

# Care-giver Tailoring of IT-based Healthcare Services for Elderly at Home: A Field Test and its Results

Mohammad Zarifi Eslami, Alireza Zarghami, Marten van Sinderen, Roel Wieringa  
Department of Electrical Engineering Mathematics and Computer Science  
University of Twente, The Netherlands  
{m.zarifi,a.zarghami,m.j.vansinderen,r.j.Wieringa}@utwente.nl

**Abstract**—One of the biggest drivers behind IT-based homecare solutions is the increasing aging population. We are specifically interested in “service tailoring” for the homecare domain, where healthcare professionals (care-givers) do the tailoring of services to support elderly (care-receivers). Our goal is that, using our approach, care-givers can create or modify services with less IT skills, time and/or effort, and care-receivers get services that are better suited for their specific and personal needs. As a proof of concept, we developed a software prototype of our approach. The prototype was subsequently used in a real-world field test at a care institution in the Netherlands to validate the approach. The validation focused on the usability aspects of the approach in terms of effectiveness, efficiency, learnability and satisfaction. This paper describes the design of the field test and reflects on the outcome of the validation experiments.

## I. INTRODUCTION

IT has penetrated every aspect of human life in the current society and healthcare is no exception. IT-based homecare systems [1]–[7] are employed to provide care services to elderly in their home environment. Supporting independent living of elderly people through IT-based homecare services is seen as one way to deal with the consequences of an aging population, which include rising healthcare expenditures and a required shortage of healthcare professionals.

Application functionality provided to users as services are usually designed for a general purpose, user, or situation. In reality, different people have different needs. Thus, application functionality provided to users as services should (1) be aligned with the uniqueness of each user’s needs, (2) evolve with changes in these needs, and (3) take the dynamic context of the user into account. Ideally this would call for tailor-made services.

Provisioning of tailor-made turn-key services, based on dedicated design, implementation and deployment of software and hardware, is infeasible because of the cost and time involved for doing so. Instead, homecare systems should provide a set of patient-neutral healthcare-related functions which can be configured and combined according to the needs of each individual patient (*tailorability*). Previously we proposed a “service tailoring” approach [8], [9]. The service tailoring, as proposed in this paper, is a way of creating new services, and adapting previously created services, involving healthcare professionals (care-givers) in the creation process and targeting elderly people (care-receivers) as the primary users of the created services.

The top level goal of our approach is to improve homecare systems. We want to improve homecare systems by facilitating

the service creation process, in terms of reducing the IT skills, time, and effort needed by the care-givers to create new services that suit the individual needs of care-receivers. The improvement criteria can be classified as: (a) reduced or same costs of care provisioning to elderly persons, (b) better or same quality of care provided to elderly persons, and (c) improved or same quality of life experienced by elderly persons.

There are several existing works dealing with supporting independent living of elderly using IT-based systems [10]–[15]. However, little work has been done about service tailoring for homecare systems and validating such systems in a real-world settings. We proposed and prototyped a service tailoring approach for homecare. The contribution of this paper can be listed as: (a) it describes the design and execution of a field test, (b) it presents the collected and analyzed data from the field test, and (c) it reports on interesting results we obtained from the field test. The field test is designed in two series of experiments to study the usability of the approach in terms of *effectiveness*, *efficiency*, *learnability* and *satisfaction*. Because of the small size of the experiments (with a limited number of participants), the results that we obtained are primarily qualitative. However, we found them interesting, since they provided insights into the social and motivational mechanisms underlying the use of IT-based homecare services, which in turn could be used to further improve our approach.

Section II describes our proposed tailorable IT-based homecare system. Section III describes an evaluation strategy, which we followed to evaluate our approach. Section IV presents the details of the experiments’ setup and their organization under which the evaluation is carried out. Section V presents the implementation and results of the first series of experiments and Section VI presents the implementation and results of the second series of the experiments and finally, Section VII presents our conclusions.

## II. TAILORABLE IT-BASED HOMECARE SYSTEM

Tailorability of a homecare system means that a care-giver can configure the behavior of the system without help from technical personnel. The outcome of a service tailoring process is called a *service plan*, which represents a composite service tailored to the specific needs of a specific care-receiver as understood by the care-giver. A tailoring platform is responsible to enhance the creation of service plans.

As a proof of concept, we developed a prototype of our service tailoring platform, as part of the U-Care<sup>1</sup> project. Fig. 1

<sup>1</sup><http://www.utwente.nl/ewi/ucare/>

presents the U-Care system. The U-Care system comprises three main platform components which are: (a) a *tailoring platform* (to enhance creation of the service plans by care-givers), (b) a *provisioning platform* (to execute the service plans, and to integrate and orchestrate the application services as required by the service plans) [16] and c) a service repository platform (a collection of application services). Some of these application services are implemented by the U-Care system such as reminder, calendar and alert services, while others are implemented by third-party providers outside the U-Care system, such as blood pressure monitoring and medication dispensing services.

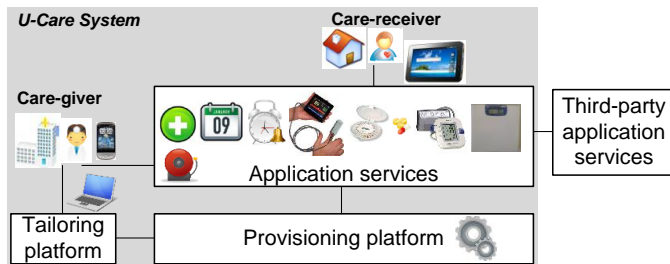


Fig. 1. U-Care system plus application providers

The care-giver drives the tailoring by making constrained decisions, based on his professional knowledge, concerning the configuration of predefined basic homecare support actions. The homecare support actions are represented as user-level service descriptions, and referred to as Service Building Blocks (SBBs). As presented in Fig. 2, Each SBB corresponds to functionality that has been implemented by a device and/or software application, and is available for use by the care-receiver. Each SBB has configuration parameters for specifying behavior constraints.

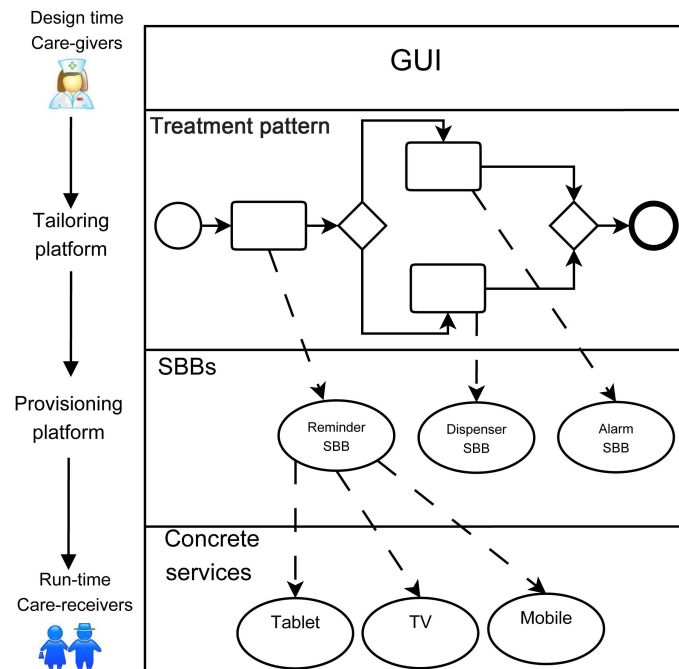


Fig. 2. Treatment patterns and Service Building Blocks

To simplify the creation of the service plan, we propose the

use of treatment patterns. A treatment pattern, which consists of references to one or a composition of several SBBs, is a pattern for a homecare task (for example, blood pressure monitoring task). A pattern can be provided by the service tailoring platform based on a selection of applicable homecare tasks from a menu with a list of common homecare tasks. To personalize the selected treatment pattern for a specific care-receiver, the care-giver provides values for the configuration parameters of the SBBs in the selected pattern. The service plan thus created, if confirmed by the care-giver, is deployed on the provisioning platform. This approach requires minimal technical knowledge and skills from the care-giver, since the SBBs hide the details of concrete implementations and the treatment patterns simplify the selection and composition of required SBBs.

### III. VALIDATION CRITERIA

The goal of the experiments of the field test was to evaluate the *usability* of the approach. Since only a few subjects participated in the experiments, we cannot generalize statistically from the results, and we regard this as an experimental case study. We will use the measurements from these experiments to identify possible improvements to our approach, but will not generalize as to the set of (all) possible applications of the current version of the approach. For this reason, we explain our observations as this allowed us to understand which parts of our approach needed improvement.

To evaluate the usability of our approach, we use the usability criteria of the *NISTIR 7432* standard [17]. It defines usability (in compliance with ISO 9241-11) as: “The extent to which a product can be used by specified users to achieve specified goals (an intended outcome) with *effectiveness*, *efficiency*, and *satisfaction* in a specified context of use”. The standard provides guidelines to measure effectiveness and efficiency, which results in objective data, and to measure satisfaction, which delivers subjective data. Furthermore, as another aspect of the usability, we measured the *learnability* of the service tailoring user interface.

#### A. Effectiveness

Common measures of effectiveness include *task completion rate*, *frequency of errors*, and *frequency of assists* to the participant from the testers. The effectiveness measurement indicates the accuracy and completeness with which users achieve specified goals (here viewed as task completion by users). It does not take into account of how the goals were achieved, only the extent to which they were achieved. Effectiveness can be scored on a scale of 0 to 100% based on specified criteria.

#### B. Efficiency

Efficiency relates the level of effectiveness achieved to the quantity of resources expended. Efficiency is mainly assessed by the mean time taken to achieve a task. It may also relate to use of other resources (e.g. total cost of usage). Task time values are useful when making comparisons between systems.

In order to measure efficiency, we measure the task completion time of end-users (care-givers) and we compare this with the task completion time of an ‘expert’ (someone who

is familiar with the system and the technology used, but not a domain expert). The comparison will give an indication of whether or not the used technology is a hindrance for using the system. This is similar to *relative user efficiency* as defined in the literature, i.e., how long a user takes in comparison with an expert [17]. However, here the values obtained from the experiments have no statistical relevance due to the small number of participants.

### C. Learnability

According to ISO/IEC 9126, learnability is the capability of a software product to enable the user to learn how to use it. Learnability is an important aspect of usability, i.e., if users cannot easily learn to use a system, for example, by following simple instructions and/or just by trying out, they will simply ignore the system. In our field test, we want to measure the degree to which the user interface of the service tailoring platform can be learned quickly and effectively. To do so, we analyze the learning time where we measure the service plan creation time by the care-givers and we observe if this time is decreased during the experiments for same or similar tasks. Second, we performed the second series of the experiments two months after the first one, and we observed if the care-givers still remember how to create service plans without our assistance.

### D. Satisfaction

Satisfaction describes a user's subjective response when using the system (expression of perceived usability). The satisfaction measurement shows if the experience was freed from discomfort and gives positive/negative evaluation of the experience of using the system. In order to measure user satisfaction, we used a *questionnaire* method. We prepared a questionnaire based on the *Computer System Usability Questionnaire* [18].

The questionnaires we prepared contain two types of questions: close-ended questions directly related to the subjects and open-ended attitudinal questions to uncover people's beliefs and thoughts on a subject. The close-ended part includes 19 questions and each question can be answered using a 7-point scales, anchored at the end points with the terms "Strongly disagree" for 1 and "Strongly agree" for 7, and a Not Applicable (N/A) point outside the scale. Thus, higher numbers are used to represent higher usability of the system.

Following the guidelines in Lewis [18], the results from user satisfaction are summarized into the following factors: perceived *overall system usability* (*OVERALL*), perceived *system usefulness* (*SYSUSE*), perceived *information quality* (*INFOQUAL*), and perceived *interface quality* (*INTERQUAL*). These factors are reported as mean values, following the guidelines in Lewis [18].

There were three questions for the open-ended part, where we ask the care-givers for the positive and negative points, as well as any suggested improvement that they see possible in the tailoring platform.

## IV. SETUP OF THE EXPERIMENTS

The U-Care system was used in a field test with two series of experiments to validate the properties of the approach and

to improve it. The field test is an action case study [19], in which we aim to improve the current situation of providing care by using a tailorable IT system. We follow the guidelines described by Wieringa in [20] to perform the experiments systematically.

Each series of experiments was conducted in a near real-world setting in a care institute in the Netherlands and each series lasted for two months. The setting of the experiments is *near* real-world, because some real-world aspects are present, such as real care-receivers, a real care institution, real nurses, and realistic scenarios, but some other aspects are absent, such as only a single homecare institution, a limited number of end-users, and no use of real medicines. In this section, we first explain which actors participate in the experiment and the role of each one, then we describe the scenarios that were used in the experiments, we explain how we collect data, and finally we describe which instruments and services were used in the experiments.

### A. Actors

Several actors have participated to cover the range of expertise required and to provide the facilities and environments needed to perform the experiments. The participants consists of one research partner (UT-CTIT), a healthcare partner, Orbis<sup>2</sup>, and three information technology partners (IBM<sup>3</sup>, MobiHealth<sup>4</sup>, and Innospense<sup>5</sup>).

UT-CTIT (the research partner), which includes three research groups, studied the goals of the project in terms of four different aspects: service tailoring, service provisioning, user interfacing, and business feasibility.

The information technology partners provided knowledge of IT solutions, application services, and industry standards. They also have experience applying such solutions in practical settings, including healthcare. For example, MobiHealth provides IT services for vital signs monitoring. Their services were used in the experiments to measure blood pressure, oxygen saturation, and weight. Innospense provides an electronic medicine dispenser device and service which were used in the experiments to guide a person in taking the proper medicine at the correct time and in the correct amount. There exist simple medicine dispensers that only function as a pill sorting box (we refer to this as a manual medicine dispenser). Also dispensers exist which can automatically dispense medication to persons and can give notifications at specified times (we refer to this as an automatic medicine dispenser). Our experiments employed both manual and automatic medicine dispensers.

Orbis, the healthcare partner, owns residential blocks where the elderly can live and receive care services provided through professional care-givers. The aim of this institution is to provide round the clock services to their care-receivers and at the same time to enable them to live an independent life as much and as long as possible. For performing the experiments, Orbis provided the application context and a test-bed for the research. They also participate in the development

---

<sup>2</sup><http://www.orbisconcern.nl/>

<sup>3</sup><http://www.ibm.com/nl/nl/>

<sup>4</sup><http://www.mobihealth.com/home/en/home.php>

<sup>5</sup><http://www.innospense.com/index.html>

of scenarios and the derivation of user requirements, as well as in evaluation of the usability of the prototype. Five care-receivers (identified as Client 1, Client 2, ..., Client 5) and three care-givers volunteered to use the U-Care system in two series of experiments. None of these care-receivers had ever used any IT-based system (e.g., computer, laptop, Tablet-PC, or smart phone) previously.

### B. Scenarios

We considered four different care services in the validation experiments: blood pressure monitoring (BP), oxygen saturation monitoring (OX), weight monitoring (WT), and medication intake support (MD). Together with the care-givers we defined the following scenarios for the experiments with Clients 1 to 5. We involved the care-givers in defining the scenarios in order to make them as realistic as possible. The motivation for using these specific scenarios in the experiments are the individual needs of Clients 1 to 5; the scenarios combine different use of the configured services to meet the needs of the care-receivers.

- *Manual MD + BP + WT*: In this scenario, a manual medicine dispenser is used; the care-receivers are asked to take their medicines (in the experiments, we used candies instead of medicines) using manual MD, and to measure their blood pressure and body weight using BP and WT, respectively, based on the plan created by the care-givers.
- *Manual MD + BP*: Same as previous scenario but only supporting medicine intake and monitoring blood pressure using MD and BP, respectively.
- *BP + WT*: Same as previous scenario but only monitoring blood pressure and body weight using MD and WT, respectively.
- *Automatic MD + OX + BP*: In this scenario, an electronic medicine dispenser is used (only for routine medication and not for medicines in hazard situations); the care-receivers are asked to take their medicines (candies instead of medicines) using Automatic MD, and to measure their oxygen saturation and blood pressure using OX and BP, respectively.

### C. Measurement Instruments

To collect data, we installed a screen capture software package, and asked the care-givers to run this software whenever they want to create a service plan. Capturing the screen helped us to record the care-givers' activities and behavior while creating service plans, so we could analyze and see in which part of the application they have difficulty understanding and performing the required actions. We also used screen capturing (with time recording) to measure the amount of time spent creating a service plan. Furthermore, after each series of the experiments, we interviewed the care-givers and care-receivers who participated in the experiment to collect their opinion about the service tailoring.

### D. Devices and Services Used in the Experiments

The care-givers received a laptop with the installed service tailoring software on it which they use to create service plans.

The created service plans are sent through the Internet to a tailoring server, and deployed and executed in the provisioning server (both servers were located at the Computer Science department of UT-CTIT). The care-givers also had a smart phone on which they received alert messages.

The care-receivers received three type of sensors to measure their own vital signs (blood pressure, oxygen saturation, and/or weight), one medicine dispenser, one smart phone, and one Tablet-PC. The measured vital signs' values are transmitted from the sensors to the smart phone via Bluetooth and then the smart phone transmits this data through the Internet to MobiHealth's local servers. The MobiHealth server pushes these values (without storing them) to the U-Care provisioning server. Finally, after successfully receiving the values, the provisioning server analyzed the values and send them to be presented in the care-receiver's Tablet-PC.

## V. FIRST SERIES OF THE EXPERIMENTS AND RESULTS

In the first series of the experiments, all five clients based on their needs and care-giver's recommendations participated in the following scenarios:

- *Client 1* participated in scenario *Manual MD + BP*
- *Client 4* participated in scenario *BP + WT*
- *Client 5* participated in scenario *Manual MD + BP + WT*
- *Client 2* participated in scenario *BP + WT*
- *Client 3* participated in *[Automatic MD] + OX + BP*

Due to the limited number of sensors, we could not test the scenarios in parallel and instead we scheduled them in three different time periods each lasting three weeks (one week of instructions and two weeks of usage). In period 1, *Client 1 & Client 4*, in period 2 *Client 5 & Client 2*, and finally in period 3, *Client 3* use the care services. We could not integrate the automatic MD with the U-Care system for the first series of the experiments, thus *Client 3* participated only in the BP and the OX tasks and we postponed the use of automatic MD to the second series of the experiments.

Before the start of the experiments, we trained the care-givers how to use the overall U-Care system including tailoring the services, measuring the vital signs, and checking the results using the care-receiver's application. The care-givers instructed the care-receivers how to measure their vital signs using the sensors and smart phones, and how to use the care-receiver's application using a Tablet-PC. It was because of privacy issues and because care-receivers trust their care-givers.

Since a limited number of care-receivers participated in the experiments, only one care-giver (we refer to this care-giver as *Care-giver 1*) was responsible for providing care activities to those care-receivers and she participated in creating the service plans. The other two care-givers only created test plans for an imaginary care-receiver.

### A. Usability Results

In the first series of the experiments, *Care-giver 1* created 11 service plans in total for 5 clients and covering four tasks: BP (5 plans), OX (1 plan), WT (3 plans), and MD (2 plans). We present the usability results of the tailoring platform based on the validation criteria which were introduced in Section III.

1) *Effectiveness*: To measure the *effectiveness*, we counted the number of the service plans created by the care-giver without our assistance. Only one service plan (for MD) out of 11 needed our assistance to be created. The care-giver thought that in order to schedule a medication task for two times per day, she should create two different service plans. We explained that the second service plan would overwrite the first plan and she could schedule the medication task for two times per a day with one service plan and specifying it in the configuration parameters. This shows that care-givers should be properly instructed concerning this point (this does not require an improvement in the system itself, but of the instructions provided to the users of the system). Note that this measurement of effectiveness is only based on the plans which were created by one care-giver.

2) *Efficiency*: To evaluate efficiency, we measured the relative user efficiency. To do so, we asked a colleague who is a partner in the U-Care project to create a service plan for the BP task (we did not repeat it for the other tasks, since creating service plans for all tasks have similar steps). This choice was made to model the role of an IT-expert as he is familiar with the underlying technology and the idea behind the services that were to be created. Then, we compared the ratio of time taken by the IT-expert with the care-givers creating the same service plans. The IT-expert created the BP service plan for an imaginary care-receiver in 4:48 minutes, while 3 care-givers participating in the experiments created the same service plan in: 3:34, 9:23, and 2:23 minutes respectively.

The care-giver, who created the service plan in 2:23 minutes, is younger and has more experience using computers than the other two. And the care-giver, who created the service plan in 3:34 minutes, had previously created 11 service plans and thus, had the most experience with creating plans among our subjects. This indicates that experience with computers and/or with the tailoring interface are two important variables to measure if we want to predict the time needed to complete a service plan. A care-giver who has general experience with using computers might create service plans faster than an IT expert with in-depth knowledge of the technology used in the system. We believe that in this case domain knowledge is more important than IT knowledge. However, a person with no computer experience at all needs more time than an IT expert (who lack domain knowledge). In this case having more domain knowledge does not compensate for the lack of IT knowledge (computer experience). Whether we can draw a conclusion from this concerning possible improvements of the system depends on the learnability of the system. If learnability (for people with no initial experience with using computers) is good, then there is no need for improvement. Our measurement of completion times of a few minutes for all plans and all nurses was encouraging, but cannot be generalized to other plans or other nurses.

3) *Learnability*: As illustrated in Fig. 3, the time required for creating service plans decreased after the first two tries by our care-givers and then stayed roughly the same. Because this kind of interface was familiar to the care-givers, we think this is learning behavior would happen with other care-givers too.

*Service plans 5-9* were created two weeks after creating *service plans 1-4*, and *service plans 10-11* were created two weeks after creating *service plans 5-9*. Nevertheless, looking

at Fig. 3, it is evident that the care-giver could remember the steps required in creating service plans even after two weeks, hence later plans took less time than the initial ones. Another interesting observation is that even though there was only one plan for the oxygen saturation task, *service plan 10* was created in a relatively short time due to the similarity of steps required to create service plans for the different tasks.

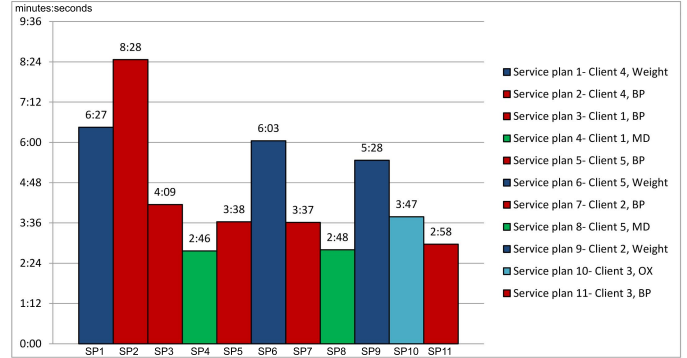


Fig. 3. Created service plans in the first series of the experiments

The service plans for the weight monitoring task took on average more time than other plans. This is because the care-giver created the plans for the weight monitoring task for Clients 5 & 2 (i.e., *service plans 6 & 9*) after creating the plans for the blood pressure monitoring task (i.e., *service plans 5 & 7*), and hence, she took some time to decide about the time of weight measurement so that it does not conflict with the blood pressure measurement time. This shows that the tailoring platform should provide some abstract information to a care-giver about previously created plans for a care-receiver while s(he) creates a new service plan, such that interdependencies can be tackled during the plan creation process. Another possible improvement is that a service plan could cover more than one task, allowing better handling of interdependencies between tasks.

4) *Satisfaction*: After the execution of the first series of the experiments, we interviewed the care-givers and asked their opinions about the U-Care system in general and the service tailoring platform in particular. To measure *satisfaction*, we asked the care-givers to fill in the questionnaires, which were introduced in Section III. They filled in these questionnaires two times: once for the whole U-Care system and second time for the service tailoring. Table I summarizes the results for the System Usefulness (SYSUSE), Information Quality (INFOQUAL), and Interface Quality (INTERQUAL) (for the close-end questions). The presented values are the mean-values calculated from the values provided by the three care-givers.

Since the care-givers used the service tailoring platform themselves, the numbers reflect their opinion about the platform. While judging the whole system (care-receiver's application, third party services, and provisioning platform), they provided their answers from two perspectives: first, the amount of time and effort they needed to (re)train and assist the care-receivers in using the system, and second, the care-receiver's opinion about the system as perceived by the care-givers.

Looking at the Table I, we can see that the care-givers were more satisfied with the tailoring platform than with the whole system. The care-givers only interacted directly with

the service tailoring platform. It had an interface of the kind they are familiar with in other hospital applications, and they felt comfortable with it. Perhaps when considering the whole system, they also assessed the interaction between the system and the care-receivers. Moreover, looking at the satisfaction factors of the tailoring platform, we can see that *information quality* was judged low (5.17), because they could not get a list of service plans and parameters that they had already entered. This relates to what has been said under learnability.

TABLE I. SATISFACTION RESULTS FOR THE FIRST SERIES

Score Name (1-7)	Whole System	Tailoring platform
OVERALL	4.75	5.67
SYSUSE	4.37	6.00
INFOQUAL	4.93	5.17
INTERQUAL	5.32	5.67

Through subjective opinions from the second part of the questionnaire, we observed that the care-givers found the tailoring interface and the process easy and straight forward and it did not take too much of their time to create a service plan. They mention that “We create a plan in less than 5 minutes and it was enough for the whole period”. Moreover, they found the re-tailoring of previously created plans quite easy and quick. The re-tailoring was required if case they monitored the behavior of the care-receivers and wanted to change some part of the created plan based on this experience.

The care-givers found the user interface of elderly person’s application not mature enough to be used by the care-receivers. For example, the care-givers say that “Client 4 has fear of using the system alone”. This took the care-givers extra time because they were frequently called to assist Client 4 in using the system, thus defeating the time-saving purpose of introducing the system. None of the care-receivers who participated in the experiments had ever used any IT-based system previously, and in a similar situation we expect similar extra calls to care-givers to be made. Designing a suitable user interface for the elderly remains a challenging task.

The care-receivers themselves, on average, have more positive opinions about the U-Care system. Client 1 (73 years old) and Client 5 (93 years old) were enthusiastic users of the U-Care services without the care-givers’ help. But not all clients were satisfied. Client 4 (98 years old) found it difficult to use the system: “Too many steps (button pressings), not enough loud reminder sound to notify me, etc. I lived 98 years without using this types of services, I could survive the rest of my life without them”. However, she liked using the Skype service to communicate with her family members (who filled in the contact lists), friends, and care-givers. She also liked reading E-books. In fact, she ignored reminders from the system, because she wanted to continue reading an E-book.

### B. Requested Changes

The care-givers requested a number of changes be implemented in the tailoring platform. These changes were expected to be implemented before the second series of the experiments.

- 1) They asked to add a feature that would give the possibility of scheduling the tasks to be executed at

different times of a day. In the current prototype, they could add multiple times, but only by indicating time intervals between two events. For example, they could specify in the blood pressure service plan to create an event twice per day, starting from 9:00 in the morning and repeating it after 8 hours. However, they would like to specifically indicate the time of two or more (up to four) events per day.

- 2) For the weight monitoring task, the initial treatment pattern was to compare the measured value from today with that from yesterday, and if the difference exceeds a predefined threshold, the system should raise an alert. However, the care-givers mentioned that the measured values must be compared with a reference point (e.g., 85 Kg), which could be configured for each individual. They also preferred to have the tailoring interface in local language, which was Dutch in our experiment/validation environment.
- 3) As mentioned before, we captured the screen of the care-givers’ laptop to monitor their behaviour while creating service plans. Based on the analysis of these screen shots, we identified another possible change. We saw that most of the time, which is consumed by a care-giver to create a plan, was spent checking other created plans for that care-receiver. This can be eliminated by providing a list of already created service plans (for a care-receiver) and a summary of their parameters to a care-giver, while (s)he is creating a new service plan.

The care-givers also requested some changes regarding the user interface of the applications by the elderly. The improved version should have fewer buttons and options and require a minimum number of interactions with the care-receivers.

## VI. SECOND SERIES OF THE EXPERIMENTS AND RESULTS

Two months after the end of the first series of the experiments, the improved U-Care system based on the requested changes from the care-givers and care-receivers was validated in a second series of experiments. The same clients were supposed to participate in the same scenarios.

Unfortunately between the first and second series of the experiments, *Client 5*, who was very enthusiastic about using the system and appreciated it, had passed away. *Client 4* stopped using the system because of a lack of interest. Client 3 initially was willing to use the system for the second experiment, however, she lost her interest because of receiving bad news (she was diagnosed with cancer). For the second series of the experiments, Orbis introduced a new volunteer client (Client 6) to use the U-Care system. Thus, only 3 clients in total participated in the following scenarios:

- *Client 1* participated in scenario *Manual MD + BP*.
- *Client 2* participated in scenario *BP + WT*.
- *Client 6* participated in scenario *BP + WT*.
- *Client 1* participated in scenario *Automatic MD*.

As in the case of the first series of the experiments, due to the limited number of sensors and smart phones, in the second series of the experiments the scenarios are also

executed sequentially in three periods: in period 1, *Client 1*; in period 2, *Client 2 & Client 6*; and finally in period 3, *Client 1* (for the second time in the second series of the experiments) used the care services. Before the start of these experiments, we informed the care-givers about the refinements made in different parts of the system based on their feedback.

#### A. Usability results

In the second series of the experiments, the care-givers created seven service plans in total for 3 clients and for three tasks: BP (3 plans), WT (2 plans), and MD (2 plans). The creation of service plans for automatic MD and manual MD have same number of steps and therefore, the care-givers saw no difference in creating their plans. This is one of the benefits of having a service tailoring platform and using the concept of SBBs for the different concrete services provided by various providers. This way, the care-givers did not have to learn different configuration options, user interfaces of different vendors, but only used the tailoring interface that we proposed.

Out of the seven service plans, five were basically re-tailoring of the service plans previously created (for *Client 1 & 2*) during the first series of the experiments. Creating these five service plans was easy and each took less than half a minute to create. The care-givers only created 2 service plans for *Client 6* from scratch, of which the service plan for *BP & WT* were created in 2:38 and 2:43 minutes respectively.

Since only 2 service plans were created from scratch, measuring the *effectiveness* and the *efficiency* is not that meaningful. However, we could observe the *learnability* of the tailoring platform and its user interface. The service creation time required by the care-givers in the second series of the experiments was compared to that in the first series of the experiments (see Fig. 3). This comparison suggests that the learning curve is shorter, i.e., once trained, the care-givers not only could remember how to create the service plans for the second series of the experiments themselves and did so without expert assistance, but they also created the service plans faster.

After the completion of the second series of the experiments, we interviewed the care-givers to measure their *satisfaction*. In this final interview, we asked the care-givers to fill in a questionnaire. But unlike in the first series of the experiments, this time we used a different questionnaire which had only open-ended questions and had no close-ended questions. The reason behind this was that after analyzing the results of the first series of experiments, we noticed that the open-ended questionnaires provide more information about the usability of the system than the close-ended questions.

The new questionnaire has 15 questions. Q1-Q3 evaluate the functionality of the system, Q4-Q13 evaluate quality aspects, and Q14-Q15 ask whether the care-giver is willing to use such a system in practice or not. Filling in the questionnaires took almost an hour, and then after a short break, we discussed the questions and answers all together. The results of these interviews are listed below:

- **Usability of the Service Tailoring Platform:** The care-givers found the service tailoring platform as usable the second time as they did the first time. They could create service plans in about 3 minutes, and they did not feel a need for reducing this further.

- **Usability of GUI of Applications for Elderly:** The care-givers still found the care-receiver GUI not sufficiently usable for the elderly. They stated that there should be fewer options and buttons than the GUI that we implemented, and they expressed a preference for a voice interface to the elderly. Not only the software, but also hardware should be designed specifically for elderly. Current Tablet-PCs technology was not considered convenient for elderly who are not familiar with concepts such as scrolling, tapping, or sweeping. The care-givers mentioned the automatic medicine dispenser (MD) as an example of a usable technology for the elderly in compare to the manual MD, where in order to get the medicine from the automatic MD, the care-receivers did not have to press any button on the Tablet-PC, but instead after receiving a reminder message, just pressed a big button on the dispenser and they could confirm receiving the medicine.
- **Saving Time:** The care-givers mentioned that if an IT-based system works correctly, indeed it could save their time, since they do not have to measure the vital signs of care-receivers or dispense their medication in person. However, if elderly need help of care-givers to operate the system, then systems like U-Care would create more work for care-givers rather than less work.
- **Quality of Care:** The care-givers believed an IT-based homecare system could increase and at the same time decrease the quality of care. It could increase the quality of care, when a care institution has too many clients or when the elderly live in their own home and receive care services at home. In those situations, using the IT-based systems could save care-givers' time and increase the quality of care by providing services 24/7. On the other hand, for a care institution with a limited number of elderly, the quality of care could be decreased, because of *less attention* and *less physical contact* between care-givers and care-receivers. The physical contact is reassuring for elderly and reduces stress. Due to this less attention, the elderly may feel uncertain and keep measuring their vital signs repeatedly, when the values are too high/low. Furthermore, when there is a hazardous situation, the elderly do not get immediate help if there is a need. The care-givers believed that using a video communication service together with other care services could compensate for these negative effects of less attention and less physical contact.
- **Quality of Life:** The care-givers also believed that an IT-based system could increase as well as decrease the quality of life of the elderly. For some elderly, it gives *a sense of independence*, since they could measure their own vital signs without the help of care-givers. For other elderly persons, it has negative effects as the system restricts their behavior in the sense that they are afraid to leave their rooms because of the fear that they might not be able to measure their vital signs at the scheduled time. The care-givers indicated that the system is mobile (so mobility is an important factor) and the care-receivers could take the

system with them for example, to the restaurants and measure their vital signs there. However, some elderly persons were scared that the devices fall and break (so another important factor is solidity and durability of the devices and using fewer devices, i.e., only a Tablet-PC or a smart phone and not the both).

- **Target Group:** The care-givers believed an IT-based system is more useful in situations where the elderly still live in their own home rather than in a care institute. The care-givers indicated that they would prefer to test the system with the care-receivers in the range of 60-80 and for a longer period of time.
- **Integrated with their Current IT Systems:** Another interesting desire was that the system be integrated with their current IT system, such as their electronic patient file system, this would be more useful and have greater value than a separate stand alone system. This is because vital signs and other elderly data could be directly stored into the electronic file system without intervention from nurses, relieving them of some administrative tasks.

## VII. CONCLUSIONS

In this paper, we discussed a field test of our proposed service tailoring approach for IT-based homecare systems. The approach proposed had the following goals: (a) reduced or same costs of care provisioning to the elderly, which can be achieved by saving the time of the care-givers, (b) better or same quality of care provided to the elderly, which can be achieved by personalizing the services, and (c) improved or same quality of life experienced by the elderly, which can be achieved by supporting independent living of elderly. At least one of the above items should be better and none should be worse. This is subject to discussion. We believe, a hard criterion is costs: “care provisioning should not become more expensive”. But the other criteria might be violated to some extent (but preferably not) if the cost savings are substantial. Also the criteria are relative to the current situation. But the current situation will not last. The aging population may force changes: lack of personnel leads to less time spent per care task per elderly person, and thus, the quality of care or quality of life will decrease, unless this can be prevented by new solutions, such as tailorable IT-based services.

We performed two series of the experiments in the field test in which several care-givers and care-receivers volunteered to use the system. We defined evaluation criteria to analyze the usability of the approach. Our conclusion from the field test is that the U-Care system in general and the tailoring platform in particular is usable (by care-givers and care-receivers) at least in our field test. If the amount of time required to train the care-receivers to use the system is reduced, using the U-Care system generally could save their time and effort. We did a qualitative research on the real-world field test, and our conclusions are only propositions, where a quantitative research should be then complemented it to find empirical support for our hypotheses.

## ACKNOWLEDGMENT

This work is part of the IOP GenCom U-Care project (<http://www.utwente.nl/ewi/ucare/>), sponsored by the Dutch Ministry of Economic Affairs under contract IGC0816.

## REFERENCES

- [1] C. Müller, C. Neufeldt, D. Randall, and V. Wulf, “ICT-development in residential care settings: sensitizing design to the life circumstances of the residents of a care home,” in *Conference on Human Factors in Computing Systems*, ser. CHI '12. ACM, 2012, pp. 2639–2648.
- [2] M. Batet, D. Isern, L. Marin, S. Martinez, A. Moreno, D. Sanchez, A. Valls, and K. Gibert, “Knowledge-driven delivery of home care services,” *Journal of Intelligent Information Systems*, pp. 1–36.
- [3] I. Korhonen, J. Parkka, and M. Van Gils, “Health monitoring in the home of the future,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 22, no. 3, pp. 66 – 73, May-June 2003.
- [4] S. Zhang, S. I. McClean, B. W. Scotney, X. Hong, C. D. Nugent, and M. D. Mulvenna, “Decision Support for Alzheimer’s Patients in Smart Homes,” in *CBMS*, 2008, pp. 236–241.
- [5] N. Malanowski, R. Ozcivelek, and M. Cabrera, “Active Ageing and Independent Living Services, The Role of Information. and Communication Technology,” European Community, 2008.
- [6] K. Gaßner and M. Conrad, “ICT enabled independent living for elderly, A status-quo analysis on products and the research landscape in the field of Ambient Assisted Living in EU-27,” prepared by VDI/VDE Innovation und Technik GmbH, March 2010.
- [7] European Commission, “Ageing well in the information society - an i2010 initiative - action plan on information and communication technologies and ageing,” EU, Tech. Rep., Jun. 2007.
- [8] M. Zarifi Eslami, A. Zarghami, B. Sapkota, and M. van Sinderen, “Personalized Service Creation by Non-technical Users in the Homecare Domain,” *Procedia CS*, vol. 5, pp. 409–417, 2011.
- [9] M. Zarifi Eslami, A. Zarghami, B. Sapkota, and M. van Sinderen, “Flexible Homecare Application Personalization and Integration Using Pattern-Based Service Tailoring: Supporting Independent Living of Elderly with IT,” in *11th IEEE International Conference on Computer and Information Technology (CIT)*, 2011, pp. 467–474.
- [10] CLEAR, “The Clinical Leading Environment for the Assessment of Rehabilitation protocols in home care,” available at: <http://www.habiliseurope.eu/?q=node/5>, last visited: Feb. 2013.
- [11] T-SENIORITY, “Expanding the benefits of Information Society to Older People through digital TV channels,” available at: <http://tseiority.idieikon.com/index.php/lang-en/the-project>, last visited: Feb. 2013.
- [12] Dreaming, “eIDeRly-friEndly Alarm handling and MonitorING,” available at: <http://www.dreaming-project.org/>, last visited: Feb. 2013.
- [13] MATCH, “Mobilising Advanced Technologies for Care at Home,” available at: <http://www.match-project.org.uk/main/main.html>, last visited: Feb. 2013.
- [14] MPOWER, “Middleware Platform for eMPowering cognitive disabled and elderly,” available at: <http://www.sintef.no/Projectweb/MPOWER>, last visited: Feb. 2013.
- [15] Amigo, “Ambient intelligence for the networked home environment,” available at: <http://www.hitech-projects.com/euprojects/amigo/>, last visited: Feb. 2013.
- [16] A. Zarghami, M. Zarifi Eslami, B. Sapkota, and M. van Sinderen, “Dynamic Homecare Service Provisioning Architecture,” in *9th International Conference on Service-Oriented Computing and Applications (SOCA)*. IEEE, 2011, pp. 1–8.
- [17] National Institute of Standards and Technology, “Common Industry Specification for Usability - Requirements,” June 2007, available at: <http://zing.ncsl.nist.gov/iusr/documents/CISU-R-IR7432.pdf>, last visited: Feb. 2013.
- [18] J. R. Lewis, “IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use,” *International Journal of Human-Computer Interaction*, vol. 7, no. 1, pp. 57–78, 1995.
- [19] R. Wieringa, “Designing Technical Action Research and Generalizing from Real-World Cases,” in *Advanced Information Systems Engineering*, ser. LNCS, 2012, vol. 7328, pp. 697–698.
- [20] R. Wieringa, “A Unified Checklist for Observational and Experimental Research in Software Engineering (Version 1),” <http://eprints.eemcs.utwente.nl/21630/>, Enschede, Technical Report TR-CTIT-12-07, March 2012.