federal University of Ceará. This application runs on the visitor's mobile device and provides information about the laboratory's rooms that s-he is visiting, using texts, images and videos.

To collect the measures presented in Table 2, we first identified the context information and adaptations considered by GREat Tour. It presents two adaptations (N=2) and two context informations (M=2). The first adaptation is the laboratory map view according to user's location. To this purpose, the system identifies the room through the user's mobile device that reads the QR Code installed in all the doors of the laboratory's rooms. With this input, GREat Tour updates the user's map. The second adaptation is about showing media according to the device battery level. When the battery level is low (0-9%), only text appears, when it is medium (10-20%), texts and images are displayed and when it is high (21-100%), text, images and videos are displayed.

Thus, the *Adaptation Correctness* measure takes into account the laboratory map view (i=1) and the media view (i=2) as adaptations. The *Context Correctness* measure takes into account user's location through QR Codes (j =1) and battery level (j =2) as context information. The *Context Frequency* takes into account changes in location (F) and the *Adaptation Time*, the time required to show the new map to the user (T).

The data needed to compute the measures were collected both automatically and manually. Automatic data were recorded in logs that contain: the URL of map presented according to QR Code captured, the time taken to show the map after the capture of the QR Code, the hour, minute, second and millisecond that the context was collected to calculate the context changing, the device battery level and the media presented with this value. Manual data was collected in forms filled in by evaluators that followed users during the tours, observing if the application worked correctly, in other words, identifying if the system performed a correct adaptation with correct information.

Twelve users participated in the evaluation. All of them have experience with MCAA and are from computer science domain. Their tasks were divided into three laboratory tours, with each tour consisted of three rooms to be visited. The visit consisted of updating the user map and viewing all the available informations. Each tour was done with a device in different battery charge levels for the user to experience the batteries level-based adaptation. The twelve users were equally divided in three groups to execute the test with different sequences of battery level, as presented in Table 3.

Table 3. The performed scenarios in our evaluation

		Group 1	**Group 2**	**Group 3**
Tour	**Visited Rooms**	**Battery Level**		
1	Seminars Room Library Administrative Room 1	High	Low	Medium
2	Prototyping Room Software R&D Lab 1 Meeting Room	Medium	High	Low
3	Kitchen Administrative Room 2 Research Lab	Low	Medium	High

The final result was calculated by the average of all users tours. The result is shown in Table 4. The measures about correctness had high results, only one of them was not 100% (*Adaptation Correctness*). This happened because with adaptation i=1, the wrong map was displayed. We investigated this result and identified that the instability of the wireless network was a possible cause for this adaptation problem. The *Context Frequency* measure was low, because the result was in minutes, i.e, it takes about one minute for a context change to happen. If changes occur frequently, the adaptation may not take place before another change occurs, influencing measures of adaptation correctness. The *Adaptation Time* was short (milliseconds). It is interesting to note that this result was inferior than *Context Frequency*, favouring adaptation correctness.

Table 4. Results

Measures	Results		Interpretation
AdaptationCorrectness	when i=1, 96% when i=2, 100%	98%	High Correctness
ContextCorrectness	when j=1, 100% when j=2, 100%	100%	High Correctness
ContextFrequency	00:01:37		Low Frequency
AdaptationTime	539 ms		Short Time

Based on the collected results, we can see that the high degree of correctness and low adaptation time provide to the GREat Tour application, a good HCI, regarding context-awareness measures defined by the TRUU Quality Model.

5 Conclusion and Future Work

This paper presented a model for the HCI quality evaluation in ubiquitous systems. We also presented a case study focused on the context-awareness characteristic. Currently, we are working on the execution of several case studies in order to evaluate the whole model and also on qualitative evaluations to know the user's perception about the characteristics from our model.

Acknowledgments. We thank for the technical support to this work: CTQS - Software Quality and Testing Cell of the GREat.

References

1. Rocha, L.S., Filho, J.B.F., Lima, F.F.P., Maia, M.E.F., Viana, W., Castro, M.F., Andrade, R.M.C.: Past, Present, and Future Perspectives on Ubiquitous Software Engineering. Journal of Systems and Software (2012)
2. ISO/IEC 9126. Software engineering Product quality Part 1 (2001)
3. Weiser, M.: The Computer for the 21st Century. Scientific American (1991)
4. Lee, J., Song, J., Kim, H., Choi, J., Yun, M.H.: A User-Centered Approach for Ubiquitous Service Evaluation: An Evaluation Metrics Focused on Human-System Interaction Capability. In: Lee, S., Choo, H., Ha, S., Shin, I.C., et al. (eds.) APCHI 2008. LNCS, vol. 5068, pp. 21–29. Springer, Heidelberg (2008)
5. Scholtz, J., Consolvo, S.: Toward a Framework for Evaluating Ubiquitous Computing Applications. IEEE Pervasive Computing (2004)
6. Kitchenham, K., Charters, S.: Guidelines for Performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report (2007)
7. Abowd, G.D., Dey, A.K., et al.: Towards a Better Understanding of Context and Context-Awareness. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, pp. 304–307. Springer, Heidelberg (1999)
8. Basili, V., Rombacj, H.: Goal Question Metric Paradigm. Encyclopedia of Software Engineering 2 (1994)
9. Lee, J., Yun, M.H.: Usability Assessment for Ubiquitous Services: Quantification of the Interactivity in Inter-personal Services. In: International Conference on Management of Innovation and Technology (2012)

10. Abi-Char, P.E., Mhamed, A., El-Hassan, B., Mokhtari, M.: A Flexible Privacy and Trust Based Context-aware Secure Framework. In: Lee, Y., Bien, Z.Z., Mokhtari, M., Kim, J.T., Park, M., Kim, J., Lee, H., Khalil, I. (eds.) ICOST 2010. LNCS, vol. 6159, pp. 17–23. Springer, Heidelberg (2010)
11. Ranganathan, A., Al-Muhtadi, J., et al.: Towards a Pervasive Computing Benchmark. In: International Conference on Pervasive Computing and Communications Workshops (2005)
12. Wu, C.-L., Fu, L.-C.: Design and Realization of a Framework for HumanSystem Interaction in Smart Homes. Transactions on Systems, Man, and Cybernetics: Systems and Humans (2012)
13. Kemp, E.A., Thompson, A.-J., Johnson, R.S.: Interface Evaluation for Invisibility and Ubiquity: An Example from E-learning. In: International Conference on Human-Computer Interaction: Design Centered HCI (2008)
14. Jia, L., Collins, M., Nixon, P.: Evaluating Trust-Based Access Control for Social Interaction. In: International Conference on Mobile Ubiquitous Computing, Systems, Services, and Technologies (2009)
15. Sun, T., Denko, M.K.: Performance Evaluation of Trust Management in Pervasive Computing. In: International Conference on Advanced Information Networking and Applications (2008)
16. Zhang, Y., Zhang, S., Tong, H.: Adaptive Service Delivery for Mobile Users in Ubiquitous Computing Environments. In: Ma, J., Jin, H., Yang, L.T., Tsai, J.J.-P. (eds.) UIC 2006. LNCS, vol. 4159, pp. 209–218. Springer, Heidelberg (2006)
17. Cappiello, I., Puglia, S., Vitaletti, A.: Design and Initial Evaluation of a Ubiquitous Touch-Based Remote Grocery Shopping Process. In: International Workshop on Near Field Communication (2009)
18. Iqbal, R., Sturm, J., et al.: User-centred Design and Evaluation of Ubiquitous Services. In: International Conference on Design of Communication - Documenting and Designing for Pervasive Information (2005)
19. Ross, T., Burnett, G.: Evaluating the Human-Machine Interface to Vehicle Navigation Systems as An Example of Ubiquitous Computing. International Journal of Human Computer Studies (2001)
20. Weihong-Guo, A., Blythe, et al.: Using Immersive Video to Evaluate Future Traveller Information Systems. IET Intelligent Transport Systems (2008)
21. Wagner, S., Toftegaard, T., Bertelsen, O.: Requirements for an Evaluation Infrastructure for Reliable Pervasive Healthcare Research. In: International Conference on Pervasive Computing Technologies for Healthcare (2012)
22. Kim, H.J., Choi, J.K., Ji, Y.: Usability Evaluation Framework for Ubiquitous Computing Device. In: International Conference on Convergence and Hybrid Information Technology (2008)
23. Kourouthanassis, P.E., Giaglis, G.M., Karaiskos, D.C.: Delineating the Degree of Pervasiveness in Pervasive Information Systems: An assessment framework and design implications. In: Pan-Hellenic Conference on Informatics (2008)
24. Sousa, B., Pentikousis, K., Curado, M.: UEF: Ubiquity Evaluation Framework. In: Masip-Bruin, X., Verchere, D., Tsaoussidis, V., Yannuzzi, M. (eds.) WWIC 2011. LNCS, vol. 6649, pp. 92–103. Springer, Heidelberg (2011)
25. Cheng, B.H.C., et al.: Software Engineering for Self-Adaptive Systems: A Research Roadmap. In: Cheng, B.H.C., de Lemos, R., Giese, H., Inverardi, P., Magee, J. (eds.) Software Engineering for Self-Adaptive Systems. LNCS, vol. 5525, pp. 1–26. Springer, Heidelberg (2009)
26. Marinho, F.G., Andrade, R.M.C., Viana, W., Maia, M.E.F., Rocha, L.S., et al.: MobiLine: A Nested Software Product Line for the Domain of Mobile and Context-aware Applications. Science of Computer Programming (2012)