

Integration Architecture of User Models

Aivars Niedritis, Laila Niedrite, Natalija Kozmina

Faculty of Computing, University of Latvia, Raina boulv.19, Riga
{Aivars.Niedritis, Laila.Niedrite, Natalija.Kozmina}@lu.lv

Abstract. In this paper we have presented integration architecture of user models of different systems for two purposes: (i) to give recommendations for the data warehouse user model and personalization, and (ii) to find out the recommendations for data warehouse model evolution and further development. A general user model has been proposed based on the concept of user describing profiles that represent the interaction between a user and a system and reflect different aspects of their interaction. User characteristics, interaction, temporal, spatial, preferential, and recommendation profiles have been developed. The proposed general user model is employed as a main component of the user profile integration architecture.

Keywords: user model, model integration, data integration, data warehouse.

1 Introduction

A user profile is a set of information that describes a user. Attributes of the user profile typically store information about user's identity, preferences, etc. We can treat the user profile as a simplified user model. The user model at the same time describes user from diverse aspects, e.g. Interaction, Geographical, Demographical, etc., that are considered to be important in the context of a given system. So, different user profiles can be used to collect the information about users and also to develop a comprehensive user model.

User models are utilized for the adaptation of systems [1], for personalized answers to user requests [2]. They are used in different types of systems, e.g. recommender systems, search engines, and others. Also, in the area of information integration, for example, in data warehousing, the personalization aspects are studied [3], [4] and user models can be used for personalization purposes. Some approaches use ontology as a basis for the creation of the user model [5], [6]. Other authors provide architectures to build and apply a user model [7].

Users usually interact and employ different systems available on the Web as well as internal systems of organizations. All these systems might contain different pieces of user information that could be used to gather a comprehensive view about users. There exist works done to integrate distributed publicly available user data, e.g. in the case of Social Web [8]. In this paper we will consider internal systems of the organization and a data warehouse as one of these systems.

The user profile life-cycle consists of three phases: (i) information gathering about the user, (ii) personalization that constructs integrated view of the user (integrated profile), and (iii) usage that provides personalized services to the user based on the user profile [9]. We will concentrate mainly on the first two phases in this paper. The usage possibilities are shown in the motivation section further in this paper.

The rest of the paper is organized as follows. The 2nd section gives an overview of the method for constructing a user model based on the concepts of user-describing profiles. The 3rd section describes an integrated user profile and its development. The 4th section ends the paper with conclusions and future work.

2 The Method for Construction of the User Model

In our previous work [10] we have introduced a concept of user-describing profiles for data warehouses. These user-describing profiles are meant to be utilized for personalization purposes of data warehouses.

2.1 The Concept of User-Describing Profiles

User-describing profiles form a set of profiles that characterize interactions of the user with a data warehouse system from different viewpoints. We based the definition of the profiles on Zachman Framework concept [11], using the following questions: *who*, *what*, *how*, *when*, *where* and *why*. So, the set of user-describing profiles consists of six different profiles:

- *Preferential* profile represents the expectations of users – *what* are the interests of the user,
- *User* profile represents users' basic data, e.g. personal data – *who* is the user,
- *Spatial* profile represents physical location of the user – *where* is the user,
- *Temporal* profile represents the time - *when* does the user interact with the system,
- *Interaction* profile mostly represents the way of interaction – *how* the interaction is performed, e.g. by means of a particular technical device and by means of a particular program version.
- *Recommendation* profile represents the goals of the interaction – *why* the user is interested in the particular system; this profile is generated based on other profiles to develop recommendation for the user.

Proposed profiles do not describe the user himself/herself but the whole environment of the interaction between the user and the data warehouse system. User, spatial, temporal, interaction and preferential profiles altogether compose a comprehensive description of the user.

2.2 Steps and Outcomes of Constructing a User Model

Each user-describing profile is based on the collection of comprehensive lists of attributes gathered during literature studies and documenting our practical experience.

Potential sources of attributes are: (i) books about particular type of systems (e.g. data warehousing, see [12]), (ii) standards in the corresponding field (e.g. CWM, see [13]), (iii) scientific and technical papers, (iv) practical experience in case studies, for example, in particular data warehousing projects, and in usage of particular tools and web-services (e.g. [14]).

The attribute set of each profile has been logically split into classes and a class diagram has been developed for each profile. Time profile is one of the constructed profiles; its class diagram is depicted in Figure 1.

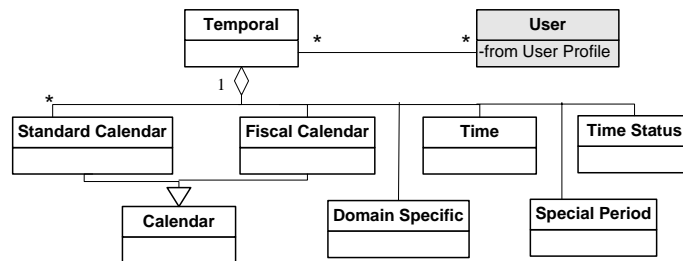


Fig. 1. Class diagram of the time profile.

Detailed description of user-describing profile classes is given in [10]. However, for the better understanding of the method and the presented general user model, a short description of user-describing profiles is given.

- User profile classes: *Role* attributes describe system role of the user; *Personal* contains personal information; *Work* and *Education* attributes describe user work and education; *Session* describes user session characteristics; *Activity* attributes describe user activity on a certain webpage in a certain period of time, *Rights* contains attributes that describe user rights for certain objects of the system.
- Temporal profile classes: *Standard Calendar* and *Fiscal Calendar* are two Calendar types, *Time* contains non-calendar attributes and attributes that represent date as a number, *Time Status* contains attributes of yes/no type, *Domain Specific* attributes are specific for a domain (e.g. semester for education, etc.), *Special Period* contains attributes that describe certain planned or spontaneous global or local events.
- Spatial profile has two classes: *Physical Location* (attributes that describe person's physical address) and *Location by IP* (attributes derivable from user IP-address by means of web-services).
- Interaction profile classes are *Web Access*, attributes of which describe operating system, web-browser and Internet connection properties, *Functional*, attributes of which describe web-browser functional properties and supported applications, and *Visual Layout*, attributes of which describe visual layout properties in web-browser.

Each class may be complemented with more attributes, if necessary.

For the construction of preferential and recommendations profiles a different method is used. The preferential profile is a set of attributes from other profiles that is composed by user, while he/she expresses his/her own preferences. Recommendations profile contains sets of preferences that belong to different users.

User-describing profile connections are depicted in Figure 2.

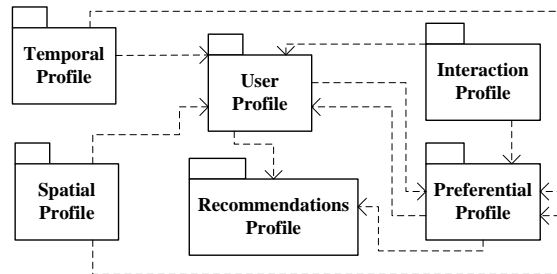


Fig. 2. Connections of user-describing profiles.

One user may have more than one spatial, temporal, interaction, preferential and recommendations profile when he works in different environments. Preferential profile is connected with all other profiles, because of the definition of the profile, but recommendations profile is used to describe offers to the user from other users' preference sets and the choices of the users among them.

3 Integrated User Profile

Based on the user-defining profiles described in the previous section, we enhanced this model to get a general user model that represents the needs of more than one system that have their own user profiles. Due to the space limitations, we will analyze in more detail only one component of the general user model – the general and also the integrated user profile. We will consider the case when the user interacts with internal systems of an organization and a data warehouse as one of these systems.

3.1 Motivation

There are two major goals of producing a general user model in the data warehousing field:

- *Improvement and extension of a data warehouse model (developer perspective).*

Gathering user activity in distinct data sources is useful for keeping track of functions executed and data explored by each user. Users may be grouped in a certain way – for instance, by the level of skills of the user, by access rights, by the role of the user, etc. Thus, understanding the common tendencies of different groups of users, it is possible to improve the data warehouse, i.e., to add new data marts to existing ones that would reflect statistical data on various components of data sources, most actively employed by different groups of users.

- *Personalization of the data warehouse reporting tool (user perspective).*

Sometimes a user while browsing data warehouse reports has no notion about what kind of data he/she is able to find in data warehouse reports. Moreover, a user might be unaware of potentially useful report, because, for instance, it has been

created recently and the user hasn't examined it yet. It is essential to involve personalization in data warehouse reporting tool, i.e., to acquire and process user preferences in order to suggest him/her other reports that might be helpful. As a result, a user will be able to orient in a variety of data warehouse reports, saving time and effort.

3.2 User Model Properties

To develop a general user model we used the structures introduced in our previous work on user profiles [10] as a basis and merged them with some structures of the learner profile [15]. Here we provide a brief description of each structure that is a component of an integrated, general and also extendible user model (see Figure 3).

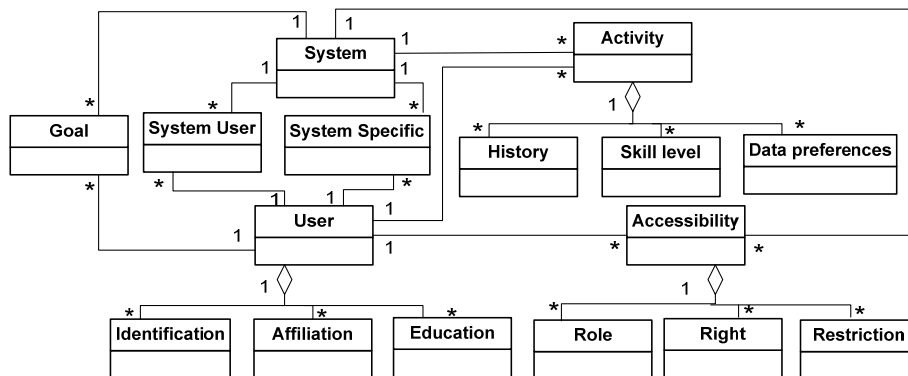


Fig. 3. Class diagram of user profile.

The *Identification* structure includes user's personal information (e.g. first name, last name, date of birth), demographical information (gender, age group, etc.), contact information (telephone, e-mail, address) and others.

The *Affiliation* contains user professional attributes (e.g. position, company name, total years of experience, etc.)

The *Education* class describes the education of the user (e.g. currently student, educational institution, year of graduation, diploma nr., honors, etc.).

The *Accessibility* describes the user's ability to interact with the data source or data warehouse environment of the user. This structure consists of such components as the *Role* attributes of the user, user's *Rights* attributes in each data source, and *Restrictions* (e.g. user's disabilities or languages chosen). Interaction from the system party (e.g. web-browser functional properties, etc.) is described within the Interaction profile.

The *Activity* is used for keeping track of either detailed or summarized history of user actions and frequency of use of each system's components. For example, hit count and time spent on a certain component in a specified period of time, the overall number of sessions, last session start/end time, user's IP address, the total number of system components accessed, the number of system components accessed relative to

the total number of components available for the user with certain rights, etc. *Skill level* attributes defines user's experience in working with data sources and data warehouse reporting tool. For instance, it is possible to follow if the user has worked with data sources and how he/she has worked (time, frequency etc.). The *Data Preferences* include user preferences on data in data sources and in DW reporting tool, preferences on reports' structure, and others.

Goals' attributes are associated with user's role in each system. Also, monitoring the set of components accessed by the user and capturing the frequency of access we can determine the goals of usage of the system – e.g. business, personal, learning, analytical, data entry, and others.

System specific attributes are meant to be used for the extension possibilities of the described general model, when a system with different important attributes that extend the user model is introduced in an organization.

System user attributes describe the users' login information for each particular system.

3.3 Integration of User Profiles

Let us consider a situation when one physical person interacts with two different source systems for a data warehouse. Both systems have their own local different user profiles. The Figure 4 shows the integration processes of these two user profiles for the further use, according to the goals described in the motivation section.

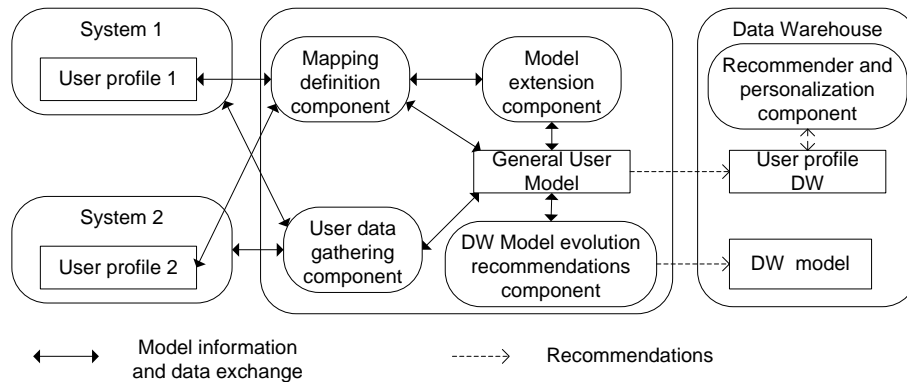


Fig. 4. Integration architecture of user profiles.

The main steps of the user profile integration process are the following:

- Mapping definition – mapping is defined for each particular user model against the general user model. Three different situations can happen: (i) there exist automatically detected corresponding attributes by identical attribute names, (ii) semantically equal attribute exist with different attribute names, and (iii) corresponding attribute can not be found. Last two cases should be processed manually and in the (iii) case the next step is performed.

- User Model extension activity is used to decide which attributes from local user models from the previous step in the (iii) case are essential for the general user model and then the necessary changes are made.
- User data gathering step is performed when all local user profiles are processed, mappings defined and, if necessary, the general user model is extended. According to the mapping information, the data is collected and moved into corresponding integrated general user model.
- Recommendation component of the proposed architecture is used to analyze the gathered information (to identify contradicting user profile data for the same general user model attribute, e.g. different rights on the same type of objects in both systems). Further two activities are performed:
 1. Recommendations for the data warehouse user model are prepared – mostly this is useful, when new data warehouse reports or new data warehouse users are introduced to define access rights. The integrated profile information could be used also for the personalization of the data warehouse system to provide the most appropriate information to the user.
 2. Recommendations for further development of the data warehouse are given, e.g. which new data marts could be most interesting for the user based on the user activity information gathered in the source systems.

4 Conclusions and Future Work

This paper describes the general user model taking one of the user-describing profiles – i.e. user profile – as an example. This component is the most essential in the whole user model, however, other user-describing profiles should be considered as well in the context of integration.

Also, a user profile integration architecture that employs the general user model is proposed and described for an example situation with two data source systems and a data warehouse. The number of the involved systems is not limited, but with more systems involved the manual adjustments in the integration process to solve the integration conflicts could be more often possible. We have not described in this paper the likely conflicts of attribute values of integrated user profiles; different solutions are possible depending on the semantics of the analyzed attribute.

The experimental environment, which is used for the implementation of the general user model development and integration of user profiles, is the university information system, e-learning system and university data warehouse. The evaluation of recommendations about data warehouse model development is the next step of the research to be performed. The data warehouse user model will be utilized as one of the data sources for the ongoing research about data warehouse personalization.

Acknowledgments. This work has been supported by ESF project No. 2009/0216/1DP/1.1.1.2.0/09/APIA/VIAA/044.

References

1. Brusilovsky, P.: Adaptive hypermedia. *Journal of User Modeling and User Adapted Interaction*, 11 (1/2), pp. 87--110 (2001)
2. Koutrika, G., Ioannidis, Y. E.: Personalization of Queries in Database Systems. In *Proceedings of 20th International Conference on Data Engineering*, pp. 597--608, (2004)
3. Jerbi, H., Ravat, F., Teste, O., Zurfluh, G.: Preference-Based Recommendations for OLAP Analysis. In *Proceedings of DaWaK'09*, pp. 467-478, (2009)
4. Golfarelli, M., Rizzi, S.: Expressing OLAP Preferences. *Scientific and Statistical Database Management, LNCS*, vol. 566/2009, pp. 83-91, Springer, Heidelberg (2009).
5. Razmerita, L., Angehrn, A., Maedche, A.: Ontology Based User Modeling for Knowledge Management Systems. *Proceedings of the 9th International Conference on User Modeling*, pp. 213--217, Springer-Verlag (2003)
6. Zhang, H., Song, Y., Song, H.: Construction of Ontology-Based User Model for Web Personalization. *Proceedings of the 11th International Conference on User Modeling*. pp. 67--76. (2007)
7. De Bra, P., Aroyo, L., Chepegin, V.: The Next Big Thing: Adaptive Web-Based Systems. *Journal of Digital Information*, 5 (1), (2004).
8. Abel, F., Henze, N., Herder, E., Houben, G.J., Krause, D., Leonardi, E., Building Blocks for User Modeling with Data from the Social Web. *Proc. of International Workshop on Architectures and Building Blocks of Web-Based User-Adaptive Systems at UMAP*, (2010)
9. Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A.: User Profiles for Personalized Information Access. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.): *The Adaptive Web. LNCS 4321*, pp. 54--89. Springer, Heidelberg (2007)
10. Kozmina, N., Niedrite, L.: OLAP Personalization with User-Describing Profiles. In: Forbrig, P., Günther, H. (eds.) *BIR 2010. LNBIP*, vol. 64, pp. 188--202. Springer, Heidelberg (2010)
11. Zachman, J. A.: *The Zachman Framework: A Primer for Enterprise Engineering and Manufacturing*. Zachman International, (2003).
12. Imhoff, C., Gallemmo, N., Geiger, J. G.: *Mastering Data Warehouse Design: Relational and Dimensional Techniques*, Wiley Publishing, USA, (2003)
13. Poole, J., Chang, D., Tolbert, D., Mellor, D.: *Common Warehouse Metamodel Developers Guide*, Wiley Publishing, USA (2003)
14. IP Address Geolocation to Identify Website Visitor's Geographical Location. <http://www.ip2location.com/>
15. IMS Learner Information Package Information Model Specification (2001), <http://www.usit.uio.no/prosjekter/eSU/eSU-revisjon/rapporter-referater/eksterne-rapporter/lip/lipv1/lipinfo01.pdf>