

User Profiling in a SPOC: A method based on User Video Clickstream Analysis

<https://doi.org/10.3991/ijet.v14i01.9091>

Naima Belarbi^(✉), Nadia Chafiq, Mohammed Talbi, Abdelwahed Namir,
Elhabib Benlahmar
University Hassan II - Casablanca, Morocco
naima.belarbi-etu@etu.univh2c.ma

Abstract—In the present paper, we address to construct a structured user profile in a Small Private Online Course (SPOC) based on user’s video clickstream analysis. We adopt an implicit approach to infer user’s preferences and experience difficulty based on user’s video sequence viewing analysis at the click-level as Play, Pause, Move forward the Bayesian method is used in order to infer implicitly user’s interests. Learners with similar clickstream behavior are then segmented into clusters by using the unsupervised K-Means clustering algorithm. Videos that could meet the individual learner interests and offer a best and personalized experienced learning can therefore be recommended for a learner while enrolling in a SPOC based on his videos interactions and exploiting similar learners’ profiles.

Keywords—Learner/User profile, SPOCs, MOOCs, video Clickstream analysis, Recommendation, Personalization, Bayesian method, Clusters, k-means algorithm.

1 Introduction

The phenomenon of Massive open online courses (Moocs) attracted great attention at the end of 2012, which is known as the year of Moocs. In recent years, MOOCs are said to dominate remarkably online learning in higher education [1, 2] thanks mainly to the flexibility in time and place that they provide to learners and according to [1] these environments will still shape online learning in the future. Small Private Online Courses (SPOCS) like MOOCs, are remarkably growing in universities and corporate education especially in the field of blended learning and flipped classroom learning. Contrary to openness in MOOCs, SPOCs aim to offer a tailor-made course intended for small group of learners [2].

However, the “one-size-fits-all” learning model provided by these environments is not suitable meeting with a diversity of student profiles. The various students learning needs present a real challenge either to MOOCs [3, 4] or to SPOCS [5]. In order to overcome these challenges and to improve the average completion rate for MOOCs, adaptive learning and personalized learning are recently explored within MOOCs. Adaptivity and personalization concepts have great potential to deliver the best learn-

ing experiences to learners while enrolling in a MOOC [3, 4, 6, 7, 8, 9, 10, 11]. In fact, Personalization techniques are a classic solution recommended by many experts for improving learning and adapting courses to the learning preferences and characteristics of the students can enhance the learning process, leading to an increased learner satisfaction [12, 13, 14, 15]. Furthermore, in the last few years, recommendation techniques in the technology-enhanced learning (TEL) field, have invited increasing interest and are employed to enable learners in finding pertinent educational content to meet their profiles [16]. TEL recommendation systems have to be enlarged to various Web-based learning environments such as massive open online courses (MOOC) [17]

User profile is one of the issues that it is addressed by personalization systems to recommend individual items such as products, services, documents, etc. that meet the user's interests.

In fact, a user profile is a collection of information relating to user characteristics, interests, preferences and behavior within a system [18, 19]. This structure helps a system to know user requirements and behave in accordance of them [20]. Therefore, user profiling, which is the process of capturing the user relevant information, may be a really challenging task, since the values of certain attributes such as interests when they are unknown and not explicit submitted by the user. So, implicit techniques in user profiling gather more importance nowadays [21, 22].

In the present paper, we address to use an approach for user profiling in a Small Private Online Course (SPOC) based on an analogical reasoning that stipulate that a user can be characterized by determining a set of similar users [19]. The method known as clique-based filtering method [23, 19] is applied in the present work. This technique consists on matching a single profile with profiles of similar users. As Video lecture forms an extremely important part of SPOCs, we suggest building implicitly users profiles based on data collected from analyzing their video interactions while enrolling in the SPOC. We propose to track the user at the click level and we retain the following events: Play, Pause, Move Back (RW), Move Forward (FF), Replay and Download. In order to infer implicitly the user's interests, we suggest using the Bayesian method from datamining techniques. Then, for a given user, we propose to find other users who have similar video clickstream behavior by categorizing them into clusters using the K-means algorithm. The interests of users from the same cluster constitute an implicit profile of the individual user.

The objective of this work is to improve the learner's experience learning within the SPOC by recommending suitable videos that meet the personal learner's interest and match his individual needs by exploiting similar learners' profiles.

The main contributions of the present work is 1) the exploitation and the analysis of the clickstream among a SPOC, 2) The construction of a user profile based on the learner clickstream behavior, 3) the application of machine learning techniques among a SPOC to build a user profile and 4) this work is a contribution on personalization among a SPOC by recommendation of videos that could meet the learner interest and offer a best learning experience.

The structure of this paper is as follows: First, we present the various related work done either on adaptive MOOCs or on clickstream analysis. Second, we describe the elaborated user's profile and finally, we present our ongoing works.

2 Literature Review

2.1 User profile

The notion of profile appeared in the 1970s decade, which was mainly due to the need to create custom applications that could be adapted to the user [24]. User profiles were originally developed in the fields of information retrieval [18]. In fact, since a long time, User profiles have played a significant role in recommender systems, retrieval information systems, Search Personalization. But nowadays; user profiling is becoming a widely used technique in many applications like Adaptive Websites, e-business applications [22], information seeking [21], web browsing, one-to-one marketing [25], e-commerce websites [26, 27], web personalization systems[28], Adaptive MOOCs [3, 4, 6, 8, 9, 29, 30], E-learning systems [16, 31-35].

2.2 User Profiling

User Profiling is the process that consists on capturing user information related to his preferences, characteristics, activities and the representation of the user within a system. User profiling helps systems to behave according to users' individual interests in accordance to enhance user experience by providing personalization, adaptivity or suitable recommendation tailored to user's needs [20, 26]. A user profile can be static when the user information that it contains such user's common attributes like name, demographics data, learning style rarely changes or dynamic when the user profile's information alters frequently [27]. Various approaches and techniques were used to construct a user profile. Explicit user profiling approaches consist on submitting information directly by users using online forms, surveys, or rating [4, 6, 9, 29]. But even this technique is simple and easy way to create a user profile; it presents a risk of user entering deceitful or incomplete information [22]. Nowadays, recent research is more focused on getting user's data implicitly based on observing user system interaction and infer about relevance of the information captured in such a way [20, 21, 27, 31, 36, 37, 38].

2.3 In-Video behavior literature

As Video lectures form an extremely crucial part of MOOCs and SPOCs, we were interested in this paper to build a user profile based on an implicit analysis of data collected from learner video interactions within a SPOC at the click-level. Previous research has already recognized the benefits of user-based analysis related to viewed videos such as explicit comments or tags [39] and much of the available work concern video behavior modelization and prediction with lack of empirical studies [40]. Therefore, there is limited work on implicit indicators deduced from a click-level events analysis within a video [39, 40]. In fact, video interactions such as Playing, Pausing, replaying and peaks in re-watching sessions and play events are useful:

- To reflect either user's interest in a video or confusion [41]

- To reflect experience difficulty [40];
- To tell more about learner video engagement in MOOCs [42, 43, 44]
- To give important drop out indicators [41, 44]
- To generate video summaries [45] and fast video previews [46]
- To identify interesting video segment based on user interaction [39]
- To give insights to MOOC instructors about video production to enhance the learner experience [40, 42, 44]
- To identify video interaction patterns [41]; Video interaction profiles [40] and video watching profiles [44]
- To reveal the relationship that rely the complexity of videos and student video behavior in a MOOC [47].

3 Methodology

3.1 Study context

In the present work, we address to build users profiles in a Small Private Online Course (SPOC) named “UNIVTICE” intended for the teachers of the University Hassan 2–Casablanca-Morocco. This course is created and designed on Moodle and aims to accompany and train teachers to acquire techno-pedagogical skills. Our objective is to improve the teacher’s experience within the SPOC UNIVTICE by recommending suitable videos that meet his personal interest and match his individual needs by exploiting similar learners’ profiles. The present work is the first step where we describe the approach to construct the user profile. In an ongoing work, we intend to test the proposed method on a real dataset collected from the SPOC UNIVTICE.

3.2 The proposed User profiling approach

As videos are tremendously important components in the SPOC, we intent to build users profiles based on data collected from analyzing their video interactions while enrolling in the SPOC. We propose to track the learner at the click level and for this, we retain the following events: Play (PL), Pause (PA), Replay (RP), Move Forward (FF), Move Back (RW), Download (DL) and Stop (ST).

3.3 Notation

The following notation is used in the present paper (Table 1):

Table 1. Notation

Notation	Description
SPOC	Small Private Online Course
V_i	The i th video in the SPOC
n	Number of videos of the SPOC
M	Number of users
i	Subscript for videos, $i=1,2,\dots,n$
j	Up script for users, $j=1,2,\dots,m$
e	e is an event in (PL, PA, RP, FF, RW, DL, ST) Events: Play (PL), Pause (PA), Replay (RP), Move Forward (FF), Move Back (RW), Download (DL) and Stop (ST).
f_e	is the duration of the event e
$e_i^j = \frac{f_e}{\sum_{r \in (PL, PA, RP, FF, RW, DL, ST)} f_r}$	The weight of the event e related to user j and corresponding to video V_i : $PL_i^j, PA_i^j, RP_i^j, FF_i^j, RW_i^j, DL_i^j, ST_i^j$
VVS_i^j	Video Viewing Sequence related to user j on video V_i
\overline{VVP}_i^j	Vector Video Viewing Profile related to user j on video V_i
$[VVH]^j$	Matrix Video Viewing History for user j
VI_i^j	Video Interest related to user j in a video V_i
\overline{VIH}^j	Vector Video Interest History related to user j corresponding for all videos
UP^j	User Profile for user j
\overline{VT}_i	Mean of Total Viewing Time on a video V_i for all users
$VT_i^j = \begin{cases} 1 & \text{if } VT_i^j \geq \overline{VT}_i \\ 0 & \text{otherwise} \end{cases}$	Total Viewing Time for a user j on a video V_i :
\overline{MB}_i	Mean of Move Back on a video V_i for all users
$MB_i^j = \begin{cases} 1 & \text{if } MB_i^j \geq \overline{MB}_i \\ 0 & \text{otherwise} \end{cases}$	Number of Move Back events performed by a user j on a video V_i
DW_i^j	Download event performed by a user j on a video V_i $DW_i^j = \begin{cases} 1 & \text{if the user } j \text{ downloads the video } V_i \\ 0 & \text{otherwise} \end{cases}$
$RP_i^j = \begin{cases} 1 & \text{if the user } j \text{ replays the video } V_i \\ 0 & \text{otherwise} \end{cases}$	Replay event performed by a user j on a video V_i

3.4 The proposed User Profile UP^j

Let define a User Profile UP^j as a 3-tuplet $(j, [VVH^j], \overline{VIH^j})$:

Where:

j is the learner identifier;

$[VVH^j]$ is the Video Viewing History related to the learner j

and $\overline{VIH^j}$ is the Video Interest History related to the learner j .

Video Clickstream tracking

Video Viewing Sequence VVS_i^j : Let define a visioning session when a learner begins visioning a video and finishes when he performed another event or action within the SPOC. The learner video interactions are described as a sequence of the events performed by the learner on a video. Each event is described by its name and its duration in second. We adopt the discretization approach introduced in [48] for the temporal dimension. Therefore, we consider a unity of time equal to 1 second, so if event duration is equal to t seconds, we consider that the event is performed by the learner t times. We then normalize the formulation by retaining the event's weight as an event attribute relating to the total events achieved by a learner during a visioning session.

Therefore, for a learner j , we note VVS_i^j , his viewing sequence related to a video V_i that is described as a sequence of events $((e, e_i^j))$ performed on video V_i . Each event e in $(PL, PA, RP, FF, RW, DL, ST)$ is characterized by its weight e_i^j :

$$VVS_i^j = ((PL, PL_i^j), (PA, PA_i^j), (RP, RP_i^j), (FF, FF_i^j), (RW, RW_i^j), (DL, DL_i^j), (ST, ST_i^j)) \quad (1)$$

$$e_i^j = \frac{f_e}{\sum_{r \in (PL, PA, RP, FF, RW, DL, ST)} f_r}$$

(2)

Where e_i^j is the weight of event e , f_e is the event e duration.

Video Viewing Profile \overline{VVP}_i^j : Let define the Video Viewing Profile \overline{VVP}_i^j for a learner j related to a video V_i as a vector composed by all event weights:

$$(\overline{VVP}_i^j)^T = (PL_i^j, PA_i^j, RP_i^j, FF_i^j, RW_i^j, DL_i^j, ST_i^j) \quad (3)$$

Learner Video Viewing History $[VVH^j]$: Therefore, for a learner j , we can describe his Viewing Video History $[VVH^j]$ as a matrix composed by all the Videos Viewing Profiles related to the learner j :

$$DW_i^j = \begin{cases} 1 & \text{if the user } j \text{ downloads the video } V_i \\ 0 & \text{otherwise} \end{cases}$$

(7)

RP_i^j : Replay event for a user j on a video V_i :

$$RP_i^j = \begin{cases} 1 & \text{if the user } j \text{ replays the video } V_i \\ 0 & \text{otherwise} \end{cases}$$

(8)

The posteriori Probability of an Interest in a video V_i

Let denote VI_i^j the assumption that a user j has an interest in a video V_i . The posteriori probability $P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j)$ can be assessed using Bayesian method:

$$P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j) = \frac{\prod_{r \text{ in } (VT, MB, DW, RF)} P(r_i^j | VI_i^j) * P(VI_i^j)}{\sum_{r \text{ in } (VT, MB, DW, RF)} (P(r_i^j | VI_i^j) * P(VI_i^j) + P(r_i^j | \neg VI_i^j) * P(\neg VI_i^j))}$$

(9)

Where:

$P(VI_i^j)$ and $P(\neg VI_i^j)$ denotes the priori probability of an interest for a user j in a video V_i . They can be assessed with a statistical analysis on the training data.

$P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j)$ is the posteriori probability for an interest given the independent interest indicators $VT_i^j, MB_i^j, DW_i^j, RP_i^j$.

$P(r_i^j | VI_i^j)$ where $r \text{ in } (VT, MB, DW, RF)$ is the probability of the observing interest indicator given an interest VI_i^j in the video V_i . It can be computed with a statistical analysis on the training data.

$P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j)$ can then be estimated using the equation (10):

$$P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j) = \max(P(VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j), P(\neg VI_i^j | VT_i^j, MB_i^j, DW_i^j, RP_i^j))$$

(10)

Video Interest History $\overline{\mathbf{VIH}}^j$

For a learner j , the Video Interest History $\overline{\mathbf{VIH}}^j$ related to a learner j is composed by his interest values corresponding to all the videos can be defined as a vector:

$$(\overline{\mathbf{V}\mathbf{H}^j})^T = (V_1^j, V_2^j, \dots, V_i^j, \dots, V_n^j) \quad (11)$$

3.5 Users Clustering

The k-means algorithm is one of the widespread unsupervised learning algorithms [49] that is commonly used for classification and pattern recognition [50]. In the present work, we propose to use the k-means algorithm to classify learners given their Video Viewing History $[\mathbf{V}\mathbf{V}\mathbf{H}^j]$ into k clusters where each learner belongs to the cluster that minimize the Euclidian distance between his Video Viewing profile vector and the cluster centroid.

The k-means algorithm:

Lets k be the number of clusters.

Step 1) Initialization of clusters centroids:

Let $C = (C_1, C_2, C_3, \dots, C_k)$ denotes the set of clusters centers.

C_s Represents the centroid of cluster s, where $s = 1, 2, \dots, k$.

Step 2) Calculate the Euclidian distance between each learner video viewing profile given by (equation 3) and each cluster center C_s :

$$\begin{aligned} D(\mathbf{V}\mathbf{V}\mathbf{H}^j, \mathbf{V}\mathbf{V}\mathbf{H}^{C_s}) &= ((d(\overline{\mathbf{V}\mathbf{V}\mathbf{P}_1^j}, \overline{\mathbf{V}\mathbf{V}\mathbf{P}_1^{C_s}}), d(\overline{\mathbf{V}\mathbf{V}\mathbf{P}_2^j}, \overline{\mathbf{V}\mathbf{V}\mathbf{P}_2^{C_s}}), \dots, d(\overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^j}, \overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^{C_s}}), \dots, d(\overline{\mathbf{V}\mathbf{V}\mathbf{P}_n^j}, \overline{\mathbf{V}\mathbf{V}\mathbf{P}_n^{C_s}})) \\ &= \sum_{i=1}^n \left\| \overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^j} - \overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^{C_s}} \right\| \end{aligned} \quad (12)$$

Step 3) Assign the learner to the cluster that minimizes the Euclidian distance from all the cluster centers:

The user j is then affected to the cluster x that minimizes:

$$D(\mathbf{V}\mathbf{V}\mathbf{H}^j, \mathbf{V}\mathbf{V}\mathbf{H}^{C_x}) = \min_{s=1,2,\dots,k} \left(\sum_{i=1}^n \left\| \overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^j} - \overline{\mathbf{V}\mathbf{V}\mathbf{P}_i^{C_s}} \right\| \right) \quad (13)$$

Step 4) Recalculate cluster centers C_s :

Let x_{C_s} be the number of cluster s elements.

$$\begin{aligned}
 & V_1 \quad V_2 \quad \dots \quad V_i \quad \dots \quad V_n \\
 & \dots\dots\dots \\
 [\mathbf{vvh}^{C_s}] &= \frac{1}{x_{C_s}} \begin{bmatrix} \sum_{x=1}^{x_s} PL_1^x & \sum_{x=1}^{x_s} PL_2^x & \dots & \sum_{x=1}^{x_s} PL_i^x & \dots & \sum_{x=1}^{x_s} PL_n^x \\ \sum_{x=1}^{x_s} PA_1^x & \sum_{x=1}^{x_s} PA_2^x & \dots & \sum_{x=1}^{x_s} PA_i^x & \dots & \sum_{x=1}^{x_s} PA_n^x \\ \sum_{x=1}^{x_s} RL_1^x & \sum_{x=1}^{x_s} RL_2^x & \dots & \sum_{x=1}^{x_s} RL_i^x & \dots & \sum_{x=1}^{x_s} RL_n^x \\ \sum_{x=1}^{x_s} FF_1^x & \sum_{x=1}^{x_s} FF_2^x & \dots & \sum_{x=1}^{x_s} FF_i^x & \dots & \sum_{x=1}^{x_s} FF_n^x \\ \sum_{x=1}^{x_s} RW_1^x & \sum_{x=1}^{x_s} RW_2^x & \dots & \sum_{x=1}^{x_s} RW_i^x & \dots & \sum_{x=1}^{x_s} RW_n^x \\ \sum_{x=1}^{x_s} DL_1^x & \sum_{x=1}^{x_s} DL_2^x & \dots & \sum_{x=1}^{x_s} DL_i^x & \dots & \sum_{x=1}^{x_s} DL_n^x \\ \sum_{x=1}^{x_s} ST_1^x & \sum_{x=1}^{x_s} ST_2^x & \dots & \sum_{x=1}^{x_s} ST_i^x & \dots & \sum_{x=1}^{x_s} ST_n^x \end{bmatrix} \\
 &= \frac{1}{x_{C_s}} \sum [\mathbf{vvh}^{C_s}] \\
 &= (\overline{\mathbf{VVP}}_1^{C_s}, \overline{\mathbf{VVP}}_2^{C_s}, \dots, \overline{\mathbf{VVP}}_i^{C_s}, \dots, \overline{\mathbf{VVP}}_n^{C_s})
 \end{aligned}$$

(14)

Go to step 2)

Step 5) If no learner was reassigned then stop, otherwise, repeat from step 3).

3.6 Video Recommendation

For a new user J_0 , let $\overline{\mathbf{VVP}}_{i_0}^{J_0}$ be his video viewing profile related to video V_{i_0} . Let C_{s_0} be the cluster to which the user J_0 belongs so as the Euclidian distance to the cluster center is the lowest distance.

The Video Interest $VI_{i_0}^{J_0}$ for the user J_0 related to video V_{i_0} can be estimated as described as follows:

As possible Video Interest values are 1 or 0, let f1 be the frequency of value 1 and f2 designs the frequency of value 0.

For each user x belonging to a cluster C_{s_0} , the video interest VI_i^x concerning the video V_i

$$VI_i^x, x=1,2,\dots,x_{C_{s_0}}$$

where $x_{C_{s_0}}$ is the number of cluster C_{s_0} members .

Let then define VI_{mode} , the most common value (0 or 1) for the video interest for all cluster's members. The "mode" designs the value that occurs most often:

$$VI_{mode} = Max(f_1, f_2)$$

(15)

For the user j_0 , the video interest $VI_i^{j_0}$ concerning the video V_i is evaluated by the following relation:

$$VI_i^{j_0} = VI_{mode} \quad i=1,2,\dots,n$$

(16)

Therefore, a list of videos for that the interest value is equal to 1 can be proposed to the user J_0 .

4 Conclusion and Future work

In this work, we proposed an approach to construct a learner profile within a SPOC. The suggested approach is based on learner video clickstream analysis and on the use of machine learning techniques to implicitly assess the learner video interest. Our objective is to improve a learner experience while enrolling within a SPOC by recommending videos that could meet his personal interest and match his individual needs by exploiting similar learners' profiles.

For future work, we plan to test the proposed approach by using a real dataset in the SPOC UNIVTICE in order to improve the learner experience while enrolling within the SPOC. We intend also the exploitation of the dataset to Clustering learners' video behaviors, to assess implicitly learners' difficulties and finally the comparison of the results with similar works.

5 References

- [1] Baturay, M. H. (2015), An overview of the world of MOOCs, Procedia – Social and Behavioral Sciences 174, 427 – 433 <https://doi.org/10.1016/j.sbspro.2015.01.685>
- [2] Kaplan, A. M., Haenlein, M., Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster. Business Horizons, 2016. 59(4), 441-450. <https://doi.org/10.1016/j.bushor.2016.03.008>

- [3] Onah, D. F. O. and Sinclair, J., (2015) Massive open online courses: an adaptive learning framework. In: 9th International Technology, Education and Development Conference, Madrid, Spain, 2-4 Mar 2015. Published in: INTED2015 Proceedings, 1258-1266.
- [4] García-Peñalvo, F.J., Fidalgo-Blanco, A., Sein-Echaluce, M.L. (2017). An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education, *Telematics and Informatics* <https://doi.org/10.1016/j.tele.2017.09.012>
- [5] Renée M. Filius; Renske A.M. De Kleijn; Sabine G. Uijl; Frans J. Prins; Harold V.M. Van Rijen; Diederick E. Grobbee, Challenges concerning deep learning in SPOCs, *International Journal of Technology Enhanced Learning*, 2018, 10(1/2), 111-127.
- [6] Sonwalkar, N. (2013). The First Adaptive MOOC: A Case Study on Pedagogy Framework and Scalable Cloud Architecture—Part I. In *MOOCs Forum*, 1(P), 22-29). 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: Mary Ann Liebert, Inc.
- [7] Leris, D., Sein-Echaluce, M. L., Hernandez, M., & Bueno, C. (2017). Validation of indicators for implementing an adaptive platform for MOOCs. *Computers in Human Behavior*, 72, 783-795. <https://doi.org/10.1016/j.chb.2016.07.054>
- [8] Clerc, F., Lefevre, M., Guin, N., & Marty, J. C. (2015). Mise en place de la personnalisation dans le cadre des MOOCs. In 7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2015), 144-155.
- [9] Sein-Echaluce, M. L., Fidalgo-Blanco, A., García-Peñalvo, F. J., & Conde, M. A. (2016). iMOOC Platform: Adaptive MOOCs. In P. Zaphiris, & A. Ioannou (Eds.), Volume 9753 of the series *Lecture Notes in Computer Science Learning and Collaboration Technologies* (pp. 380-390). Switzerland: Springer International Publishing.
- [10] Daniel, J., Vázquez Cano, E., & Gisbert, M. (2015). The Future of MOOCs: Adaptive Learning or Business Model? *RUSC. Universities and Knowledge Society Journal*, 12(1), 64-73. <https://doi.org/10.7238/rusc.v12i1.2475>
- [11] Li Y.-H., Zhao B. and Gan J.-H. (2015), Make adaptive learning of the MOOC: The CML model, 10th International Conference on Computer Science and Education, ICCSE 2015, art.no.7250398, 1001-1004.
- [12] Henning, P. A., Heberle, F., Streicher, A., Zielinski, A., Swertz, C., Bock, J. and Zander, S., (2014), Personalized web learning: Merging open educational resources into adaptive courses for higher education, *Proc. Personalization Approaches Learn. Environ.*, p. 55, 2014.
- [13] Brusilovsky, P. (1999), Adaptive and intelligent technologies for web-based education. *KI*, 13(4), 19-25.
- [14] Brusilovsky, P., Peylo, C. (2003), Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education (IJAIED)*, 13, 159-172. <hal-00197315>
- [15] Alshammari, M. , Anane, R. and Hendley, R. J. (2014), Adaptivity in E-learning systems' in *Proceedings - 2014 8th International Conference on Complex, Intelligent and Software Intensive Systems, CISIS 2014*, 79-86. IEEE. DOI: 10.1109/CISIS.2014.12 <https://doi.org/10.1109/CISIS.2014.12>
- [16] Muruganandam, S. and Srinivasan, N. (2017), Personalised e-learning system using learner profile ontology and sequential pattern mining-based recommendation, *Int. J. Business Intelligence and Data Mining*, 12(1), 78–93. <https://doi.org/10.1504/IJBIDM.2017.082704>
- [17] Khrabi, K.M., Jemni, M. and Nasraoui, O. (2015), Recommendation systems for personalized technology-enhanced learning?, *Ubiquitous Learning Environments and Technologies*, 159–180, Springer, Berlin Heidelberg.

- [18] Brusilovsky, P. and Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In: Brusilovsky P., Kobsa A., Nejdl W. (eds) The Adaptive Web. Lecture Notes in Computer Science, vol 4321. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72079-9_1
- [19] Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalised hypermedia presentation techniques for improving online customer relationships. The knowledge engineering review, 16(2), 111-155. <https://doi.org/10.1017/S0269888901000108>
- [20] Kanoje, S., Girase, S. and Mukhopadhyay, D. (2014). User profiling trends, techniques and applications. International Journal of Advance Foundation and Research in Computer (IJAFRC) 1(1), Jan 2014. <https://doi.org/10.1504/IJAIP.2009.026760>
- [21] Shtykh, R. Y., & Jin, Q. (2009). Dynamically constructing user profiles with similarity-based online incremental clustering. International Journal of Advanced Intelligence Paradigms, 1(4), 377-397. <https://doi.org/10.1504/IJAIP.2009.026760>
- [22] Kanoje, S., Mukhopadhyay, D., & Girase, S. (2016). User Profiling for University Recommender System Using Automatic Information Retrieval. Procedia Computer Science, 78, 5-12. <https://doi.org/10.1016/j.procs.2016.02.002>
- [23] Alspector, J, Kolcz, A and Karunanithi, N (1997). Feature-based and clique-based user models for movie selection: a comparative study, User Modeling and User-Adapted Interaction 7(4) 279–304. <https://doi.org/10.1023/A:1008286413827>
- [24] Bouneffouf, D. (2013). Towards user profile modelling in recommender system. arXiv:1305-1114.
- [25] Lakiotaki, K., Matsatsinis, N. (2012). Analysing user behavior in recommender systems, International Journal of Electronic Business 2012 – 10(1), 1 – 19.
- [26] Wang, H., Liu, Y. and Yin, P. (2012) ‘Study on user preferences modelling based on web mining’, Int. J. Information Technology and Management, 11(4), 307–322. <https://doi.org/10.1504/IJITM.2012.049995>
- [27] Inbarani H., Thangavel, K., (2011). Discovery of user profiles using fuzzy web intelligent techniques, International Journal of Web Based Communities Jan 2011, 7 (3), 357-375 <https://doi.org/10.1504/IJWBC.2011.041204>
- [28] Hawalah, A., Fasli, M. (2015), Dynamic user profiles for web personalisation, Expert Systems with Applications, 2015, 42(5), 2547-2569, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2014.10.032>
- [29] Watson S.L., Watson W.R., Yu J.H., Almarh H. & Mueller C., Learner profiles of attitudinal learning in a MOOC: An explanatory sequential mixed methods study, Computers & Education (2017). <https://doi.org/10.1016/j.compedu.2017.07.005>
- [30] El Mhouth, A., Nasseh, A., & Erradi, M. (2016). Stimulate Engagement and Motivation in MOOCs Using an Ontologies Based Multi-Agents System. International Journal of Intelligent Systems and Applications, 8(4), 33-42. <https://doi.org/10.5815/ijisa.2016.04.04>
- [31] Schiaffino, S., Garcia, P., & Amandi, A. (2008). eTeacher: Providing personalized assistance to e-learning students. Computers & Education, 51(4), 1744-1754 <https://doi.org/10.1016/j.compedu.2008.05.008>
- [32] Kritikou, Y., Demestichas, P., Adamopoulou, E.F., Demestichas, K.P., Theologou, M.E., & Paradia, M. (2008). User Profile Modeling in the context of web-based learning management systems. J. Network and Computer Applications, 31, 603-627. <https://doi.org/10.1016/j.jnca.2007.11.006>
- [33] Golemati, M., Katifori, A., Vassilakis, C., Lepouras, G., & Halatsis, C. (2007). Creating an ontology for the user profile: Method and applications. In RCIS, 407– 412.
- [34] Xu, D., Wang, H., & Su, K. (2002). Intelligent student profiling with fuzzy models. In Proceedings of the 35th Hawaii international conference on system sciences.

- [35] Abel, F., Herder, E., Houben, G.-J., Henze, N. and Krause, D. (2013). Cross system user modeling and personalization on the social web. *User Modeling and User Adapted Interaction*, 23, 169–209. <https://doi.org/10.1007/s11257-012-9131-2>
- [36] Premlatha, K. R., Dharani, B., & Geetha, T. V. (2014). Dynamic learner profiling and automatic learner classification for adaptive e-learning environment. *Interactive Learning Environments*, 24(6), 1054-1075. <https://doi.org/10.1080/10494820.2014.948459>
- [37] Bagher, R. C., Hassanpour, H. and Mashayekhi, H. (2017). User trends modeling for a content-based recommender system. *Expert Systems with Applications*, 87, 209-219. <https://doi.org/10.1016/j.eswa.2017.06.020>
- [38] Chen, L., Chen, G. and Wang, F. (2015). Recommender systems based on user reviews: the state of the art. *User Modeling and User-Adapted Interaction*, 25(2), 99-154. <https://doi.org/10.1007/s11257-015-9155-5>
- [39] Chorlianopoulos, K. (2013), Collective intelligence within web video. *Human-centric Computing and Information Sciences* 3:10. <https://doi.org/10.1186/2192-1962-3-10>
- [40] Li, N., Kidzinski, L., Jermann, P., & Dillenbourg, P. (2015). How do in-video interactions reflect perceived video difficulty? In *Proceedings of the European MOOCs Stakeholder Summit 2015*, No. EPFL-CONF-207968, 112-121. PAU Education
- [41] Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014, March). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, 31-40. ACM.
- [42] Guo, P. J., Kim, J., & Rubin, R. (2014, March). How video production affects student engagement: An empirical study of mooc videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, 41-50. ACM.
- [43] Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*, 170-179. ACM. <https://doi.org/10.1145/2460296.2460330>
- [44] Sinha, T., Jermann, P., Li, N. and Dillenbourg, P. (2014). Your click decides your fate: Inferring information processing and attrition behavior from MOOC video clickstream interactions. arXiv preprint arXiv:1407.7131.
- [45] Peng W-T, Chu W-T, Chang C-H, Chou C-N, Huang W-J, Chang W-Y, Hung Y-P (2011) Editing by viewing: automatic home video summarization by viewing behavior analysis. *Multimedia, IEEE Transactions on* 13(3):539–550
- [46] Syeda-Mahmood, T., & Ponceleon, D. (2001, October). Learning video browsing behavior and its application in the generation of video previews. In *Proceedings of the ninth ACM international conference on Multimedia*, 119-128. ACM. <https://doi.org/10.1145/500141.500161>
- [47] Sluis, F.V., Ginn, J., and Zee, T.V. Explaining Student Behavior at Scale: The Influence of Video Complexity on Student Dwelling Time. In *Proc. ACM L@S 2016*.
- [48] Branch, P., Egan, G., & Tonkin, B. (1999). Modelling interactive behavior of a video based multimedia system. In *Communications, 1999. ICC'99. 1999 IEEE International Conference*, 2, 978-982.
- [49] MacQueen, J. B. (1967): Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1:281-297
- [50] Davidson, I. (2002). Understanding K-Means Non-Hierarchical Clustering, SUNY Albany Technical Report 02–2.

6 Authors

Naima Belarbi is an Engineer in Automatic and Computer from Mohammadia School of Engineers (EMI Rabat) Morocco and a Teacher trainer in Regional Education and Training Trades Center (CRMEF) Casablanca. She is also a PhD student in Laboratory of Technological Information and Modelisation (LTIM) -Faculty of Sciences Ben M'Sik, University Hassan II - Casablanca, Morocco. Her area of interest: Blended learning, MOOCS, SPOCs, Mobile learning, Adaptive learning, Data Science.

Nadia Chafiq is a PhD in Educational Technology. She is a Professor at Faculty of Sciences Ben M'Sik, University Hassan II of Casablanca, Morocco, B.P 7955 Sidi Othmane. She operates in several fields of educational sciences: educational technologies, engineering of distance education and didactics. She is a Member of the Observatory of Research in Didactics and University Pedagogy (ORDIPU) and the Multidisciplinary Laboratory in Sciences and Information Communication and Education Technology (LAPSTICE). She also coordinates works on digital learning: hybrid device, inverted classes, collaborative platforms, serious games, MOOCS, SPOCS, mobile learning, learning analytics and performance management

Mohamed Talbi is a PhD in Sciences and Processes of Analysis from the University Pierre et Marie Curie of Paris. He is currently the Dean of the Faculty of Sciences Ben M'Sik at Hassan II University, B.P 7955 Sidi Othmane, Casablanca, Morocco, and the Director of the Observatory of Research in Didactics and University Pedagogy (ORDIPU) since 2014. He is an Expert in the fields of teaching and research on educational technologies, engineering of the distance training, techniques of training, information systems. He is the author of several national and international awards and has accumulated more than 30 years of scientific productions

Abdelwahed Namir is a PhD in Numerical Methods for Engineers from Mohammadia School of Engineers (EMI Rabat) Morocco. He is also the Head of Laboratory of Technological Information and Modelisation (LTIM) Faculty of Sciences Ben M'Sik, University Hassan II - Casablanca, Morocco. His research interests include Mathematics, Computers, modeling.

El Habib Benlahmar is a PhD in Computer Science. He is a professor of Higher Education at the Mathematic & Computer Science Department at the Faculty of Sciences Ben M'Sik, Casablanca- Morocco. In January 2008, he won the Prize for the best thesis in information sciences and communication His research interests include Metasearch, Information retrieval, Semantic web, Automatic Processing of Natural Language, Data Science, Mobile Applications. He is also the chief of the semantic web and knowledge retrieval team at the university Hassan II Mohammedia Casablanca, Laboratory of Technology and Information Modelisation (LTIM)–Casablanca-Morocco.

Article submitted 24 June 2018. Resubmitted 15 July 2018. Final acceptance 03 August 2018. Final version published as submitted by the authors.