MENTAL MODELS FOR PERSONALISATION OF TELCO SERVICES

SEBASTIAN MÖLLER AND MARTIN KURZE
Deutsche Telekom Laboratories
Ernst-Reuter-Platz 7
10587 Berlin, Germany
+49 30 8353-58465 (SM)
+49.30 34 97-25 96 (MK)
{sebastian.moeller@telekom.de, martin.kurze@telekom.de}

Abstract. This paper describes a novel view to the area of multimodal interfaces from the perspective of personalization. At the same time the creation and utilization of mental models for personalization is described. The whole area of cross-application multimodal personalization is discussed and practical as well as conceptual (theoretical) work is presented.

1. Introduction

What makes human factors in telecommunications so special, given that human factors and man-machine-interaction studies are a wide field of research which is covered by a large number of excellent researchers?

The answer is closely connected to the type of business, Telco companies do and the unique selling position and interests they have. Two main factors are worth mentioning:

1. Telco services are multimodal in many cases. At least the sum of all offered services deals with several modalities such as voice (e.g. in telephony), touch (e.g. touchscreens) and graphics (e.g. web based applications) (Kurze, 2007).
2. Telcos as companies and the services they offer may build on top of a huge database of user (customer) habits: Telcos know a lot about their users and they usually have a large user basis.

These preconditions make it extremely interesting to offer personalized services, or to offer personalization as a service for 3rd parties. Especially in the somewhat difficult field of personalization for multimodal applications, Telcos might want to utilize as much information about the user as possible. Unfortunately, the sheer mass of data makes proper personalization very difficult. And the challenge of multimodality adds further complexity to the task.

In this paper, we describe a novel approach to predict user interaction behaviour following an assumed 'mental model' of the user. On the basis of predicted user behavior, personal needs and preferences can be deducted and utilized in a complex personalization framework.

The paper goes into details of the conceptual work which is now being implemented at Deutsche Telekom Laboratories and briefly describes two independent projects which feed into the approach:

- One project on personalization deals with the complexity of application-overarching personalization and presents a framework which offers the necessary tools and methods to personalize application interfaces and content, in different modalities.
- The other project develops a workbench for automatic user testing based on mental models. We concentrate on the user model built during the setup of the workbench, which mainly consists of a formal representation of a user's assumed mental model.

Both projects are running right now and already produced a number of valuable results (in their particular context).

The combination of the core components of these two projects (mental models for user modeling and a personalization framework for multimodal cross-application adaptations) enables companies such as Telcos to gain maximum user acceptance and optimal usability, making use of the multimodal nature of the relationship between a customer (user) and his or her Telco service provider. So far, this is more a conceptual work showing how a future exploitation could be set up. An implementation of the feature is not yet complete. Conclusions are drawn and future work is proposed focusing mainly on a large scale implementation of a combined utilization system and the respective architecture. Possible effects of standardization in the area of personalization (and mental models) are discussed.

2. Personalization as a requirement and opportunity for Telcos

As mentioned above, Telcos already know a lot about their customers. But are all their users customers? This is one of the key questions for an effective personalization. And this question has two possible denotations:

1. Aren't there users who use the network and/or offered higher level services without paying any money to the Telco company? These users form actually an extremely interesting group since knowledge about their needs and habits could help to convert them from "users" to "paying customers" (where "paying" might have several connotations, see below).

2. Is the role of a "customer" really covering all relevant aspects of the person using a Telco network/service? As a customer, the person books certain services or products from the Telco provider and pays for it while the Telco provides the services/products and ensures a certain level of QoS (quality of service) and prevents the customer from using more than he paid for.

Both questions suggest that it might be wise to collect actual usage information from the individual user and utilize this collection for the sake of the user and the (Telco) company: The more the Telco/service provider knows about the user and his habits, the more focused offers can be made. These offers can either be enhanced user interfaces and pre-selected contents adapted to the individual user, or new product offers such as new tariff or a new service product.

These considerations make it obvious that individual sensed interactions are needed to identify individual needs (Brusilovsky, Kobsa and Nejdl, 2007). On the other side, individual data and generalized versions of the datasets might be useful to enhance the service quality as a whole, including – but not limited to – usability and interface questions. They will very likely feed into an abstract user model which also considers the results from the mental model approach described in the subsequent chapter.

Sensing and collecting user interactions (and optionally other data such as ratings etc.) form one aspect of personalization. The other task is to provide the actual benefit for the user, i.e. a better, more personalized service. To be able to deliver this enhanced service, a smart mechanism is necessary taking into consideration all known and relevant information about the user, the task, the interface, and the content of interest. And finally, a mechanism is needed to alter the user interface, the presented content, or other impressions the user might experience. For this final step, the alteration of the application, we propose "activators", plugged in into the applications which finally perform the alteration tasks and change the user's perception.

Figure 1 shows the complete system architecture: Four applications are currently using the approach and implement sensors and actuators for the technical interfacing to the personalization framework. This personalization framework uses several internal and external components to build up an individual user model with all available profiling information, and derives recommendations from the database. It is important to note that these "recommendations" are not presented to the user (directly) but rather are forwarded to the actuators and the applications, which in turn can then react to the recommendation according to the application designer's decision.

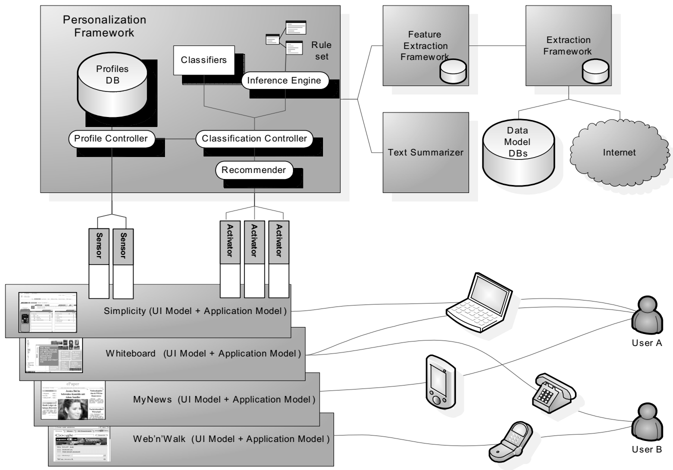

Figure 1. Architecture of the personalization framework.

This architecture offers a number of benefits for affected parties: The framework can operate on a quite “abstract” level; the application can keep its “look and feel”; and the user can (to a certain degree) influence the style and severity of the alterations.

One further main benefit of the overall architecture is the “application-overarching” personalization capability: Since the personalization framework collects data from several applications and numerous users, users’ preferences in a specific context can be transferred to other applications used by the same person, and/or these preferences can be offered to other users of the same application with a similar profile. This of course raises certain privacy issues (Kobsa and Chellappa, 2006). Finally the large amount of collected data can feed an even more ambitious approach: Evaluating and enhancing user interfaces during the interface design phase based on some sort of ‘mental model’. This approach, the core of the second project, will be described now.

3. User behaviour prediction based on mental models

This second project is related to the link between user characteristics and expected user behaviour. Its aim is to simulate user behaviour with initial prototypes of interactive systems (graphical user interfaces (GUIs), spoken dialogue systems, multimodal systems) in order to analyze and optimize the usability of the systems already in the system design phase. In this way, user characteristics collected in the personalization framework described above can be taken into account, making the prediction also a helpful tool for personalization.

The original idea of the simulation is based on the concept of so-called ‘mental models’. This internal representation of the system and its running logic are used to describe, explain and predict the behaviour of the system (Gentner and Stevens, 1983). It has to be noted that a mental model is usually acquired on-the-fly, and that it is neither stable nor necessarily ‘correct’; the user may have ideas of what to do and how to do it, whereas the system is not necessarily capable of accomplishing that. Discrepancies between the user’s ‘mental model’ and the implementation of the system affect the smooth course of the interaction and may lead to interaction failures (e.g. Möller et. al., 2007). Mental models cannot be simulated trivially with stable conceptual models, but updating and learning through interaction needs to be accounted for.

Because of these limitations, we adapted the concept of ‘mental models’ to come to executable simulations of user behavior. We finally ended up with four models which describe the actions of the user (e.g. clicking a button, writing in a text field, moving the mouse, utter a sentence, etc.) as well as the ones of the system (e.g. display a GUI, highlight an element in the GUI, utter a spoken prompt, etc.) during the interaction: the user task model, the system task model, the system interaction model, and the user interaction model. These models are used to simulate interactions between the system and one or several potential users in the so-called “MeMo workbench” described in detail in Möller et. al. (2006), see Figure 2.

User and system models provide input to an automatic testing unit – the core of the system. This unit hosts user and system models in the automatic testing cycle; it simulates interactions and generates loggings. The log-protocols serve the automatic detection of estimated (and potentially erroneous) user behavior, and include discrepancies between user and system models, resulting in problematic interactions. Using appropriate weighting of the discrepancies, a prediction algorithm derives a usability profile from the protocols. This usability profile describes the impact of errors in terms of performance and quality indicators, related to effectiveness, efficiency, and user satisfaction.

Four types of models are used by the MeMo approach. The system task model describes the tasks a user may perform with the help of the system. Despite their limitations, most state-of-the-art systems use an attribute-value description (slots) for internal task representation. Such a representation seems to be preferable also for the description of the system task model. The description may be directly derived from the system specification, using e.g. a graphical tool.

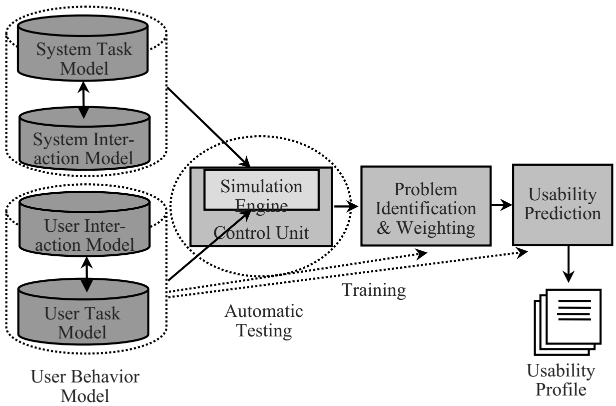

Figure 2. MeMo workbench for simulating user-system interactions.

The system interaction model describes the flow of interaction which is coded in the system. We make use of a state-chart for this purpose. Like the system task model, this model can directly be derived from a formal description of the system.

The user task model is a collection of potential user tasks with the system. The formal description is identical to the one of the system task model. Its derivation may be based on a general domain model (i.e. a model of objects and tasks conceivable for a specific domain, derived e.g. with the help of user studies or reviews of existing systems), on user studies carried out with a prototype version of the system, or simply on the system task model. The user task model reflects the frequency with which tasks are likely to be carried out. Such probabilities may be derived e.g. from user studies or marketing information (in case that the service is not yet available), or from the personalization framework (if a first version of the service, or alternative services for similar tasks, are already available). Both frequent and infrequent – but potentially error-prone – tasks should be modeled.

The user interaction model describes potential behaviour of the user when being confronted with the system. In general, user input will strongly depend on the system output in the previous turn. Thus, it is conceivable to base this model on the system interaction model: the latter determines a set of "correct" user inputs for each step of the dialogue, in terms of an intended path for each task to be solved. From the intended path, deviations are generated in a probabilistic rule-based manner. For example, there are certain probabilities that the user will deviate from the intended path, e.g. because s/he does not understand a system prompt or a button label, because s/he is irritated by several apparently similar options, or simply because s/he is lost in the course of the interaction. Each deviation from the intended path is linked to a probability.

In turn, the probabilities may be linked to user characteristics. For example, a user with poor command of the English language may show a higher probability to not click on a button with an English label, or may tend to miss-interpret English abbreviations. A user with experience in a particular technology (e.g. the user of a particular mobile phone) may try to stick to the interaction behaviour associated with this technology; this may increase the probability to deviate from the intended path in cases where the path is contrary to the interaction logic of the mobile phone the user is used to. The user characteristics may be combined to form certain classes of prototypical users which are expected to be homogeneous in their interaction behavior, cf. (Naumann et. al., 2008).

The user of the MeMo workbench – i.e. the system designer – may select one or several classes of users he would like to simulate interactions for. The simulation engine then runs a pre-defined number of (simulated) interactions between user and system models. The control unit generates logs of the simulated interactions as an output. The logs are annotated by information from all models (system task and interaction models, user task and interaction models), providing a priori information for the following discrepancy detection and weighting modules. The interactions may also be annotated for the rules which created deviations from the intended path; these rules may later assist the system developer in understanding why deviations occurred, and may be used as a basis for optimization.

The final usability prediction module provides a link between the instrumental annotation of the dialogue on the one hand, and the user opinion about quality and usability on the other. In this way, a usability profile is established which lists the simulated user and system behavior, the frequencies and types of interaction errors, as well as indices related to the effectiveness, efficiency and expected user satisfaction of the simulated interactions. These profiles can be derived for specific sets of tasks, users, and system variants. They serve a summative, diagnostic as well as comparative evaluation of the system under test.

4. Conclusion and Next Steps

The approaches described above show that complementary ways to advance user personalization and optimizing usability have an enormous potential for synergy. On the one hand, the personalization framework is able to provide user-specific information on his/her behavior and preferences across a number of multimodal applications. Making use of this information, relevant information can be provided to the user in the way he/she prefers it or is used to it. On the other hand, the information collected within the framework may be used for the early optimization of updated and future services. Using implemented versions of 'mental models', user behavior towards such applications can be anticipated and usability can be optimized at an early stage of service development, and at low costs.

It is obvious that the implementation of the personalization framework as well as the MeMo workbench remain challenging. Nevertheless, we are convinced that the personalization framework will contribute valuable information about user behavior, while the resulting implicit mental model has the potential to ease interface development and thus enables much better "first shots" that cause far less problems to the first time user. By doing this, the expected results carry an enormous value for the Telco companies in their struggle for future markets, and to their users and customers as well.

Acknowledgements

The authors would like to thank the teams of the PIUI and MeMo projects at the Quality and Usability Lab and DAI-Labor, TU Berlin, as well as at DFKI, Saarbrücken/Berlin, and Fraunhofer FIT, Birlinghofen, for their excellent conceptual and technical work that made these results possible.

References

Brusilovsky, P., Kobsa, A, and Nejdl W. (eds.): 2007, *The Adaptive Web: Methods and Strategies of Web Personalization*, Springer, Berlin Heidelberg New York.

Gentner, D. and Stevens, A. L. (eds.): 1983, *Mental Models*, Lawrence Erlbaum Associates, Hillsdale NJ.

Kobsa, A. and Chellappa, R. (eds): 2006, Privacy Enhanced Personalization, in: *Proc. of the PEP Workshop*, Montreal, Canada, April 22nd.

Kurze, M.: 2007, Personalization in Multimodal Interfaces, in: *Proc. ICMI '07 Workshop on Tagging, Mining and Retrieval of Human-Related Activity Information*, Nagoya, Japan November 15.

Möller, S., Englert, R., Engelbrecht, K., Hafner, V., Jameson, A., Oulasvirta, A., Raake, A., and Reithinger, N.: 2006, MeMo: Towards Automatic Usability Evaluation of Spoken Dialogue Services by User Error Simulations, in: *Proc. 9th Int. Conf. on Spoken Language Processing (Interspeech 2006 – ICSLP)*, Pittsburgh PA, pp. 1786-1789.

Möller, S., Engelbrecht, K.-P., and Oulasvirta, A: 2007, Analysis of Communication Failures for Spoken Dialogue Systems, in: *Proc. 8th Ann. Conf. of the International Speech Communication Association (Interspeech 2007)*, Antwerp, pp. 134-137.

Naumann, A., Hermann, F., Niedermann, I., Peissner, M. and Henke, K.: 2008, Interindividuelle Unterschiede in der Interaktion mit Informations- und Kommunikationstechnologie, to appear in: *Mensch und Computer 2007*, Oldenbourg Wissenschaftsverlag, München,.