AN APPROACH TO PERSONALISATION IN E-LEARNING SOCIAL ENVIRONMENTS

Hend Ben Hadji and Ho-Jin Choi
Korean Advanced Institute of Science and Technologies, Daejon, South Korea

Keywords: E-learning, Social tagging, Contextualization, Formal concept analysis, Similarity-based concept approximation.

Abstract: We present an approach to a personalized learning service that takes advantage of social tagging to support social learning in context and provides learning resources adapted to the abilities and needs of an individual learner. We employ graph clustering technique in order to group tags into clusters having different contexts, learner's personomy to discover the learning context of the targeted learner, and Formal Concept hierarchy theory to hierarchically cluster learning resources and retrieve the relevant resources to the learner's needs.

1 INTRODUCTION

With the increasing growth of learning resources, personalized support has become an essential component of e-learning systems. Personalized learning service aims at providing learning resources adapted to the abilities and needs of an individual learner. Several approaches in this direction have been investigated. Some approaches employ Semantic Web technologies to create semantic-rich e-learning systems. They mostly rely on ontologies to capture the semantics of the entire e-learning process, including learning content, learner's characteristics, and learning context. Other approaches propose social recommender systems for recommending learning resources without the need of understanding the semantics of the learning content or the learners' need. While other approaches combine both technologies for better personalization purpose.

Although all the diversity of approaches, e-learning systems have not reached their full potential in practice. This is because e-learning process tends to undertake continuous changes over time, such as changes to learning content or learner's behavior. Such changes are unfortunately not well supported by prior approaches. (Vassileva, 2008) studied the behaviour of the new generation of learners and found that learners mostly learn by accessing to social resource sharing systems, such as YouTube or Del.cio.us, to find information, video, or any related materials of interest. (Forte and Bruckman, 2008) state that learning happens by consuming and producing knowledge and provide the example of the collaborative writing of Wiki articles as a valuable learning experience. (Bateman, Brooks, and McCall, 2006) propose a working prototype which illustrate how socially constructed knowledge can support domain experts in defining the semantics of the learning process. Similarly, (Al-Kalifa and davis, 2007) demonstrate that metadata generated using social tagging, core product of the social resource sharing systems, is better than the metadata created by human expert annotation in terms of search and contextual coverage. Of all these works, no prior work on learning personalization over Social Web technologies has been explored yet. The main approach being taking, here, focuses on leveraging social tagging to support social learning in context. We propose a personalized learning service which can be built over any social resource sharing systems and provides personalization capabilities.

The paper is structured as follows. Section 2 presents our motivation through a simple scenario of learning over the Del.cio.us bookmarking system. Section 3 details the phases of our learning personalization process. Section 4 concludes the paper and outlines the future work.

2 USE CASE SCENARIO

In this section we describe a simple e-learning scenario over the social bookmarking system,

Ben Hadji H. and Choi H. (2010).
AN APPROACH TO PERSONALISATION IN E-LEARNING SOCIAL ENVIRONMENTS.
In *Proceedings of the 2nd International Conference on Agents and Artificial Intelligence - Agents*, pages 351-354
DOI: 10.5220/0002728103510354

Del.cio.us to motivate our approach. Consider a student, *Tom*, is taking a Web Services course. Tom has been given an assignment, which consists of implementing a Web Service client using the SOAP technology. As Tom wants learning by interacting with other people and sharing knowledge about topics of particular interest, he becomes an active member of Del.cio.us; he uploads learning resources that he is interested in and assigns them individual tags for future retrieval. Further, Tom often navigates in people's space with the same interest looking for relevant learning resources for making assignments or reviewing exams.

As the users' number accessing to Del.cio.us increasingly grows, Tom finds it difficult this time to find relevant resources about SOAP technology. Therefore, he decides to issue a query looking for learning resources which are tagged with "SOAP". As a result, thousand of resources that were annotated with a SOAP tag are returned to Tom's query. Further, not all the documents are relevant to the learner's context. For example, a document entitled "SOAP maker" was retrieved, and it is about a shopping space for soap maker. This means that SOAP tag has been used by users to refer to different concepts in different contexts.

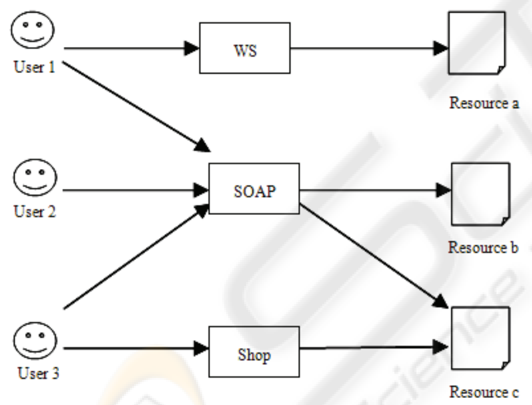

Figure 1: Social Tagging Mechanism.

Figure 1 illustrates such a situation where SOAP tag is used by user 1 and user 2 to refer to documents about Web Service, while used by user 3 to refer to Shopping context. To this end, we can say that the Del.cio.us bookmarking system does not implement an intelligent retrieval engine that tracks the user's behavior and returns relevant resources to his/her context. This limitation constitute the main barriers of adopting social resource sharing systems as learning environments or even for implementing social resource sharing system as a service of e-learning systems. (Vassileva, 2008) believes that the new generation of e-learning systems need to support social learning in context, that is, support the learner to find the right content and people while offering social environments. In this context, we propose a social learning personalization mechanism that supports social learning and provides personalization capabilities. Detailed description of the process is in section 3.

3 PROPOSED ARCHITECTURE

Figure 2 presents the personalization process over social learning environments. This process is composed of three main phases: *pre-processing phase, contextualization phase*, and *retrieval phase*.

3.1 Preprocessing Phase

This phase starts by collecting the learning resources from the delicious website, number of users who tagged the learning resources, and list of all tags assigned to the learning resources. Once folksonomies are collected, they are passed to the preprocessing module which performs a series of filters for cleaning the tags and removing ambiguous and infrequent tags. Fist, tags are converted to lower cases so that string manipulation can be performed. Stop words and Non-English words are then removed to ensure that only English tags are present when recommending learning resources to the learner. Finally, infrequent and isolated tags are filtered out. At the end of the preprocessing phase, we get concise tags, which are ready to be used in the subsequent phases of the personalization process.

3.2 Contextualization Phase

This phase consists of two steps: (1) finding the different contexts that might be related to the learner's query, and (2) identifying the suitable context to the targeted learner. Let t be the tag query and F' the subset of the folksonomy that is associated with that particular tag query defined as a tuple $F' := (L_t, R_t, T_t, A)$, where L_t is the set of learners who have used the tag t on one or more learning resources: $L_t = \{l \mid \exists r \in R, (l, t, r) \in A\}$; R_t the set of learning resources which have been annotated with the tag t: $R_t = \{r \mid \exists l \in L, (l, t, r) \in A\}$; and T_t is the set of tags which have been used together with t on some resources: $T_t = \{t' \mid \exists (l, r) \in L \times R, (l, t, r) \in A \land (l, t', r) \in A\}$. Having identified the set of tags, T_t, a weighted tag vector, v_r, can be constructed to represent a resource r, whose elements correspond to the number of times a tag has been assigned to it: $v_r = (v_{r1}, v_{r2}, v_{r3}, \ldots v_{r|Tt|})$. A similarity matrix $A = \{a_{ij}\}$ is constructed to represent the pairwise similarity of each resource by using cosine similarity measure:

$$a_{ij} = \text{cos\_similarity}(v_r, v_d) \qquad (1)$$

A bipartite graph can be then constructed, where $|L_t|$ vertices representing each of the resources and edges weighted by the similarity between these resources. This resource-based network might display the tags used by the learner to refer to different concepts in different contexts. For example, a learner who is looking for resources annotated with t = "*SOAP*" in *Web Service* context might get resources referring to *Shopping* context. Therefore, we should first discover the different contexts in which t can be used and then find out the context that matches the learner's profile.

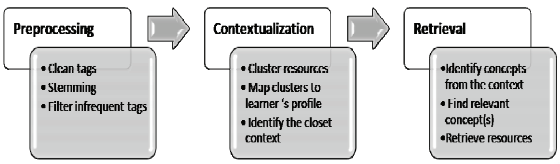

Figure 2: Pipeline Process of Social Learning Personalization.

3.2.1 Graph Clustering

The process of finding out the different contexts in which a tag query might be used is called *Tag Contextualization*, term coined by (Yeung, Gibbins, and al., 2009). This process consists of applying the fast greedy algorithm to identify groups of vertices in a network which are highly connected with those in the same group of vertices but loosely connected with those in other groups. The result of the tag contextualization process is a set of k clusters: $C = \{C_i | 1 \leq i \leq k\}$, where C_i is a set of resources clustered together to refer to a particular context.

Formally, each context is denoted as a triple $\kappa_i = (R_{\kappa i}, T_{\kappa i}, I)$, where $R_{\kappa i}$ is the set of learning resources in a cluster C_i: $R_{\kappa i} = \{r_i | i=0, 1, \ldots n \; \& \; n \leq |R_t|\}$, $T_{\kappa i}$ is the set of tags used within a cluster C_i: $T_{\kappa i} = \{t_j | j=0, 1, \ldots m \; \& \; m \leq |T_t|\}$ and $I \subseteq R_{\kappa i} \times T_{\kappa i}$ is a binary relation defined between $R_{\kappa i}$ and $T_{\kappa i}$, i.e., incidence matrix. Following FCA theory, the elements of $R_{\kappa i}$ are treated as objects, those of $T_{\kappa i}$ as the attributes of the objects collection. The relationship between objects and attributes is represented as cross tables, whose rows are headed by the objects, whose columns are headed by the attributes and whose cells are marked if the incidence relation holds for the corresponding pair of object and attribute.

3.2.2 Mapping Clusters to Learner's Profile

Having identified the different contexts used with the tag t, the next step is to find out the appropriate context for the learner's profile. The latter can be obtained from the folksonomy which are associated with that learner. This subset of the folksonomy is called personomy. A personomy of a learner is characterized as a tuple $P_l := (R_l, T_l, A_l)$, where R_l is the learner's set of learning resources: $R_l = \{r \mid (t, r) \in A_l\}$, T_l is the learner's set of tags: $T_l = \{t | (t, r) \in A_l\}$, and A_l is the set of the tags of the user: $A_l = \{(t, r) | (l, t, r) \in A\}$. The learner profile can be then described by two main set of elements, the resources tagged and their associated tags. Finding the appropriate context for the learner's profile comes then to find the similarity degree between his personomy P_l and each context κ_i. A modified Jaccard Index is used for similarity measure, defined as follows:

$$Sim(P_l, K_j) = \alpha \frac{|R_l \cap R_{K_i}|}{|R_l \cup R_{K_i}|} + (l - \alpha) \frac{|T_l \cap T_{K_i}|}{|T_l \cup T_{K_i}|} \qquad (2)$$

As shown in the above Equation, the $Sim(P_l, \kappa_i)$ function takes into account two factors, the tags and resources. The similarity between tags reveals the similarity in topics, while similarity between documents reflects similarity in content. α is a threshold, ranged between 0 and 1, which can be set by the learner for flexibility. The greater similarity value is the closet context to the learner's profile is.

3.3 Retrieval Phase

Given the appropriate context of the learner, say $\kappa = (R_\kappa, T_\kappa, I)$, we need to find out the relevant concept(s) in which the learner is interested. For discovering the different concepts of a given context and their relationships, we propose to use FCA method because of its simplicity and effectiveness.

3.3.1 Tags Conceptualization

Central to FCA is the notion of *formal concepts*. A formal concept of formal context κ is defined as a pair (R_j, T_j), where $R_j \subseteq R_\kappa$, $T_j \subseteq T_\kappa$, $R_j' = T_j$, and $R_j' = T_j'$. R_j' is a set of attributes common to the objects in R_j which is defined as $R_j' = \{t \in T_\kappa \mid \forall r \in R_j : (r, t) \in I\}$ and T_j' is the set of objects commonly have the attributes in T_j which is defined as $T_j' = \{r \in R_\kappa \mid$

$\forall$ $t \in T_j$: $(r, t) \in I\}$. Accordingly, (R_j, T_j) is a formal concept if the set of tags shared by the learning resources in R_j is identical with T_j and on the other hand R_j is also the set of all learning resources in the collection having all attributes in T_j. R_j is then called the extent and T_j the intent of the formal concept (R_j, T_j). For the sake of simplification, the formal concept (R_j, T_j) is denoted C_j, T_j is Intent(C_j), and R_j is the Extent(C_j).

3.3.2 Concept Approximation

The final step consists of finding out the closet conceptual cluster to the targeted learner and retrieving the relevant learning resources belonging to it. Learning resources are retrieved here using *Concept approximation* method, proposed by (Saquer and Deogun 2001). This method consists of computing the similarity degree between a set of attributes (i.e., tags added by the learner in our case) to be approximated and the formal concepts on the given context, and select the most similar formal concept approximation result. The similarity measure is defined as (Saquer, Deogun, 2001):

$$f_C(B) = \frac{\dfrac{|B \cap \text{Intent}(C)|}{|B \cup \text{Intent}(C)|} + \dfrac{|\alpha(B) \cap \text{Extent}(C)|}{|\alpha(B) \cup \text{Extent}(C)|}}{2} \quad (3)$$

The range of $f_C(*)$ is the interval [0,1]. For the attribute set T_i, $f_C(T_j) = 0$ when T_j and $\alpha(T_j)$ are disjoint from the intent and extent of C, respectively. $F_C(T_j) = 1$ when T_j = Intent(C) and, therefore, $\alpha(T_j)$ = Extent (C). In general the closer the value $f_C(T_j)$ to 1, the greater the similarity between T_j and the intent of C. To approximate a set of tags S, we find a formal concept C that maximizes the value of $f_C(S)$. In case more than one formal concept are found to approximate S with same value of f_C, we say that these concepts equally approximate S.

4 CONCLUSIONS

In the light of the advances of the Social Web technologies, we believe that there is a potential for enhancement of e-learning systems. In this paper, we present our approach to social learning in context. This approach can be applied to any social resource sharing systems which store learning resources and enable the production of broad folksonomies, but also to any e-learning systems which integrate social tagging service. The future work consists of implementing the whole process over the bookmarking system Del.cio.us and comparing results with the classical retrieval approach.

ACKNOWLEDGEMENTS

This research was supported by the MKE(Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2009-(C1090-0902-0032)).

REFERENCES

Al-Khalifa, H.S., Davis, H.C. (2007). Replacing the Monolithic LOM: A folksonomic Approach. In *ICALT'07, 7th IEEE International Conference on Advanced Learning Technologies*.

Bateman, S., Brooks, C.A., and McCall, G. Collaborative Tagging Approaches for Ontological Metadata in Adaptive Learning Systems. *In Proc. SWEL at AH'06.* (2006) 3-12.

Yeung, C.A., Gibbins, N., Shadbolt, N. (2009). Contextualising tags in collaborative tagging systems. *In the Proceedings of the 20th ACM Conference on Hypertext and hypermedia*, Italy.

Forte, A., Bruckman, A. (2008). Scaling consensus: increasing decentralization in Wikipedia governance. *In the Proceedings of Hawaiian International Conference of Systems Sciences (HICSS)*.

Mika, P. (2005). Ontologies Are Us: A Unified Model of Social Networks and Semantics. *In Proc. ISWC 2005.*

Saquer, J., Deogun, J. (2001). Concept approximations based on rough sets and similarity measures. *Journal of Applied Mathematics and Computer Science*, 11 (2001), pp. 655–674.

Vassileva J., (2008), "Towards Social Learning Environments". *IEEE Transactions on Learning Technologies*, Vol. 1 No. 4, October-December, pp. 199-214, IEEE CS & ES.