User Profile Matching: A Statistical Approach

Work in Progress Paper

Benjamin Wassermann, Gottfried Zimmermann
Wilhelm Schickard Institute of Computer Science
Eberhard Karls University of Tübingen
Germany
benjamin.wassermann@uni-tuebingen.de
gottfried.zimmermann@uni-tuebingen.de

***Abstract* - Users interact with many applications and user interfaces, optionally involving additional user interface components such as assistive technologies. Adaptations of user interfaces, including the activation of assistive technologies, typically require the user to perform manual adaptations for every new device and application. It would be beneficial for the user, if adaptation data could be automatically collected and stored in a standardized way, to be utilized by any device and application the user interacts with. In this paper we present a new approach for user interface adaptation, based on statistical methods. These methods operate on an extended set of user interaction-related properties, which may be part of user profiles. We plan to implement this approach in a prototype on real user data and evaluate its results against manual adaptations by user interface experts.**

Keywords - Adaptive user interface, user interface matching, user profile, user profile matching, statistical analysis.

I. INTRODUCTION

The Internet is a fast evolving network exposing new technologies in a fast pace. In one of the current trends, user-generated files and data are moved from the local desktop to the cloud in the Web. This means that the user stores their documents, emails, pictures, etc. on Internet-based servers rather than their personal computers. Not only is the data stored on the Internet; the application itself to create and edit the data is provided over the Internet, running directly in the Web browser. This allows for accessing and editing the same data on a broad range of devices (personal computer, mobile phones, etc.) that have access to the Internet all over the world.

Currently, many new and versatile Web applications are emerging in the Internet. The user is confronted with new dynamic user interfaces on various devices that differ from classical Websites and local applications. These new Web applications give rise to new opportunities in user interaction, in that dynamic user interfaces can be adapted not only for the device and environmental contexts, but also for the needs and preferences of any user. Such user interface adaptations have the potential to increase the usability of Web applications in general, and particularly for people with disabilities.

We identify three different approaches of automatic user interface adaptations: The rule based approach uses heuristics to make decisions on adaptations; the statistical approach uses large collections of recorded user data to make decisions based on statistical analysis of the data. A third approach combines both previous approaches as a hybrid.

Statistical approaches for matchmaking are used in different fields of applications. One important application field is image and video processing. Statistical algorithms are used to find patterns and objects, improve quality, predict movement, etc. [1].

On the Web, statistical approaches are currently being used for various purposes:

- Amazon is collecting data from the users while they are using their shopping portal. Amazon uses statistical algorithms on the collected data to find user preferences on products and align this information with similar users to create selective shopping suggestions for a user [2].
- Dating agencies are using statistical methods to find good matches between people. They analyze the user profiles and their collected data about successful meetings to continuously improve their matching [3].

While these applications use statistical approaches for making recommendations to the user, we want to use statistical algorithms for user interface adaptations based on user profiles.

Research on user interface adaptations is being conducted in a number of projects, but none of them uses a statistical matching approach. The European FP7 projects MyUI [4] and GUIDE [5] are examples of such projects.

Before we describe our concept, we will set the scene with a scenario that describes an instance of what the concept of this paper could achieve. After that we will give a brief introduction to the technical backgrounds for user profiles, user profile matching and user interface matching based on user profiles. In section 4 we will discuss the evaluation methods we intend to use to prove the concepts. Finally, we give an outlook for how we plan to implement and validate the concepts of this paper.

 ISBN: 978-1-61208-167-0

II. Example Scenario

Martha (31 years old) lives alone in her flat in the city center. She has a notebook for surfing in the Web and for writing emails. She also has a smartphone, but uses it only to make calls or to write text messages. At her work, she has a personal computer and writes product reviews.

Martha uses an office suite for writing texts, both at home and at work. She is glad to have managed to use this application, and doesn't know much about other software. To perform the functions in her office suite, she uses the computer mouse in the office and the touchpad on her notebook at home. She doesn't use keyboard shortcuts, except for copy and paste. For other features, such as text formatting, she uses the toolbar of the office application.

At work she got an assignment to write a report, together with her colleagues. They have agreed to use a new Internet-based tool that allows them to work in parallel on the same text. It also promises to adapt automatically to the user's needs and preferences.

Martha goes to the website and registers for the new tool, so that she can use it. Before she can start, she has to fill out a questionnaire, asking a few questions about her preferences for different applications, such as what software she uses, whether she uses shortcuts and, if so, what kind of shortcuts.

In fact, the user interface of the Web application looks very similar to her desktop program. She can start working immediately without learning the new interface. There are a few functions she doesn't know. However, the program informs her about the new functionality and tells her that other users (that have a similar profile to hers) have rated these functions as very practical. Martha reads the descriptions and decides that two of the four functions might be useful for her and disables the other two functions. Internally, the system recognizes the user adaptations and stores them in the user profile, so the functions will no longer appear.

Over the time, Martha starts to use other Web applications and is very pleased how well the user interfaces adapt to her behavior and needs. In particular, changes done in one application are also reflected in the other applications. Meanwhile, she wants this functionality for all of her applications.

III. Technical Background

A. User Profile

The foundation for adapting a user interface to a specific user in general, and in particular for above scenario, is the user profile. It stores various data that can be used to derive an optimal adaptation of the user interface.

Data mining of user profiles is essential if we want to apply statistical methods to user interface adaptation. We need to operate on rich information items on the user, which is more than mere demographic data. Information about the user preferences and patterns of typical user interactions are required to make informed and reasonable decisions. This can be achieved by user modeling techniques [6] [7].

A user profile should therefore include:

- Demographic data;
- Preferences on features, settings, shortcuts, etc.;
- Preferred user interaction strategies (user interaction patterns);
- User interaction enhancements, such as assistive technologies.

Of course, as the profile holds sensitive user data, security and privacy issues about data storing and data access need to be thoroughly considered in an actual implementation. However, in this paper we focus on the concept of user profile matching.

B. User Profile Space

An initial user profile is normally incomplete, because user interaction patterns or user preferences are recorded gradually while the user works with various applications. In general, only demographic data and simple preference data, which is present at the beginning, is collected by an initial questionnaire. This data is not sufficient to produce any reliable results by the means of statistical analysis. One possible solution to this problem is to borrow the missing data from other similar, though more complete, user profiles. For this task a user profile space is needed, that is generated from a high number of user profiles. Such profiles are collected over a long period of time to yield enough interactions and preferences of the users. It is also possible to import user profiles from existing profile databases or to create them from virtual users or personas.

A new user profile is classified by a profile matching component and is afterwards integrated into the user profile space, using the available data stored in the profile (see fig. 1). After the classification, similar user profiles can be used to fill the gaps in the new user profile. It is important to note that at this point, this data is only a hypothesis by the system, and is therefore assigned with appropriate probability levels. However, with every new interaction of the user with any application, this hypothetical information becomes either corrected or enforced, resulting in higher probability scores.

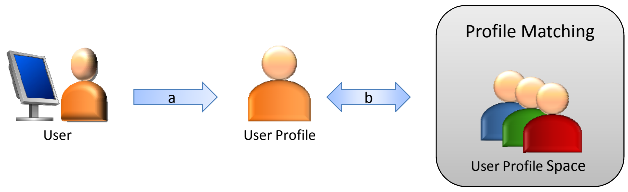

Figure 1. Sequence for (a) creating a new user profile and (b) matching it to the user profile space. After the matching the user profile is supplemented with data from similar profiles.

By following this process, we can constantly extend and improve the user profile space, with new user profiles being integrated and existing profiles validated by incoming user data (see fig. 2).

We will use a statistical approach for matching user profiles into the user profile space. One candidate is the principal component analysis (PCA) [8]. This method transforms a high dimensional feature space, derived from the user profiles, to so-called principal components. Ideally,

 ISBN: 978-1-61208-167-0

no data gets lost with the transformation, and redundant data gets combined to correlated data.

A new user profile can be aligned along the principal components to get an optimal classification in the user profile space. Now, the correlated data can be used for the completion of the missing information of a user profile. Additionally, a plausibility weight in form of a probability is assigned for any derived information item of the user space. This probability is continuously adjusted to user profile space updates, e.g., when new user profiles are added or profiles get updated with new data from a user.

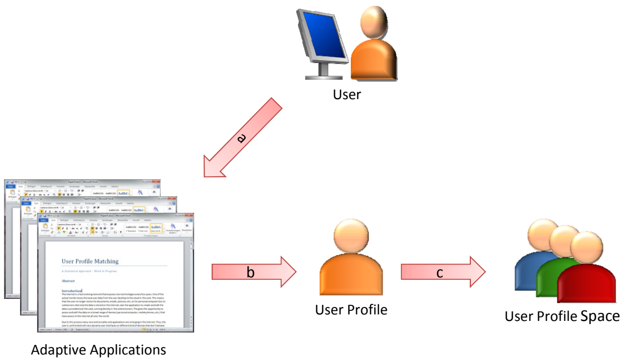

Figure 2. Sequence for updating the user profile and user profile space.
a. The user starts a new application.
b. Manual adaptations made by the user are stored to the user profile.
c. Updating the user profile also updates the user profile space.

C. *User Interface Matching*

For the matching of a user interface based on a user profile, two different means of adaptation should be considered:

- Static configuration of (assistive) tools, which enhance an application (e.g., screenreader)
- Direct customization of an application (e.g., changing font size, or reordering menus)

For direct customization, an interface matching component obtains a list of possible adaptations of an application. From this list, it provides an optimal adaptation, as it compares the customization options with the user profile. The result is passed on to a special service (interface adaptor), which is responsible for performing the adjustments on the corresponding platform or application (see fig. 3).

For the matching itself, different approaches can be used: rule-based, statistical or a hybrid approach combining the rule-based and statistical approaches.

In the statistical approach, appropriate algorithms from the field of machine learning and statistical analysis will be used for the matching. A classical candidate approach is based on Support Vector Machines [9].

They perform a classification that can make a statement about which adaptation options make a good match to a user profile. This classification is based on previous adjustment decisions made by the user or similar users. Therefore, extensive training datasets must be created before this approach yields reliable results.

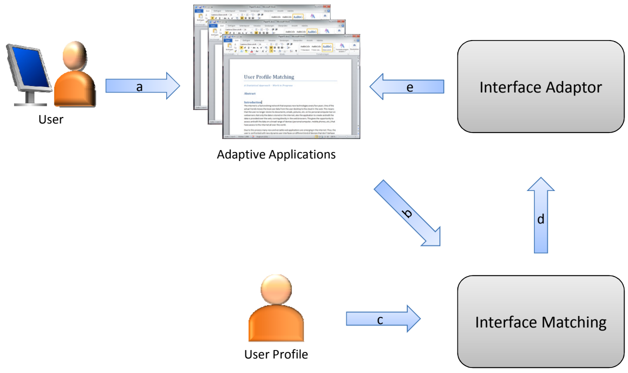

Figure 3. Sequence for matching a user interface to a user profile
a. The user starts a new application.
b. The application starts the user interface adaptation by transfering the adaptation feature list to the interface matching algorithm.
c. The matching algorithm also receives the user profile of the current user to perform the matchmaking.
d. The matchmaking results are forwarded to the interface adaptor.
e. The interface adaptor applies the adaptations on the running application.

In addition, the user can correct wrong decisions of the classifier by manually undoing or changing adjustments. These user adjustments are stored in the user profile and will influence future decisions of the matching algorithm.

IV. Evaluation

The implementation and fine-tuning of the proposed approach requires an iterative approach based on real user data and extensive assessments of the results achieved.

At the beginning, we will develop an initial set of training and validation data. Based on this data, we will experiment with existing statistical methods and their parameters. The results of these approaches will be compared against the set of validation data. As more data becomes available, we will conduct additional optimization through parameter adjustments, and conduct small-scale user tests to compare this approach to the other approaches (rule/heuristics-based, hybrid, expert).

We will use manual adaptations by user interface experts as a benchmark to compare and evaluate the automatic adaptation strategies. The quantitative evaluation will be based on a compound metric, consisting of measurements on effectiveness, efficiency and user satisfaction. Our goal is for the statistical approach to come as close as 80% to the benchmark of the expert-provided adaptations.

V. Outlook

In this paper, we introduced a concept for adapting user interfaces based on user profiles using statistical methods. For this, not only the matching algorithm is important, but also user profiles and their structure, collecting and storing user information and the classification of new user profiles within a user profile space.

ISBN: 978-1-61208-167-0

One of the main tasks is to find optimal statistical methods for the various problems introduced in this paper. In order to do this, statistical approaches will be implemented on real profile data, applied, evaluated and compared among each other.

We intend to research and implement the described concept for matching between user interfaces and user profiles in a future European project. Together with other project partners we will also investigate additional approaches, including a rule-based and a hybrid approach, which combines the rule-based with the statistical approach.

REFERENCES

[1] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2007.

[2] T. Segaran, Programming Collective Intelligence, O'Reilly, 2007, pp. 7-28.

[3] T. Segaran, Programming Collective Intelligence, O'Reilly, 2007, pp. 197-224.

[4] Mainstreaming Accessibility through Synergistic User Modelling and Adaptability (MyUI). European FP7 project. http://www.myui.eu/

[5] Gentle User Interfaces for Elderly People (GUIDE). European FP7 project. http://www.guide-project.eu/

[6] A. Kobsa, Generic User Modeling Systems. User Modeling and User-Adapted Interaction 11, 2001, pp. 49-63.

[7] J. Sobecki, Consensus-Based Hybrid Adaptation of Web Systems User Interfaces. Journal of Universal Computer Science 11, no. 2, 2005, pp. 250-270.

[8] I. T. Jolliffe, Principal Component Analysis, 2nd ed., Springer, 2002.

[9] B. Schöllkopf, A. J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, The MIT Press, 2001.

 ISBN: 978-1-61208-167-0