



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Faculty of Engineering and Information Sciences -
Papers: Part B

Faculty of Engineering and Information Sciences

2017

Ontological Learner Profile Identification for Cold Start Problem in Micro Learning Resources Delivery

Geng Sun

University of Wollongong, gs147@uowmail.edu.au

Tingru Cui

University of Wollongong, tingru@uow.edu.au

Jun Shen

University of Wollongong, jshen@uow.edu.au

Dongming Xu

University of Queensland, d.xu@business.uq.edu.au

Ghassan Beydoun

University of Wollongong, beydoun@uow.edu.au

See next page for additional authors

Publication Details

Sun, G., Cui, T., Shen, J., Xu, D., Beydoun, G. & Chen, S. (2017). Ontological Learner Profile Identification for Cold Start Problem in Micro Learning Resources Delivery. IEEE International Conference on Advanced Learning Technologies (ICALT 2017) (pp. 1-5). United States: IEEE.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Ontological Learner Profile Identification for Cold Start Problem in Micro Learning Resources Delivery

Keywords

start, problem, micro, learning, resources, delivery, ontological, profile, learner, identification, cold

Disciplines

Engineering | Science and Technology Studies

Publication Details

Sun, G., Cui, T., Shen, J., Xu, D., Beydoun, G. & Chen, S. (2017). Ontological Learner Profile Identification for Cold Start Problem in Micro Learning Resources Delivery. IEEE International Conference on Advanced Learning Technologies (ICALT 2017) (pp. 1-5). United States: IEEE.

Authors

Geng Sun, Tingru Cui, Jun Shen, Dongming Xu, Ghassan Beydoun, and Shiping Chen

Ontological Learner Profile Identification for Cold Start Problem in Micro Learning Resources Delivery

Geng Sun¹, Tingru Cui¹, Jun Shen¹, Dongming Xu², Ghassan Beydoun³, Shiping Chen⁴

1, School of Computing and Information Technology, University of Wollongong, Wollongong, Australia

2, UQ Business School, The University of Queensland, Brisbane, Australia

3, School of Systems, Management and Leadership, University of Technology Sydney, Sydney, Australia

4, Data61, CSIRO, Sydney, Australia

Email: gs147@uowmail.edu.au, tingru@uow.edu.au, jshen@uow.edu.au, D.Xu@business.uq.edu.au,

Ghassan.Beydoun@uts.edu.au, Shiping.Chen@data61.csiro.au

Abstract—Open learning is a rising trend in the educational sector and it attracts millions of learners to be engaged to enjoy massive latest and free open education resources (OERs). Through the use of mobile devices, open learning is often carried out in a micro learning mode, where each unit of learning activity is commonly shorter than 15 minutes. Learners are often at a loss in the process of choosing OER leading to their long term objectives and short term demands. Our pilot work, namely MLaaS, proposed a smart system to deliver personalized OER with micro learning to satisfy their real-time needs, while its decision-making process is scarcely supported due to the lack of historical data. Inspired by this, MLaaS now embeds a new solution to tackle the cold start problem, by opening up a brand new profile for each learner and delivering them the first resources in their fresh start learning journey. In this paper, we also propose an ontology-based mechanism for learning prediction and recommendation.

Keywords—Micro Learning; Open Education Resources; Ontology; Cold Start Problem

I. INTRODUCTION

In the information age, the development and dissemination of learning resources are booming in a much faster speed and wider range than their traditional shapes. People have shown increasing willingness and interests in getting access to online learning resources and getting involved in online learning activities via electronic, especially mobile devices [1]. On the other hand, the accessibility of learning resources is obviously expanded along with more and more universities having opened up their courseware, lecture notes, audios, videos, reading materials and other associated learning resources to the public [2]. This leads to an emerging trend, open learning, which attracts substantial attention by both IT practitioners and investors and obtains popularity in the educational sector, particularly in tertiary and vocational sections.

Nevertheless, the effectiveness of OER delivery to e-learners in diverse learning environments remains a challenge to the success of open learning. Prior studies indicate that open learning is currently suffering from low completion rate [3]. Some researchers believe that this is because learners often feel difficult to get the exact learning resources, what they want promptly and instantly [3] [4].

In our previous work, we proposed a system, Micro Learning as a Service (MLaaS), to deal with the personalized

learning resource tailoring and delivery, oriented to the micro learning in open learning environment [5]. In this paper, we introduce an ontological approach to technically support MLaaS to make the first decision in learning resource adaptation. Innovatively, this approach addresses the cold start problem in micro learning through OER, targeting at freshly joined learners in brand new educational settings. This approach relies on semantic learner profile identification by taking into account learners' similarities, and predicts learners' personal factors, which can affect their micro learning experiences.

II. BACKGROUND

A. Micro Learning

Micro learning refers to short-term learning activities on small learning units. From educational viewpoint, the learning period in micro learning usually covers a time span from a few seconds (e.g. in mobile learning) to up to 15 minutes. In the contemporary Web era, micro learning pertains to small pieces of knowledge based on pervasive Web resources [6]. With mobile devices, learners are able to accomplish learning mission in a short time period. As a completely new derivation of mobile learning, micro learning is generally carried out through mobile means and it shares some similar characteristics with mobile learning as both of them are individually referable, self-contained, reusable and re-mixable [7]. Micro learning becomes a major learning channel in mobile environments and is becoming a mainstream for next generation learners, who learn on the move, with easy access to the 'cloud' or Internet of Things [8]. Compared to traditional learning modes, now learners' overall efforts to go through an entire concept will proceed in a continuous, or even intermittent, way rather than a consecutive way.

B. Cold Start Problem in Learning Resource Recommendation

In computer science literature, widely used adaptive recommendation methods generally consist of two main categories, i.e., memory-based and model-based algorithms [9]. Although they have been found in many successful cases of recommender systems, for example, Amazon online store, it is usually difficult to provide reliable recommendations due to the insufficiency of initial data of ratings or preferences.

This leads to the occurrence of cold start problem. Commonly the cold start problem is triggered by three factors: new community, new item and new users.

The cold start problem becomes more severe in the open learning, especially in micro learning through OERs [10]. Both open learning and OER are relatively new products which are emerging in the very recent years. Meanwhile, the followers of this novel trend, no matter new education pursuers or regular learners migrated from other online learning modes, are forming as a completely new community. On the other hand, the learning demands and expectations of learners engaged in open learning are much more practical than conventional university students. In other words, they are mostly self-regulated so that it is totally flexible for them to decide when to join or quit the online course at their own willingness, and switch among courses frequently [4]. Consequently, for OER providers, it is difficult to establish a model and update it accordingly for any individual learner because they do not have historical data in hand.

In micro open learning, or micro learning over OERs, it is very normal to find that learners take part in and deviate from the learning scenarios frequently, as well as turning on and off the learning activities at their own willingness. That is to say, the overall situations of micro learning vary all the way, from individual to individual. Moreover, it is very common that freshmen join into open learning or existing learners unfold a brand new course learning profile, at any time. All in all, there are a large number of new learners in open learning; and a new learner usually initiates access of new learning resources; and learners who went through learning resources in a same branch of a discipline will form as a new community.

C. Challenges in Cold Start Problem

If treated carelessly, the cold start problem may lead to the loss of learners who are previously engaged in open learning and then decide to stop using the OER delivery system or adopting the learning mode [11]. The reasons behind the situation are mainly due to the lack of accuracy in the recommendations received in that first stage, in which the learners have not yet cast a significant number of votes or rating to feed the recommender systems. Basically, the sparsity of data affects the user satisfaction and then it can further affect the user acceptance of the new open learning mode.

In this paper we will primarily focus on the cold start problem of new learners launching micro open learning, by constructing learner profile from very limited information. In addition, we will use the terms ‘micro learning through OERs’ and ‘micro open learning’ interchangeably.

III. ONTOLOGY CONSTRUCTION

A. MLaaS and Ontology Construction

In our previous work, we proposed a system, MLaaS, to deal with the adaptive micro OER delivery problem [5]. It aimed to deliver learners adaptive micro learning resources in terms of their time availabilities. It took into account the specialties

of the micro learning environment as well as the learners’ ‘highly-mobile’ learning behaviors. MLaaS was designed by making use of the semantics approach. We combined a pattern and rule discovery process of micro learning, together with a survey of education literature for the features that could affect learning experience and outcomes in mobile environment [12]. Consequently, we made use of the ontology to build conceptual graphs in order to profile the features playing significant roles on an ongoing micro learning process. These graphs also depicted how features were mutually affected and interrelated to each other. According to our design, the profiling procedure is carried out from two sides, the learner side and OER side. To technically operate the semantic learner profile construction for micro learning, data filled in the graphs come from two sides, the explicit data collection (e.g. through mandatory requests) and implicit data tracking (e.g. automatic extraction) [13]. In addition, rather than developing the domain ontology for OERs by ourselves, a general structure of courseware ontology were built jointly by making use of existing ontologies, which had been extracted from main OER providers, such as universities involved in major open courseware alliances (e.g. participating institutions in edX¹), or from the Linked Open Data Cloud community² [14].

However, as stated in Section 2, the profile construction is impossible with insufficient information about the learner at the commencement of open learning. Therefore, the learner profile cannot be fully identified with valid data. This will be treated as a cold-start problem and tackled by filling in the gaps with predicted data.

B. Augmented Micro OER Ontology

From the item-based view, we will deepen the sights into the micro learning environment in particular and, for this reason, the general ontology of OER is augmented to adapt the needs of micro learning.

In the augmented micro OER ontology, an annotation of a micro OER is self-describing with metadata exploring its educational parameters, such as typology (video, audio, text, etc.), type of interaction (expositive, active, mixed, two-way), didactic model (e.g. inductive, deductive, learning by doing, etc.), semantic density and so on [15]. Each node in the augmented OER ontology indicates a micro OER chunk. A chunk is the smallest unit in the micro learning settings, normally a fine-cut piece of an OER from its provider, and it has an apparently shorter time length (preferably less than 15 mins) than its original shape. It can be a mini concept or knowledge point, tinier than what the teachers used to deliver; or it can be a cut of course video or lecture notes, or a course settings come along with a concept, such as assessment, task, reading material and so on [16].

There is not any totally independent chunk and each of them is part of a relational web rather than merely a conceptual object [13]. This ontology is used to explicitly classify the OERs to recommend among a pedagogically defined set of distinctive main concepts, fed as the raw material in the reasoning process of MLaaS [12][13].

¹ <https://www.edx.org/schools-partners>

² <http://lod-cloud.net/>

C. Micro Learning Learner Profile Taxonomy

Adopting ontologies as the basis of the learner profile is crucial in addressing the cold start problem in micro OER delivery. It allows the initial learner behavior to be matched with existing and pre-known knowledge in the ontologies and relationships among them.

From the user-based view, a main ontology, on which all learner profiles are based, is named as the *Benchmark* ontology, where the element *Learner* is put at the center of the graph [12]. Acting as the instances of the presetting domain ontologies, a specific learner profile oriented to micro learning is a set of nodes from the *Benchmark* ontology versus a node in the augmented micro OER ontology. It contains plenty of annotations in terms of their learning behaviors and context.

The learner profile is managed by MLaaS by two parts: static part and dynamic part. The static part can be represented by a vector, which contains the demographic and educational information. By matching these two augmented ontologies, respectively for item and user, the dynamic part of a learner node is denoted as a pair, $L_j = \{MR_u, ML_j\}$, $L_j \in L$. Herein, the element MR_u denotes the u^{th} micro OER, as introduced in Section C, which is a particular version of the micro OER ontology, and a three dimensional element $ML_j\{P_{u,j}, TA_j, D_j\}$ is exclusive to j^{th} learner during the micro learning process. Herein, the element $P_{u,j}$ indicates the learner's preference, TA_j indicates the j^{th} learner's instantly time availability, and D_j denotes the level of distraction in terms of the given learning environment and surroundings.

Whenever MLaaS gathers any information from the learner's learning process over OER, the learner profile will be updated in regards to ML_j .

IV. ONTOLOGICAL SOLUTION FOR COLD START PROBLEM

A. Preference Propagation

Provided the cold start condition for the first micro OER delivery, a learner is required to quickly mark down a preference on a specific micro OER. Consequently, a spreading activation approach is applied to maintain the preference against its parent node (i.e., the R_v is the v^{th} OER where the MR_u derived from) as well as updating learner profile. It propagates the learners' preference upwards the hierarchy of micro OER ontology based on activation values. In other words, the preference has been obtained from a micro OER to its ancestor and spread in its superclass (i.e., OER) level. An example of the spreading activation is shown in Figure 1.

A partial view of augmented micro OER ontology in 'information technologies' area is shown in Figure 1. Particularly, it describes an 'e-business' OER from an Australia provider, OpenLearning³. At the bottom level of the ontology, nodes which are depicted with oval shape typically

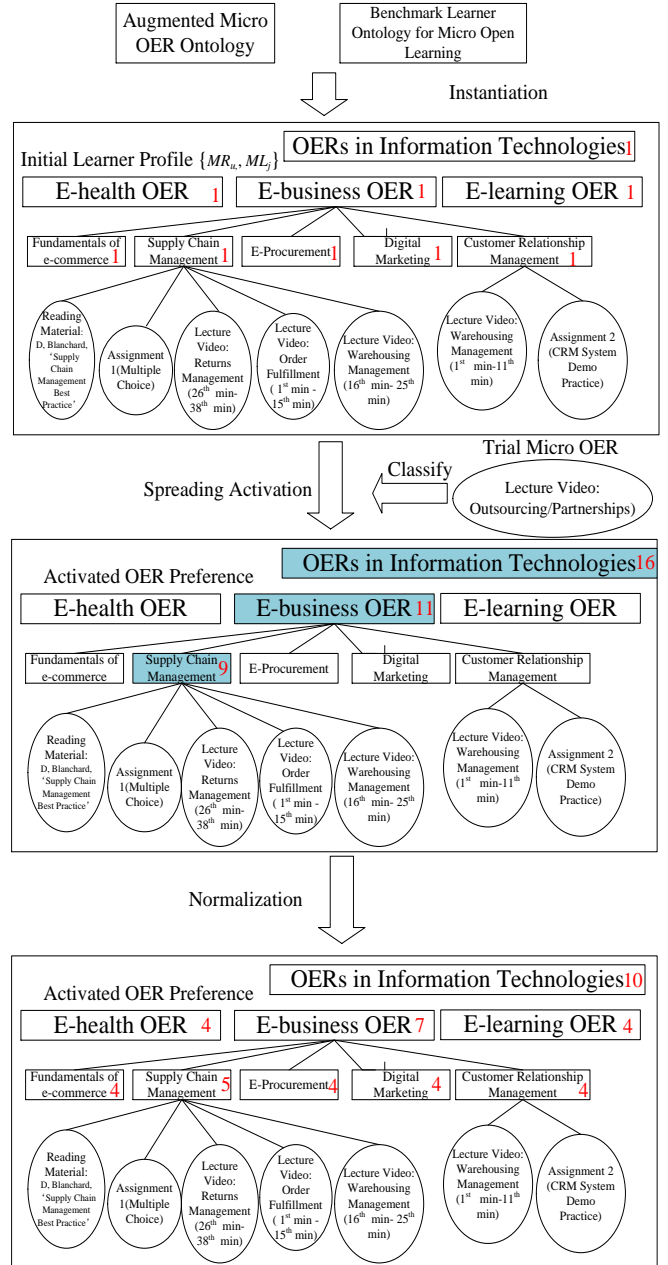


Figure 1. Partial View of the Augmented Micro OER Ontology and Spreading Activation for a Learner's Preference on OER conform to the standard of micro OER. Red integers shown in nodes with rectangle shape are preference values from a learner versus target OERs. The algorithm 1 is proposed to execute the process of preference propagation.

Algorithm1: Preference Propagation

Input: Dynamic part of learner profile $L_j = \{MR_u, ML_j\}$, a trial micro OER MR_u , $L_j \in L$

Output: Updated dynamic part of learner profile with updated $P_{u,j}$ value in the triple dimensional set ML

$P(R_v)$ and $Activation(R_v)$, preference value and activation value for the OER R_v

³ <https://www.openlearning.com/>

```

//Step 1: Spreading Activation
begin: Initialize PriorityQueue; //PriorityQueue is the set of OERs
      within the same discipline where  $R_v$  belongs to
      Set Activation of all micro OER to 0
for each  $L_j \in L$  do
    if ( $MR_u \in R_v$ ) then
      Activation( $R_v$ ) =  $P(R_v)$ 
      PriorityQueue.Add( $R_v$ )
    end if
end for
while PriorityQueue.Count > 0 do
  Sort PriorityQueue; //activation values in descending order
  Select the first item  $MR_u$  in PriorityQueue //  $R_v$  with highest
  P value
  Remove  $R_v$  in PriorityQueue
  for each  $R_v$  do
    LinkedOERs = GetLinkedOERs( $R_v$ ) //get linked nodes of  $R_v$ 
    for each  $R_w$  in LinkedOERs do //propagate activation to its
    neighbors
      Activation( $R_w$ ) += Activation( $R_v$ ) * Weight( $R_v, R_w$ )
      PriorityQueue.Add( $R_w$ )
    end for
  end for
end while
//Step2: Learner Profile Normalization
for each  $L_j \in L$  do
   $P(R_v) = P(R_v) + \text{Activation}(R_v)$ 
   $k = 1 / \sqrt{\sum_{i=1}^v (R_v)^2}$  //normalization factor
   $P(R_v) = P(R_v) * k$  //normalization
end for
end

```

The normalization factor acts on preventing the propagated preferences from escalating continuously to such an extent that exceeds a reasonable range, which could result in difficulty of data processing in the forthcoming process. The confidence degree for the propagated preference of OER is recorded as $CD(R_v)$.

B. Instant Time Availability

The system is able to obtain explicit information on how long the learner can (or would like) to spend on a micro OER through mobile devices in the real time. As a mandatory request, a learner is required to input his instant time availability at the beginning of every micro learning activity.

According to the system setting, suggestively the instant time availability, TA_j , is represented by an integer from 1 to 15. However, if the learner is not pretty sure how long he is able to spend on the micro OER at once, he is free to leave a time span, which can be continuous integers in the same range.

The confidence degree, $CD(TA_j)$, of the TA_j is associated with its time span, $CD(TA_j) = \{1/|\text{time span}|\}$.

C. Distraction Prediction

1) Demographic Classification

In [16], we have discussed the main issues that might cause distraction in micro open learning, which generally came from two sides, the social side and environmental side.

In addition, MLaaS investigates existing learners' degree of distraction as reference, and senses every learner's location information through built-in functions in mobile devices. Based on the given taxonomy and augmented ontology, we carry out a demographic classification that aims to cluster learners into cohorts, in order to match them with micro-pieces of OERs.

The mechanism of classification is designed as, learners who have similar static information, involving employment and/or education background, occupation, and similar learning environment/location, are more likely to face similar level of distraction. For the same reason, their overall time availabilities would more likely fall in the same range. Herein MLaaS tries to associate a learner into a pre-clustered learner group, by applying the stereotyping technique to fulfill the requirement of demographic classification. Combined the approaches introduced in [17] and [18], a one-against-all model and a Bayes classifier are built to realize the demographic classification.

2) Similarity Measure between Two Learners

Once new learners join into the open learning scenario, MLaaS responds immediately to classify them into clusters with similar static and location information. MLaaS is responsible to find the similar existing learners, so as to recommend them micro OERs that were recognized as suitable to learn in a given time span, situation and environment.

Their learning location information is sensed from the location service embedded in the mobile devices. Thus, the similarity between two learners L_i and L_j is evaluated using the equation (1).

$$\text{sim}(i, j) = [(\sum_{l=1}^m S_l W_l)^2 + (SLo_{i,j} W_{i,j})^2]^{1/2} \quad (1)$$

where S_l is the similarity value of the l^{th} attribute in the static part of learner profile and the W_l is its corresponding weight. $SLo_{i,j}$ denotes their similarity on location and $W_{i,j}$ denotes the weight for location factor.

3) Distraction Prediction

Thus, in terms of the equation (2), the distraction value can be estimated in accordance with the action that any member in a same cluster indicates the distraction level.

$$D_{i,Lo_a} = \frac{\sum \text{sim}(i, j) \cdot d_{j,Lo_a}}{\sum \text{sim}(i, j)} \quad (2)$$

where d_{j,Lo_a} is the degree of distraction the learner L_j felt in the Location Lo_a . This follows the expectation that the learners who have similar general situation (i.e. social factors) and surroundings (i.e. environmental factors) are in high probability to have similar degree of distraction.

The confidence degree for the predicted distraction is depicted as $CD(D)$.

D. Integration of Recommendation Results

In section B we have merely obtained the preference of a learner on an 'entire' OER rather than on a micro OER, now the preference values are again propagated downwards the

ontology hierarchy. Consequently, each micro OER node receives an estimated preference value from its ancestor. This propagation process is executed with a decay factor. For each micro OER the final preference value, $P_{u,j}$, can be calculated use the following equation (3).

$$P_{u,j} = \frac{\sum P_R CD(R) + \sum P_{R_v} CD(R_v) D(u,v)}{\sum CD(R) + \sum CD(R_v) D(u,v)} \quad (3)$$

where R is the set of all the nodes in the higher hierarchy than MR_u , R_v is a direct ancestor of node MR_u and $D(u,v)$ is the count of level between MR_u and R_v .

As far as all values of the three attributes, denoting preferences, instant time availability and degree of distraction, in the set ML are settled, a complete learner profile is constructed from the initial little-known information by the MLaaS.

MLaaS consumes the value P and D in conjunction with their TA to compare with the attributions and requirements annotated in the metadata of augmented OER ontology. The next step is to integrate the outcomes from section A , B and C , a fitness function will convert these selected multidimensional arrays into one variable. Hence, this problem is hereby properly transferred to a multi-objective optimization problem. The calculated $CD(TA)$ value will act on the resource screening. In addition, a heuristic algorithm can be employed to infer a suitable micro OER as their first attempt in the novel open learning experience through MLaaS, in which both $CD(TA)$ and $CD(D)$, will play a role in the fitness convergence. Limited by the scope of this paper, we will not elaborate the details of the algorithm of reasoning process here.

Along with the successful launch of solution to the well-known cold start problem in micro learning, learners' upcoming behaviors will be continuously acquired by MLaaS to feed to the reasoning engine.

V. CONCLUSION

In this paper, we introduced an ontological approach towards solving the cold start problem in the delivery of personalized micro OER through mobile devices. We described the system design principle and ontology structure, and we discussed two augmented ontologies exclusive to micro learning through OER. By composing them, we organized a system profile identification from a cold start condition, taking into account a learner's instant time availability, preference and extent of distraction. Hence, a heuristic inference process is able to be set accordingly.

Our future work will include prototyping this ontological approach into its corresponding component in MLaaS. This approach will be evaluated by measuring the prediction accuracy and we will also engage real learners to compare the quality of recommendation.

REFERENCES

[1] J. Hilton, 'Open Educational Resources and College Textbook Choices: a Review of Research on Efficacy and Perceptions', Educational

Technology Research and Development, vol.64, no. 4, pp.573-590, 2016.

[2] J. Hylen, D.V. Damme, F. Mulder and S. D' Antoni, 'Open Educational Resources: Analysis of Responses to the OECD Country Questionnaire', OECD Education Working Papers No. 76, June 2012.

[3] I. Nawrot and A. Doucet, 'Building Engagement for MOOC Students: Introducing Support for Time Management on Online Learning Platforms', Proceeding of WWW'14 Companion, 2014,

[4] S. Miranda, G. R. Mangione, F. Orciuoli, M. Gaeta and V. Loia, 'Automatic Generation of Assessment Objects and Remedial Works for MOOCs', 12th International Conference on Information Technology Based Higher Education and Training, Antalya, Turkey, October 2013.

[5] G. Sun, T. Cui, S. Chen, W. Guo and J. Shen, 'MLaaS: A Cloud Based System for Mobile Micro Learning in MOOC', 4th IEEE International Conference on Mobile Service (MS), New York, US, July 2015, pp.120-127.

[6] D. Kovachev, Y. Cao, R. Klamma and M. Jarke, 'Learn-as-you-go, New Ways of Cloud Based Micro-learning for the Mobile Web', 10th International Conference on Web-based Learning (ICWL), Hongkong, December 2011, Springer LNCS7048 pp.51-61.

[7] T. Hug and M. Lindner, 'ML: Emerging Concepts, Practices and Technologies after e-Learning', Proceedings of Micro Learning 2005, Austria, June 2005, pp. 8-11

[8] M. I. Souza and S. F. D. Amaral, 'Educational Micro Content for Mobile Learning Virtual Environments', Creative Education, vol. 5, pp. 672-681, 2014

[9] A. Sieg, B. Mobasher and R. Burke, 'Ontology-Based Collaborative Recommendation', the 8th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems, Hawaii, USA, June 2010, pp. 20-32.

[10] P.A.Bruck, L.Motiwalla and F. Foerster, 'Mobile Learning with Micro-content: A Framework and Evaluation', 25th Bled eConference, Bled, Slovenia, 2012, pp.527-542.

[11] B. Lika, K. Kolomvatsos, S. Hadjiefthymiades, 'Facing the Cold Start Problem in Recommender Systems', Expert Systems with Applications, vol. 41, no. 4, pp. 2065-2073, 2014.

[12] G. Sun, T. Cui, G. Beydoun, S. Chen and J. Shen, 'Profiling and Supporting Adaptive Micro Learning on Open Education Resources' the 4th International Conference on Advanced Cloud and Big Data (CBD), Chengdu, China, August 2016, pp.158-163.

[13] A. Moreno, A. Valls, D. Isern, L. Marin and J. Borras, 'SigTur/E-Destination: Ontology-Based personalized Recommendation of Tourism and Leisure Activities', Engineering Applications of Artificial Intelligence, vol.26, no.1, pp.633-651, 2013.

[14] N. Capuano, L. Dell'Angelo, F. Orciuoli, 'Ontology Extraction from Existing Educational Content to Improve Personalized e-Learning Experiences', 3rd IEEE International Conference on Semantic Computing, Berkeley, USA, September 2009, pp. 577-582.

[15] S. Miranda and G. Albano, 'Personalized Learning in Mathematics', Journal of E-learning and Knowledge Society, vol. 11, no.1, pp.25-42, 2015.

[16] G. Sun, T. Cui, J. Yong, J. Shen and S. Chen, 'MLaaS: a Cloud-based System for Delivering Adaptive Micro Learning in Mobile MOOC Learning'. IEEE Transactions on Services Computing, online first <http://dx.doi.org/10.1109/TSC.2015.2473854>

[17] J. Milgram, M. Cheriet and M. Sabourin, 'One against One or One against all: Which One is Better for Handwriting Recognition with SVMs', 10th International Workshop on Frontiers in Handwriting Recognition, France, October 2006.

[18] L. A. Dalton and E. R. Dougherty, 'Optimal Classifiers with Minimum Expected Error within a Bayesian Framework-Part II: Properties and Performance Analysis', Pattern Recognition, vol. 46, no. 5, pp. 1288-1300, 2013.

