An Evaluation Method for Context–Aware Systems in U-Health

Nayat Sanchez-Pi, Javier Carbó, and Jose Manuel Molina

Abstract. Evaluations for context-aware systems can not be conducted in the same manner evaluation is understood for other software systems where the concept of large corpus data, the establishment of ground truth and the metrics of precision and recall are used. Evaluation for changeable systems like context-aware and specially developed for AmI environments needs to be conducted to assess the impact and awareness of the users. E-Health represents a challenging domain where users(patients, patients' relatives and healthcare professionals) are very sensitive to systems' response. If system failure occurs it can conducts to a bad diagnosis or medication, or treatment. So a user-centred evaluation system is need to provide the system with users' feedback. In this paper, we present an evaluation method for context aware systems in AmI environments and specially to u-Heatlh domain.

1 Introduction

AmI environments are integrated by several autonomous computational devices of modern life ranging from consumer electronics to mobile phones. AmI has several spheres of application like: Transportation, Health, Education, Business, etc, but recently the interest in Ambient Intelligence Environments has grown considerably due to new challenges posed by society, demanding highly innovative services such as vehicular ad hoc networks (VANET), Ambient Assisted Living (AAL), e-Health, Internet of Things and Home Automation among others. These society challenges force developers to take into account growing demands of users.

Furthermore, the increase in ageing of European population and the treatment of chronic and disabled patients implies a high cost in terms of time and effort. Sometimes patients and also healthcare workers consider treatments in health

Nayat Sanchez-Pi · Javier Carbó · Jose Manuel Molina
Group of Applied Artificial Intelligence (GIAA),
Computer Science Department, Carlos III University of Madrid
e-mail: {nayat.sanchez,javier.carbo,josemanuel.molina}@uc3m.es

centres unnecessary as they could collapse national health services and increase costs. On the other hand, we face the problem of the patients living in rural areas, where is difficult to access. To face these challenges we need to differentiate medical assistance in health centres from assistance in a ubiquitous way that it is possible due to the advances in communication technologies.

Systems developed for e-Health environments need to be autonomous and self-managed. They need to adapt not only to changes in the environment, but also to the user requirements and needs. User has to take a relevant role providing an evaluation of the system behaviour while using it or at least once it has been used. One of our goals is to evaluate enhanced user experience in the course of using our system and provide automatic adaptation taking into account changes in user preferences and environment. So, generic user-centred evaluation system provides users with the possibility of having a proactive role when using the system. It is on users hands to provide the system with a feedback of the correctness of the provided services. Users will then be capable of specifying the right or wrong context information at a high-level concept so that the system could learn from it and self-adapt its behaviour for future times.

The rest of the paper is structured as followed. First section we present a general issues on evaluation and metrics for context-aware systems. Later our evaluation proposal is laid down. At the end an u-Health case study is presented and conclusions are outlined.

2   Related Work

Evaluation is a central piece of software engineering. Evaluations methodologies allow researchers to assess the quality of the findings and to identify advantages and disadvantages of a system. The goal in evaluation of conventional systems is to proof that a system is more efficient. Normally, variables associated with efficiency are the time to complete a given task or the number of errors that have been made while fulfilling the task.

However, in AmI when a system augments an environment enabling a user to do new things the metric is not straight forward anymore. So, it is important before evaluating a context-aware system to figure out what is the evaluation goal. Context-aware services must dynamically adapt to the needs of the user and to the current physical, social and task context in which those needs are formed. Developing an effective context-aware adaptive service therefore requires extensive user-centred design and evaluation as the proposed adaptive functionality for the service needs to evolve.

Since a few years, there is a growing interest in understanding specific evaluation problems that arise from context-aware systems [10, 3]. But context-aware systems are designed to provide users with services but where the main point arises in the potentiality of context. This kind of systems is designed to help users in a certain situation and provided information that is useful for a particular task [9]. Evaluation procedure becomes then a difficult problem to face when dealing with this kind of

systems. If we assume that the main purpose of context-aware system is to provide a user with information according to the contextual information, there are several problems the user currently has in accessing this information: *Distribution of the information:* concerning the manner in which the information is requested and disseminated to a user.

In AmI environments, user context and user preferences become essential aspects when deciding, which of the available services are of most interest to the user in a given situation [6]. Ubiquitous healthcare (u-health) is an emerging area of technology that uses a large number of environmental and patient sensors to monitor and improve patients' physical and mental condition. U-Health focuses on e-Health applications that can provide health care to people anywhere at any time using information and communication technologies. Besides, innovative approaches in mobile healthcare (m-Health) have also been developed as a footbridge between e-Health and u-Health.

Several initiatives, such as Mobihealth [11], XMotion [4] and MyHearth [5] have investigated the feasibility and benefits of mobile healthcare services and applications. However, these initiatives do not provide an evaluation system taking into account the contextual information as well as the user's opinion. The main contributions of this paper is to provide with an evaluation method for u-Health systems.

3   Our Evaluation Proposal

For evaluating the performance of various context-aware systems we use a three-facet approach. First, we use a taxonomy of pervasive computing systems based on our survey of proposed and prototyped systems and research projects. Second, we create a set of case scenarios which serves as a checklist of goals and functionalities for system designers to consider during both design and implementation stages. Third, we identify critical parameters for evaluating context-aware systems and a list of parameters allowing to decide what and how to evaluate each case scenario.

Taking all these into account, following we present a methodology for context-aware systems in AmI environments based on [1]. The taxonomy was constructed using a bottom-up approach and includes: architecture, modularity, geographic span, purpose and integration criteria, to define categories and key parameters to measure each one. In AmI environments the evaluation process lays the following steps:

1. **STEP 1**- Definition of the purpose of the evaluation: It is where the set of inputs is defined. This step has tremendous importance since it is here where we represent the user needs, standards and the state-of-the-art technologies. This is the reference later to match against the main objectives and intentions of the evaluation. For instance, in AmI it not the same if we evaluate the impact of the user or the performance of the system. The two approaches are lay down in next sections.
2. **STEP 2**- Design: It concerns with the set up of the evaluation plan according to the previous step. Along with the creation of the appropriate tools taking into account for instance, to perform the evaluation in an automatic, objective and respectable way.

3. **STEP 3**- Execution: Involves the measurement of the previously selected characteristics, the comparison by using the selected criteria and the assessment of the results. This step is the one that provide the feedback to the developers to serve the subsequent iteration of the design process.

Once we have decided the purpose of the evaluation, we have designed a plan of evaluation, we need to decided which metrics involve. If we would like to know the satisfaction rating of the user with a particular activity, we can measure the utility of the information delivered on attractions. So in the evaluation process we will focus on the application purpose criteria: assurance, assistive, return of investment, experience enhancement and exploration. The requirements and emphasis on various performance parameters are heavily dependent on their primary purposes. For instance in the case of return of investment, services serves primarily to increase the efficiency of users or the environments they are in, and potentially streamline the routine tasks and remove inefficiency. More prominent in business environment, like U-Commerce. Key parameters of these systems are *speed* and *efficiency*, especially of system response time and failure rate.

Another example is experience enhancement services that focus on enhancing and enriching user experiences while interacting with pervasive computing systems. They provide additional or enhanced opportunities for learning, entertainment, or sensual experience. Existing experience enhancement systems includes scenarios where personalization take place. So, usability and quality of context (QoC) represent two metrics very important for this kind of domains.

Measuring Usability

Metrics for usability are variables that are measurable in an objective manner. These variables are structured in three groups and we detail which one we use for this kind of scenario:

1. Effectiveness: Variables that allow us to measure the accuracy and completeness with which it achieves the objectives of a specific task.
2. Efficiency: Refers to the effort that a user has to do to get a goal.
3. Satisfaction: Refers to those who have more to do with the emotional or subjective.

Measuring QoC

Quality of Context (QoC) was first defined in [2] as "any information that describes the quality of information that is used as context information". While different types of contexts will have QoC attributes specific to them, there are certain attributes that will be common to most contexts and they are:

1. Precision measures how accurately the context information describes reality, e.g. location precision.
2. Probability of correctness (poc) measures the probability that a piece of context information is correct.
3. Resolution denotes the granularity of information.

4. Up-to-dateness specifies the age of context information.
5. Refresh rate is related to up-to-dateness, and describes how often it is possible or desired to receive a new measurement.

4 Case Study: Evaluation of U-Health System

In order to provide with a user-centred evaluation, we have developed a self reported data toolkit called My feedback, which will help to capture user's context, user's ratings, intentions and actions. These data can be acquired from an offline system or an online system that is installed in the users' device. Although we developed both, it is particularly useful and more accurate when collected during or shortly after key moments of interest while still fresh in the user's mind and they do not require retrieval or reconstruction data from memory but access to and accurate reporting of information available to conscious awareness. My feedback toolkit runs on Smart-Phones, PocketPCs, TabletPCs, and desktop machines running Microsoft Windows. Questionnaires are triggered based on the movements of users. To maximize user response, a numerical ratings is employed because it is much more efficient. However, this efficiency is at a cost of losing qualitative data. Thus, in this study, we also use an open question where users can recommend the correct service he would like to receive at that time and at that place. To clarify the system's functionalities, following there is a scenario where this U-health Information System can be found beneficial:

Scenario — Hospital

Hospital room, where a patient is monitored for health and security reasons. Objects in the environment are furniture, medical equipment, specific elements of the room like a toilet and a window. Users in this environment will be the patient, relatives and carers (e.g., nurses and doctors). Sensors can be movement sensors and wrist band detectors for identifying who is entering or leaving the room and who is approaching specific areas like a window or the toilet. Actuators can be microphones or speakers within the toilet to interact with the patient in an emergency. Contexts of interest can be " the patient has entered the toilet and has not returned after 20 minutes " or " frail patient left the room ". Interaction rules can consider, for example, that "if patient is leaving the room and status indicates that this is not allowed for this particular patient then nurses should be notified".

We use a multi-agent approach to implement the system published in previous work [7]. The behaviour of the system is the following: In the first phase, the Aruba Positioning system discovers the user's position(patient, doctors, nurses, medical assistant) while he enters the Wi-Fi network in the Majadahonda Hospital. Later, sensor agent provides user's positioning information to the user agent. Once user agent knows its location sends it to the facilitator agent as well as the information regarding using a specific kind of service. Following the set of phases, the facilitator agent communicates with provider agent, in this case Administration Agent

which provides a turn to the Patient's agent to see a specialist doctor that has been also previously detected. Administration agent asks the patient agent about context information (medical condition, vital signs) to be used during the interaction to provide the personalized service. Once this context information is received by the Context Manager included in the provider agent, it loads the specific context profile characteristics. This information is then consulted to personalize the provided service. Following the evaluation methodology for AmI systems and its three steps presented in the previous chapter, we define the purpose of the evaluation:

STEP 1 — Purpose

In the case u-health domain, it is a distributed system who implements, working to assist users of the hospital domain while they are carrying out their activities in different zones and with different preferences and roles. Again, taking into account the taxonomy, the purposes of evaluation in this case are:

- to measure the **application purpose** in terms of *usability*;
- to measure the **intelligence** in terms of context awareness, that means *quality of context*;

STEP 2 — Design

At the impact level, a possible metric would be the satisfaction rating of the user with a particular activity. Quality process has two distinct facets: technical quality and functional quality. Technical quality refers to the accuracy of medical diagnoses and procedures, and is generally comprehensible to the professional community, but not to patients. Patients essentially perceive functional quality as the manner in which the service is delivered; while healthcare professional can be capable of making a technical quality evaluation. There are several proposals regarding service quality measurement. Regarding this, for e-health environment we consider two groups of users: patients/caregivers/patient's relatives and health professionals. The first group will be able to make an online evaluation (OnE) of the system, for which we have defined some quality of context measurement, and the second one, an offline evaluation (OffE) with other quality of context measurement that evaluates, in this case, the technical quality of the system response.

STEP 3 — Execution

First, as the awareness of the system has been adapted for the e-health environment, the evaluation will be done based not only on the patients' location (as we did in [8], but also on his vital signs: blood pressure (BP); pulse rate (PR); respiration rate (RR) and body temperature (BT). We based on the fact that the system is composed of a set of different sensors connected to a PDA that transmits, in a secure way, all the patient data (location and vital signs) to a central server in the hospital. The authorized doctors can access this medical information from their computers (inside the hospital or even outside) afterwards. For offline evaluation, we explore the
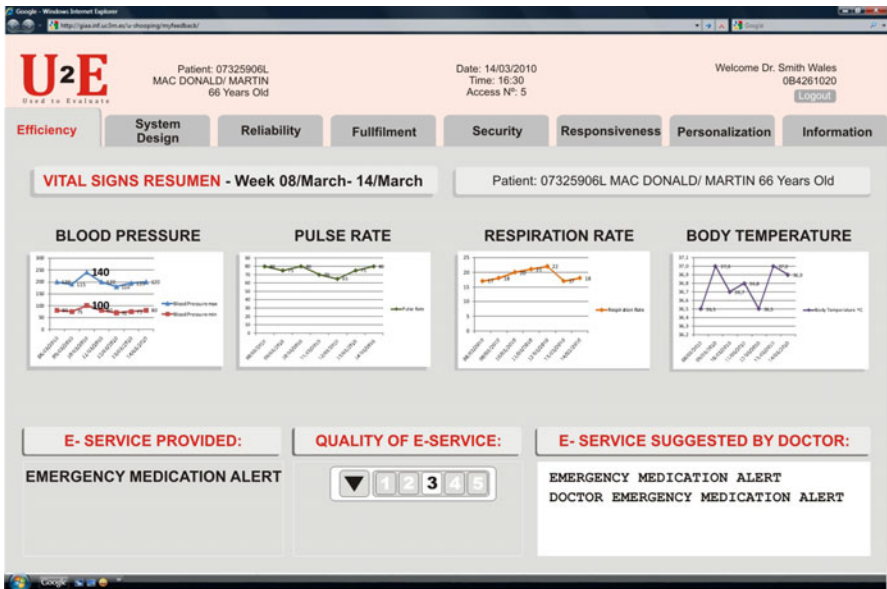
Fig. 1 Myfeedback: Offline evaluation toolkit

e-service quality dimensions based on a review of the development of e-service quality scales. So, once the e-service is provided to the patient, the OnE evaluation system is invoked by the evaluation agent and patient/caregiver/patient relatives can make the evaluation of the e-services received filling the evaluation form. Doctor or health professional in charge of following the patients' file, can also evaluate the system behaviour as see in Figure 1. In this case Dr. makes an offline evaluation of the behaviour of the system during a week. Dr. suggests the system, in a similar case, to activate the DOCTOR EMERGENCY MEDICATION'S ALERT that will send a message to the doctor, so he can be notified immediately.

5 Conclusions

Evaluations for context-aware systems can not be conducted in the same manner evaluation is understood for other software systems where the concept of large corpus data, the establishment of ground truth and the metrics of precision and recall are used. Evaluation for changeable systems like context-aware needs to be conducted to assess the impact of the users. The heterogeneity, dynamicity, and heavily context-dependent behaviors of context aware systems require new approaches of performance evaluation. Normally, apart from the simulation techniques, real-world evaluation is conducted as field studies and relies on collecting data from observation about the usability of the software in the context of use. We have applied an evaluation methodology for this kind of scenarios. The results of the application of

the evaluation method for u-Health information system show how the main characteristics of the system can be improved by taking into account users' opinion. As a future work, we intend to extend the evaluation to other application scenarios.

References

1. Abdualrazak, B., Malik, Y., Yang, H.-I.: A Taxonomy Driven Approach towards Evaluating Pervasive Computing System. 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. 32–42. Springer, Heidelberg (2010), doi:10.1007/978-3-642-13778-5_5
2. Buchholz, T., Küpper, A., Schiffers, M.: Quality of context: What it is and why we need it. In: Proceedings of the Workshop of the HP OpenView University Association, pp. 1–13 (2003)
3. Dey, A., Mankoff, J., Abowd, G., Carter, S.: Distributed mediation of ambiguous context in aware environments. In: Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology, pp. 121–130 (2002)
4. Mentrup, C.: X-motion project. Tech. rep., T-SYSTEMS NOVA Gmbh (2004), http://cordis.europa.eu/fetch?
5. Philips Research: Myheart project webpage (2010), http://www.research.philips.com/technologies/ heartcycle/myheart-gen.html/
6. Rasch, K., Li, F., Sehic, S., Ayani, R., Dustdar, S.: Context-driven personalized service discovery in pervasive environments. In: World Wide Web, pp. 1–25 (2011)
7. Sánchez-Pi, N., Molina, J.M.: Adaptation of an Evaluation System for e-Health Environments. In: Setchi, R., Jordanov, I., Howlett, R.J., Jain, L.C. (eds.) KES 2010, Part IV. LNCS, vol. 6279, pp. 357–364. Springer, Heidelberg (2010)
8. Sánchez-Pi, N., Molina, J.: A multi-agent approach for the provisioning of e-services in u-commerce environment. Internet Research 20(3) (2010) ISSN: 1066-2243
9. Schmidt, A.: Ubiquitous computing- computing in context. Ph.D. thesis (2002)
10. Scholtz, J.: Ubiquitous computing goes mobile. ACM SIGMOBILE Mobile Computing and Communications Review 5(3), 32–38 (2001)
11. Van Halteren, A., Bults, R., Wac, K., Konstantas, D., Widya, I., Dokovsky, N., Koprinkov, G., Jones, V., Herzog, R.: Mobile patient monitoring: The mobihealth system. The Journal on Information Technology in Healthcare 2(5), 365–373 (2004)