Personalized Ontology Based on Consumer Emotion and Behavior Analysis

Saranya.R[1], Sathya.R[2], Umaiyal.R.M[3], Kanmani.P[4]

Final year students, BE-Computer Science and Engineering, K.S.Rangasamy College of Technology, Tiruchengode, Tamilnadu, India[1, 2, 3]

Assistant Professor, K.S.Rangasamy College of Technology, Tiruchengode, Tamilnadu, India [4]

ABSTRACT: The relationships between consumer emotions and their buying behaviors have been well documented. Technology-savvy consumers often use the web to find information on products and services before they commit to buying. We propose a semantic web usage mining approach for discovering periodic web access patterns from annotated web usage logs which incorporates information on consumer emotions and behaviors through self-reporting and behavioral tracking. We use fuzzy logic to represent real-life temporal concepts (e.g., morning) and requested resource attributes (ontological domain concepts for the requested URLs) of periodic pattern based web access activities. These fuzzy temporal and resource representations, which contain both behavioral and emotional cues, are incorporated into a Personal Web Usage Lattice that models the user's web access activities. From this, we generate a Personal Web Usage Ontology written in OWL, which enables semantic web applications such as personalized web resources recommendation. Finally, we demonstrate the effectiveness of our approach by presenting experimental results in the context of personalized web resources recommendation with varying degrees of emotional influence. Emotional influence has been found to contribute positively to adaptation in personalized recommendation.

Keywords----Emotion and behavior profiling, behavioral tracking, recommender system, weblog mining, knowledge discovery, ontology generation, semantic web

I.   INTRODUCTION

It is generally accepted that human emotions are a major motivational factor of human behaviors. In a narrower context, the relationships between consumer emotions and their buying behaviors also seem well documented. In addition, with more and more people connected to the Internet, today's technology savvy consumers are likely to use the web to find information pertinent to products and services before they commit to a purchase. Emotions have been found to influence a person's web surfing behaviors. Discovering and modeling consumers' emotions and surfing habits and behaviors are important for many web applications such as personalized web search and recommendation, e.g., for business applications. In the authors use self-report and behavioral tracking to study how heartbeat communication can improve interpersonal intimacy. Here, we use self-report to incorporate emotions into a personalized consumer profile and web access patterns to model the consumer's mid to long-term web surfing behaviors. Specifically, users are asked to record any change in their emotional state at the end of each web access request. The information is used to gauge the emotional influence of the accessed resources on the user. To capture consumers' access patterns, one promising approach is web usage mining, which discovers interesting and frequent user access patterns from web usage logs. Many web usage mining technique have been developed for mining statistical information and user access patterns in terms of association and sequence of requested resources.

II.   RELATED WORK

A. *Semantic Web Usage Mining*

Since traditional web usage logs only record requested URLs but not the semantics of contents requested by the users, it is difficult to use such logs for tracking the users' actual web access behaviors, emotions, and interests. In response, a number of semantic web usage mining techniques have been proposed. Dai and Mobasher used domain ontology to enhance web usage mining for traditional web usage logs, but the mapping from requested URLs to ontological entities lacks reliability, especially for dynamic websites. Oberle et al.

proposed another framework for semantic enrichment of web usage logs by mapping each requested URL to one or more concepts from the ontology of the underlying website. It clusters groups of sessions with specific user interests from the semantically enhanced weblogs, and applies association rule mining to the semantically enhanced weblogs.

B. Ontology Generation

Ontology typically consists of a finite list of terms and the relationships between these terms. The terms denote important concepts (classes of objects) of the domain, while the relationships include hierarchies of classes. Ontology may also include other information, such as properties, value restrictions, and disjointedness statements. Ontology languages are semantic markup languages for defining ontology. We use Web Ontology Language (OWL) for ontology specification. OWL facilitates greater machine interpretability of web content than XML, RDF, and RDF Schema by providing additional vocabularies Ontology can be constructed manually using an ontology editor. Natural Language Processing (NLP) techniques association rule mining, hierarchical clustering, translation from relational databases and Formal Concept Analysis (FCA)

C. Fuzzy Association Rule Mining

Association rules can help discover relationships between web resources accessed by a user. They can also be used to find groups of people with similar interests. A major problem of traditional association rule mining techniques is that each item in a transaction is considered only to either exist or not. Thus, the user's preference and interest in each transaction item cannot be precisely represented. The fuzzy rule set was found to perform better in prediction accuracy and rule coverage than traditional rule set.

D. Periodic Pattern Mining

Discovering periodic patterns from time series databases is an important data mining task for many applications, such as behavioral tracking. According to the type of patterns, periodic patterns can be divided into periodic association rules and periodic sequential patterns. Periodic association rules are rules that associate with a set of events that occur periodically; such association rules hold only during certain time intervals but not others. Periodic sequential pattern mining can be viewed as an extension of sequential pattern mining by taking into account the periodic characteristics in the time series data.

III.   WEB USAGE ONTOLOGY GENERATION

We propose a process for automatic generation of Personal Web Usage Ontologies (PWUO) of individual users from semantically enriched web usage logs. Fig. 1 shows the proposed approach, which consists of four major steps: Personal Web Usage Lattice (PWUL) Construction, Global Web Usage Lattice (GWUL) Construction, Global Web Usage Ontology (GWUO) Generation, and, finally, PWUO Generation.

Step 1. Using the semantic web usage logs as inputs, it first identifies a set of *periodic attributes* (i.e., temporal concepts such as morning) and a set of *resource attributes* (i.e., useful domain ontological concepts) enhanced with user-reported emotional influence to represent periodic pattern-based web access activities. With the user's web access activities, it constructs a PWUL from the access sessions of the user.

Step 2. It constructs a GWUL to represent all periodic pattern-based *global web access activities* and hierarchical relationships between these activities. As GWUL contains a large number of global web access activities to cover all possible user web access activities, it can be very large. Generally, PWUL is just a small sub lattice of GWUL.

Step 3. It generates GWUO from GWUL by mapping global web access activities and their hierarchical relationships into activity classes and their properties.

Step 4. It generates PWUO for a user by mapping the personal web access activities in PWUL into activity instances of the corresponding activity classes in GWUO. Written in OWL, PWUO is the knowledge base that can subsequently provide personalization facility to the users.
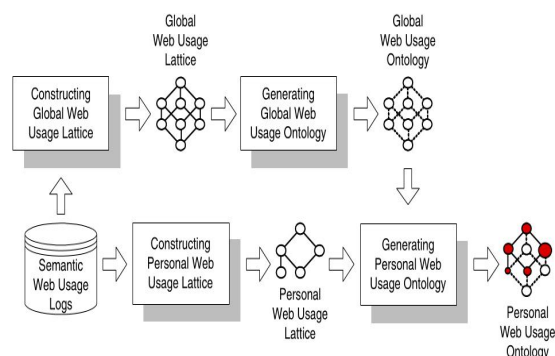

Fig.1 Automatic generation of web usage ontology

A. Semantic Web Usage Logs

Web usage logs can be semantically enriched by associating each requested URL with one or more ontological entities, such as concepts, attributes, and relations, to better describe the patterns of web navigation.

TABLE I

Semantically Enriched Web Usage Log

Use 1	21/May/2 1	8:2 : 1	UR 1	3	#Topic2, #Topic3,...
User1	21/May/2 1	08:22:32	URL2	3	#Topic7, #Topic5,...
User2	21/May/2 1	08:22:50	URL7	4	#Topic1, #Topic8,...
User1	21/May/2 1	08:27:30	URL3	2	#Topic3, #Topic1,...
User3	21/May/2 1	08:33:10	URL3	1	#Topic6, #Topic2,...
User1	21/May/2 1	09:10:02	URL5	4	#Topic7, #Topic3,...
User3	21/May/2 1	09:17:32	URL6	3	#Topic2, #Topic1,...
User2	21/M /2 1	09:26:17	URL4	5	#Topic3, #Topic7,...

B. Personal Web Usage Lattice Construction

The PWUL Construction step is presented as follows: In Section 1), we describe preprocessing performed on the weblog data to extract the most relevant information for further processing. Next, since periodicity information in crucial in our discussion, we need to define the concept of a period in the present context.

1) *Preliminaries:* We perform preprocessing tasks on the semantically enriched weblogs similarly to those for traditional web server logs as data cleaning, user identification, and session identification. The purposes are to discard unsuccessful requests, unnecessary data (e.g., scripts), and to identify all personal access sessions for each individual user.

2) *Web Usage Context:* We now construct the Web Usage Context for a user from his preprocessed user access sessions. Table 1 shows an example Web Usage Context of a user, which consists of five user access sessions, three periodic attributes "P1 (Late Afternoon)," "P2 (Evening)," and "P3 (Night)," and three resource attributes "R1 (Sports)," "R2 (Games)," and "R3 (Chat)."

Before we can construct the PWUL, we still need to define the following: the set of attributes common to user access sessions, the set of user access sessions having the same attributes, fuzzy support of a set of attributes, and web access activity.

TABLE II

A Cross Table of an Example

Fuzzy Periodic Web Usage Context of a User

SessionID	P1	P2	P3	R1	R2	R3
S1	0.6	0.5	0	0.8	0	0.6
S2	0	0.4	0.8	0	0	0.9
S3	0	0	1.0	0	0.7	0.5
S4	0	0.8	0.6	0.8	0.6	0.5
S5	0	0	1.0	0	0.9	0

3) *Personal Web Usage Lattice:* We now identify the activity relationships in order to construct PWUL. Fig. 2 shows the PWUL obtained from the data in Table 1. Each node in the lattice represents a web access activity with membership values of its attributes, and the corresponding fuzzy support and confidence values.
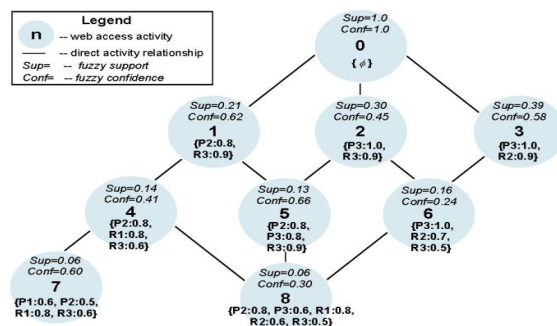

Fig. 2. Personal Web Usage Lattice.

Definition 1: A Personal Web Usage Lattice based on a Web Usage Context $K=(G,M_p,M_r,I)$ of a user is $L_p=(W_p,<w_p)$, where W_P is the set of all web access activities, and $<w_p$ is a partial order on W_P to represent the hierarchy of web access activities.

C. Global Web Usage Ontology Generation

Typically, ontology consists of taxonomy with a set of inference rules. The taxonomy can be expressed as a set of domain concepts (i.e., classes of objects) and the relationships among them (i.e., class hierarchy). Based on the formal definition of ontology defines GWUO as follows:
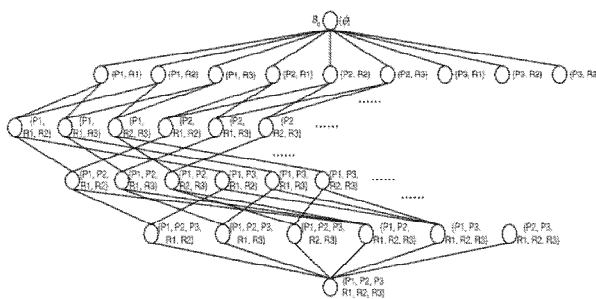

Fig. 3. Global Web Usage Lattice.

The class and hierarchy mapping is performed in the following two steps:

1) *Class mapping:* All global web access activities in the GWUL are mapped to activity classes in the GWUO. The name of each activity class is given based on the attributes involved in its corresponding global web access activity with "Activity_" as the prefix.

2) *Hierarchy mapping:* The activity hierarchy in the GWUL is mapped to the activity class hierarchy in the GWUO. All other activity classes are its (direct or indirect) sub activity classes and inherit its properties. The activity class based on the global web access activity. only direct super activity classes are declared in the "partial" part in the definition of each activity class as all indirect super activity class relationships can be derived from direct super activity class relationships

The property mapping comprises attribute property mapping, taxonomy property mapping, and quality property mapping: Attribute property mapping. All periodic and resource attributes are mapped to attribute The name of each attribute property is given according to its corresponding attribute with "during" as the prefix of periodic attributes or "access" as the prefix of resource attributes. For example, the property based on the periodic attribute P2 is named "during Evening" and the property based on the resource attribute R3 is named "access Chart".

Only the activity class "WebAccessActivity" is set with domain and range; all other activity classes are its direct or indirect sub activity classes and inherit the has Direct Sub Activity property.

Quality property mapping. Two quality properties and Confidence are created to represent values of support and confidence of activity instances.GWU can be generated after the automatic class and hierarchy mapping and property mapping processes. Representation of the activity class "Activity Evening Chat" in the GWUO which is mapped from the global web access activity in the GWUL.

OWL allows extension of imported definitions without the need to modify the original ontology and supports incremental ontology construction, making it easy to incrementally update the activity class hierarchy. When new periodic attributes or resource attributes are introduced, we just need to create new activity classes with new attributes and insert new subclass relationships between the new and existing activity classes to extend the activity class hierarchy.
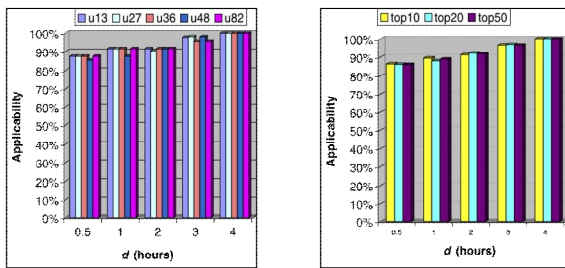
Definition 2: A Global Web Usage Ontology is OG = (C,<c;Q), where

- C is a set of activity classes (or concepts).
- <c is a partial order on C, called activity class hierarchy or taxonomy. If c_i < C c_j, for c_i, c_j $\sum$ C ($i \neq j$), then c_i is a subactivity class of c_j and c_j is a superactivity class of c_i. If there is no c_k $\sum$ C ($k \neq i$ and $k \neq j$) with c_i <C c_k <C c_j, then c_i is a direct subactivity class of c_j, and c_j is a direct superactivity class of c_i. This is denoted as c_i C c_j.
- Q is a set of properties (or relations) which consists of attribute properties QA = {q_i} to represent periodic and resource attributes of activity classes, one taxonomy property qT to represent the direct subactivity relationship between two activity instances, and two quality properties, q_{sup} and q_{conf} , to represent the values of support and confidence of each activity instance.

IV.   PERFORMANCE EVALUATION

We evaluated the performance of the proposed web usage ontology generation approach based on the effectiveness of its generated periodic access patterns for supporting web personalization. .The web forum had seven main topics (such as Past time, Sports, Computer, etc.) and 57 subtopics (such as MovieTV in Past time, soccer in Sports, etc.).

Access data of the top 50 users (measured in terms of the number of access activities) were used in the experiments. We used web access sessions of the users over a 20 day period as the training data set, and web access sessions over another 10 days as the testing data set.Three experiments (two objective and one subjective) were conducted. In the experiments, we constructed Personal Web Usage Lattices for the users from the training data set. We employed two evaluation measures in the experiments: applicability and satisfaction. These measures were derived from the well-known precision and recall measures adapted for the present setting.

(a) Top 5   (b) Averages

Fig. 4. Applicability results.

PWUL only stores typical web access activities if the period-supported activities cannot be found, the generated set of personalized resources will be empty. Therefore, applicability measures how often applicable sets of personalized resources will be generated.

A. Effectiveness of Periodic Access Patterns

In this experiment, we generated personalized resources for a set of predefined period conditions and measured the performance based on applicability and satisfaction with respect to different durations of period conditions d (from 0.5 to 4 hours) and different numbers of personalized resources N_{PR} (from 1 to 6), while keeping e_r fixed at 1. The purposes were to gauge the effectiveness of the approach.
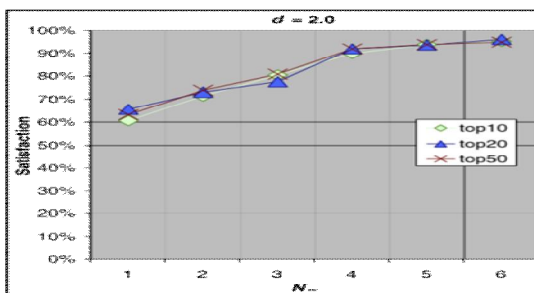

Fig.5 Satisfaction results

B. Effects of Emotional Influence

We evaluated the effects of emotional influence by adjusting the value of e_r based on the user input −E score using the three mapping rules i.e., −E (no emotional influence), +E1 suppresses negative emotional influence), and + E2 (different weights for different degrees of emotional influence). The pplicability results for $d = 0.5$ to 3.0 are summarized in Fig. 6.

(a) $d = 0.5$ (b) $d = 1.0$

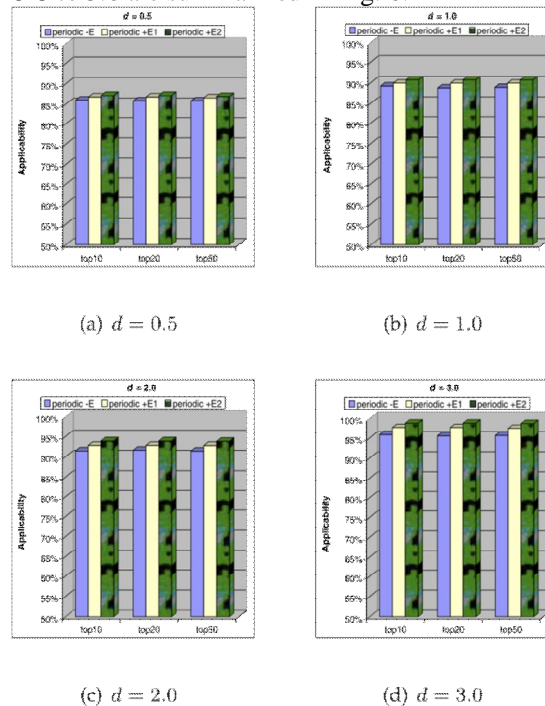(c) $d = 2.0$ (d) $d = 3.0$

Fig. 6. Comparisons of applicability results.

For our approach, the time of each request in the testing data set was regarded as the period condition and used for generating periodic personalized resources; the nonperiodic approach used the requested resource of each request in the testing data set as prior knowledge to generate nonperiodic personalized resources by matching the requested resource against the discovered association rules. Our periodic web personalization approach thus has the advantage that it does not require the user's current access information, which makes it possible to perform more compute-intensive personalized resource preparation in advance rather than in real time.

Fig. 7 shows the comparisons of −E, +E1, +E2 and the Non periodic approach [10] based on satisfaction

for the Top 10, Top 20, and Top 50 users with 0.5 and d =4.0. For example, for Top 50 and $N_{PR} = 4$, the satisfaction values for −E = 86.5%, +E1 = 87.3%, +E2 = 88.1% when d = 0.5, and −E = 86.5%, +E1 = 87.33%, +E2 = 89.8% when d = 4.0. The results again show the usefulness of emotional influence, especially +E2. Further, although −E seemed invariant to increasing d (as mentioned in Section A),satisfaction results did improve with increasing d with emotional influence.

(a) Top 10, $d = 0.5$ (b) Top 20, $d = 0.5$

(c) Top 50, $d = 0.5$ (d) Top 10, $d = 4.0$

(e) Top 20, $d = 4.0$ (f) Top 50, $d = 4.0$

Fig. 7. Comparisons based on satisfaction.

V. CONCLUSION

The presented an approach for automatic generation of Personal Web Usage Ontology of periodic access patterns from web usage logs that have been semantically enriched with information on emotional influence and resource topics. Over time the knowledge base can capture both consumer web access behavior and emotional influence of the web resources on the user. Experimental results both from objective and subjective tests have demonstrated the effectiveness of our approach in providing the user with periodic web personalization based on periodic access patterns generated. Further by varying the degrees of emotional influence we found that emotional influence contributed positively to the results. With PWUO consumers periodic access behaviors can be used by software agents to provide Semantic Web services such as web personalization and semantic search. For example personalized agents can recommend related web contents highlight important links to make navigation more effective and find interesting information from multiple websites and organize them as a personal webpage to a user. Similarly search engines can refine user queries and the search results based on the periodic behaviors discovered from PWUO. Moreover special interest groups can be formed based on users who have shared similar patterns and interests based on PWUO.
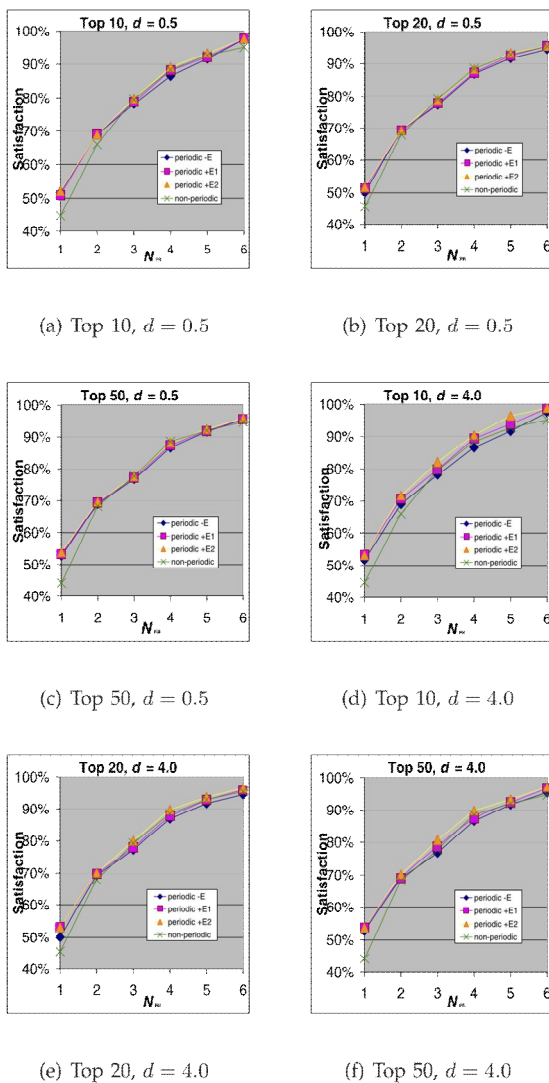
REFERENCES

[1] R.J. Dolan, "Emotion, Cognition, and Behavior," Science, vol. 298, no. 5596, pp. 1191-1194, 2002.

[2] P. Weinberg and W. Gottwald, "Impulsive Consumer Buying as a Result of Emotion," J. Business Research, vol. 10, no. 1, pp. 43-57, 1982.

[3] F. Piron, "A Comparison of Emotional Reactions Experienced by Planned, Unplanned and Impulse Purchasers," Advances in Consumer Research, vol. 20, nos. 341-344, pp. 199-204, 1993.

[4] J. Kalbach, Designing Web Navigation. O'Reilly, 2007.

[5] C. Miao, Q. Yang, H. Fang, and A. Goh, "A Cognitive Approach for Agent-Based Personalized Recommendation," Knowledge-Based Systems, vol. 20, pp. 397-405, 2007.

[6] D.H. Choi and B.S. Ahn, "Eliciting Customer Preferences for Products from Navigation Behavior on the Web: A Multicriteria Decision Approach with Implicit Feedback," IEEE Trans. Systems, Man, and Cybernetics, Part A, vol. 39, no. 4, pp. 880-889, July 2009. [7] J.H. Janssen, J.N. Bailenson, W.A. IJsselsteijn, and J.H.D.M. Westerin, "Intimate Heartbeats: Opportunities for Affective Communication Technology," IEEE Trans. Affective Computing, vol. 1, no. 2, pp. 72-80, July-Dec. 2010.

[8] J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data," ACM SIGKDD Explorations, vol. 1, no. 2, pp. 1223, 2000.

[9] "WebTrends Web Analytics,"http://www.netiq.com/webtrends, 2011.

[10] B. Mobasher, H. Dai, T. Luo, and M. Nakagawa, "Effective Personalization Based on Association Rule Discovery from Web Usage Data," Proc. Third Int'l Workshop Web Information and Data Management, pp. 9-15, 2001.

[11] J. Pei, J. Han, B. Mortazavi-Asl, and H. Zhu, "Mining Access Patterns Efficiently from Web Logs," Proc. Fourth Pacific-Asia

Conf. Knowledge Discovery and Data Mining, Current Issues and New Applications, pp. 396-407, 2000.

[12] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific Am., vol. 284, no. 5, pp. 34-43, May 2001.

[13] G. Stumme, B. Berendt, and A. Hotho, "Usage Mining for and on the Semantic Web," Proc. NSF Workshop Next Generation Data Mining, pp. 77-86, Nov. 2002.

[14] S. Middleton, N. Shadbolt, and B.D.D. Roure, "Capturing Interest through Inference and Visualization: Ontological User Profiling in Recommender Systems," Proc. Int'l Conf. Knowledge Capture, pp. 62-69, 2003.

[15] R. Missaoui, P. Valtchev, C. Djeraba, and M. Adda, "Toward Recommendation Based on Ontology-Powered Web-Usage Mining," IEEE Internet Computing, vol. 11, no. 4, pp. 45-52, July/ Aug. 2007.

[16] A. Sieg, B. Mobasher, and R. Burke, "Web Search Personalization with Ontological User Profiles," Proc. 16th ACM Conf. Conf. Information and Knowledge Management, pp. 525-534, 2007.

[17] A. Maedche and S. Staab, "Ontology Learning for the Semantic Web," IEEE Intelligent Systems, vol. 16, no. 2, pp. 72-79, Mar./ Apr. 2001.

[18] A. Todirascu, F. De Beuvron, D. Gaˆlea, and F. Rousselot, "Using Description Logics for Ontology Extraction," Proc. First Workshop Ontology Learning at the 14th European Conf. Artificial Intelligence, Aug. 2000.