

Semantic Personalization of Web Portal Contents

Christina Tziviskou, Marco Brambilla
 Politecnico di Milano
 Via Ponzio 34/5, I-20133 Milano, Italy
 0039 0223993649

tzivisko @elet.polimi.it, mbrambil@elet.polimi.it

ABSTRACT

Enriching Web applications with personalized data is of major interest for facilitating the user access to the published contents, and therefore, for guaranteeing successful user navigation. We propose a conceptual model for extracting personalized recommendations based on user profiling, ontological domain models, and semantic reasoning. The approach offers a high-level representation of the designed application based on a domain-specific metamodel for Web applications called WebML.

Categories and Subject Descriptors

D.2.2 [Software Engineering]: Design Tools and Techniques – computer-aided software engineering (CASE), evolutionary prototyping, object-oriented design methods, user interfaces.

H.5.4 [Information Presentation and Interfaces]: Hypertext / Hypermedia – architecture, navigation, theory, user issues.

General Terms

Algorithms, Design.

Keywords

Semantic Web, Personalization, Ontology, Conceptual modeling.

1. INTRODUCTION

Within traditional Web applications, the user navigation follows the predefined hypertext structure. Therefore, finding contents requires the user understanding of the Web site outline, which is not always obvious. Enriching the Web application with personalized recommendations provides alternative paths to published data, and increments the possibilities for the user to find the contents he is interested in. However, the effectiveness of personalization is based on the quality of the user profile and of the relations among the content objects. Modeling the published data and the user profile with ontologies allows to express more effectively the user interests and the relations between the pieces of information, by leveraging the advanced features of Semantic Web technologies. Such semantic relations may be exploited for more accurate personalization results.

In this work we extend the framework presented in [3], that provides: conceptual modeling of the domain ontology; new conceptual primitives that explore the ontological model and extract data; a basic ontology-based profiling mechanisms; and reasoning methods for presenting personalized contents. The current proposal presents a more accurate profiling through a complete algorithm for reasoning on the domain ontology and on the user profile in order to publish personalized data.

Copyright is held by the author/owner(s).

WWW 2007, May 8–12, 2007, Banff, Alberta, Canada.

ACM 978-1-59593-654-7/07/0005.

2. CONCEPTUAL MODELING OF SEMANTIC WEB APPLICATIONS

Our proposal of personalization of Web applications is specified using WebML (<http://www.webml.org>), a high-level language for modeling Web applications based on data [4], processes [2] and Web services [5]. The WebML specification of a Web application consists of a *data schema*, and of one or more *hypertexts* called *site views*, expressing the Web interfaces. A site view is a graph of *pages*, which in turn contain *units*, the atomic publishing primitives that extract contents from the data source. *Links* between units define the navigation paths and carry data between units. Business logics can be specified through *operation units* updating the underlying data or performing other actions.

In [6] the WebML language has been used for defining the metamodel that describes the WSMO [<http://www.wsmo.org>] components of an Ontology and has been extended with a set of new components (units and operations) for exploring ontological contents within Web applications design. The WSMO metamodel includes the class *Concept* containing the defined concepts, possibly organized in a hierarchy and related to each other through *Relations*, and having properties denoted as *Attributes*. The actual instances are managed by the classes *Instance*, *Attribute Value*, and *Relation Instance*.

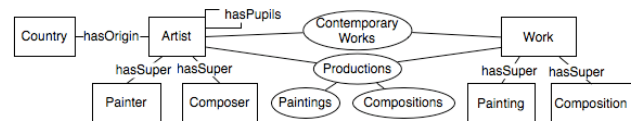


Figure 1. Ontological data model for a Web portal

Fig. 1 presents part of the ontology of a Web portal about artists and their works. Concepts are depicted as rectangles, attributes as lines, and relations as ellipses. The following instances exist: Mozart, Beethoven, and Haydn are Austrian Composers; Bach is a German composer; Elgar is an English Composer; Klimt is an Austrian Painter; Mozart is a pupil of Haydn; and Beethoven composed “Moonlight Serenade”.

The model-driven personalization mechanisms we propose are based on explicit preference declarations by the user and on an iterative process of monitoring the user navigation, collecting its requests of ontological objects, storing them in its profile, reasoning on them and on the domain ontology for delivering personalized content. The personalization is provided for both registered and unregistered users, with different quality levels. The User Profile Ontology models: (a) identification data for the user; (b) user explicit preferences on ontological objects (for usability reasons, the user can usually choose among a subset of the whole ontology); and (c) user requests on ontological objects related to its navigation.

Note that ontology object requests are not explicit. They are automatically registered by exploiting semantic annotations of visited pages. These annotations do not need to be defined completely manually, since most of them can be automatically calculated from the conceptual model of the application [1]. We enrich the WebML links with operation chains that register within the User Profile Ontology the user access to the current link and the instances of the page concept(s) that we want to monitor. Fig. 2 shows a Web page for an Artist, the list of his Works, and the details of the selected work. The user's clicks on the links (1) and (2) transparently trigger the implicit registration of the request of the selected Artist and of the selected Work respectively.

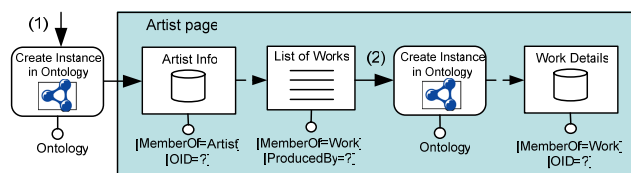


Figure 2. WebML model with implicit semantic profiling

The personalization mechanism is achieved by units placed in the pages extracting the ontological objects (i) semantically related to the objects in the current page, and (ii) contained in the user profile. An algorithm ranks the results giving higher priority to objects that are requested more often and are explicitly preferred.

3. PERSONALIZATION ALGORITHM

We propose an accurate algorithm for extracting personalized recommendations on ontological objects within Web applications. The purpose of the algorithm is to find objects *similar* to the ones published in the current page that might be of *interest* to the user. We consider similar objects as having any relations to the page objects. In the current algorithm, the page developer explicitly selects the relations for calculating the similarity in a page, and indicates their importance by assigning them a weight. Among the similar objects, the algorithm discards the ones not related to the user interests. In this step, all the relations are considered in order to extract connections between retrieved objects and user interests, although the ranking of a retrieved object gets incremented upon a similarity relation to the user interests. The algorithm presents the remaining objects in a ranking order.

The algorithm is based on the exploration of the following relations: the semantic relations between concepts; the attribute values; the ISA hierarchies of concepts/relations; and the concept type of objects (isMemberOf relationship between instances and concepts). In order to exemplify the algorithm, we assume that the user, by browsing the Austrian composers in a portal defined over the ontology of Fig. 1, has reached the page of Mozart and his compositions. He has already visited the page of Germany, and has explicitly requested to be informed on the “Moonlight Serenade” composition. Therefore, in its profile the registered interests are Mozart, Austria, Germany, and “Moonlight Serenade”. The algorithm is used to enrich the contents in the Mozart page for facilitating the user research.

In the first step, the developer explicitly specifies the relations and the correspondent weights to be used for extracting objects similar to the objects in the page: (a) ContemporaryWorks (weight 3) extracts works produced on the same period with the compositions of Mozart (“Moonlight Serenade”), and Productions (weight 3) extracts all the compositions by Mozart; (b) the

attribute hasPupils (weight 2) extracts all the pupils of Mozart and his teachers (Haydn, and the attribute hasOrigin (weight 2) extracts all the Austrian artists (Beethoven, Haydn, Klimt), (c) the hasSuper relationship (weight 2) extracts the Artist concept, and (d) the isMemberOf (weight 4) relationship extracts all the composers and painters (Beethoven, Haydn, Klimt, Bach, Elgar).

In the second step, the algorithm discards all the retrieved objects that are not related (in any direct way) or are not contained in the user interests (e.g., Elgar). The remaining objects are “Moonlight Serenade”, Haydn, Beethoven, Klimt, Bach.

In the third step, the algorithm ranks the retrieved objects by calculating the total weight of the relations that connect them to the user interests. Empirical tests showed that good values for the weights could adhere to the following guidelines: first step semantic relations from the object have the explicit weight given by the designer; the rest of relations have a weight of 0.5; containment as explicit interests have a weight of 1.5; containment as implicit interests that have a weight of 1. For instance, the weight of the “Moonlight Serenade” is 1.5 since the only relation that connects it to the user interests is the explicit preference. The weight of Beethoven is 9, of Haydn is 8, of Bach and Klimt is 6. Thus, the retrieved personalized contents are displayed in the following order: Beethoven, Haydn, Bach, Klimt, and “Moonlight Serenade”. The ranking for Haydn comes from:

$$\begin{aligned}
 w(\text{Haydn}) &= w(\text{Haydn.hasPupils}(\text{Mozart})) + \\
 &+ w(\text{Haydn.hasOrigin}(\text{Austria})) + \\
 &+ w(\text{Haydn.isMemberOf}(\text{Composer})) = 2 + 2 + 4 = 8
 \end{aligned}$$

4. CONCLUSIONS & EXPERIENCE

In this paper, we have proposed an approach and an algorithm for extracting personalized data within Web applications. The approach is integrated in a framework modeled in WebML that leverages Semantic Web techniques and software engineering solutions for Web application design.

The algorithm proposal presented has been proved valid in a Web portal for educational organizations where we enriched the pages with personalized contents. Such contents get actually selected by the users. Hence the algorithm presents data of interest to the user.

5. REFERENCES

- [1] Brambilla, M., Celino, I., Ceri, S., et Al. A Software Engineering Approach to Design of Semantic Web Service Applications, ISWC 2006, Athens, USA. LNCS 4273.
- [2] Brambilla, M., Ceri, S., Fraternali, P., Manolescu, I. Process Modeling in Web Applications, ACM Transactions on Software Eng. and Methodology (TOSEM), 15 (4), 2006.
- [3] Brambilla, M., Tziviskou, C. Modeling Ontology-based Personalization within Web Applications, JWE, submitted.
- [4] Ceri, S., Fraternali, P., Bongio, A., Brambilla, M., Comai, S., Matera, M. Designing Data-Intensive Web Applications, Morgan-Kaufmann, USA, 2002.
- [5] Manolescu, I., Brambilla, M., Ceri, S., Comai, S., Fraternali, P. Model-Driven Design and Deployment of Service-Enabled Web Applications, ACM TOIT, 5(2), May 2005.
- [6] Tziviskou, C., Brambilla, M. A Knowledge Base Management System for WSMO Ontologies, SWESE 2006 workshop, Intl. Semantic Web Conference, 2006, USA.