



**HAL**  
open science

## A two-step clustering approach for improving educational process model discovery

Hanane Ariouat, Awatef Hicheur Cairns, Kamel Barkaoui, Jacky Akoka,  
Nasser Khelifa

► **To cite this version:**

Hanane Ariouat, Awatef Hicheur Cairns, Kamel Barkaoui, Jacky Akoka, Nasser Khelifa. A two-step clustering approach for improving educational process model discovery. 25th IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 2016), Jun 2016, Paris, France. pp. 38-43, 10.1109/WETICE.2016.18 . hal-01500511

**HAL Id: hal-01500511**

**<https://hal.science/hal-01500511>**

Submitted on 3 Apr 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>  
Eprints ID : 17049

The contribution was presented at WETICE 2016 :  
<http://www-inf.int-evry.fr/WETICE/WETICE2016/>

**To cite this version** : Ariouat, Hanane and Cairns, Awatef Hicheur and Barkaoui, Kamel and Akoka, Jacky and Khelifa, Nasser *A two-step clustering approach for improving educational process model discovery*. (2016) In: 25th IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 2016), 13 June 2016 - 15 June 2016 (Paris, France).

Any correspondence concerning this service should be sent to the repository administrator: [staff-oatao@listes-diff.inp-toulouse.fr](mailto:staff-oatao@listes-diff.inp-toulouse.fr)

# A two-step clustering approach for improving educational process model discovery

Hanane Ariouat

Lab. IRIT, Toulouse 1 University, France  
Email: hanane.ariouat@irit.fr

Awatef Hicheur CAIRNS

ALTRAN Research, Vlizy-Villacoublay, France  
Email: awatef.hicheurcairns@altran.com

Kamel Barkaoui

Lab. Cedric, Cnam Paris, France  
Email: barkaoui@cnam.fr

Jacky Akoka

Lab. Cedric, Cnam Paris, France  
Email: jacky@cnam.fr

Nasser Khelifa

ALTRAN Institute, France  
Email: nasser.khelifa@altran.com

**Abstract**—Process mining refers to the extraction of process models from event logs. As real-life processes tend to be less structured and more flexible, clustering techniques are used to divide traces into clusters, such that similar types of behavior are grouped in the cluster. Educational process mining is an emerging field in the educational data mining (EDM) discipline, concerned with developing methods to better understand students' learning habits and the factors influencing their performance. However, the obtained models, usually, cannot fit well to the general students' behaviour and can be too large and complex for use or analysis by an instructor. These models are called spaghetti models. In the present work, we propose to use a two steps-based approach of clustering to improve educational process mining. The first step consist of creating clusters based employability indicators and the second step consist on clustering the obtained clusters using the AXOR algorithm which is based on traces profiles in order to refine the obtained results from the first step. We have experimented this approach using the tool ProM Framework and we have found that this approach optimizes at the same time, both the performance/suitability and comprehensibility/size of the obtained model.

**Keywords**—Educational process mining; Process Discovery; Fitness; Clustering;

## I. INTRODUCTION

Nowadays, education and training centers promote personalized curriculums where students are free to choose the skills they want to develop (from beginner to specialist), the way they want to learn (theoretical or practical aspects) and the time they want to spend. This tendency is reinforced by the emergence of "e-learning" which represents an increasing proportion of in-company training. Educational Data Mining (EDM) and Learning Analytics and Knowledge (LAK) [6] study data and analytics in education, teaching, and learning, suggesting educational priorities and undertaking high-quality research into the models, methods, technologies, and impact of analytics. One of the current promising techniques in EDM and LAK is Educational Process Mining (EPM). The idea of process mining [9] is to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs (recorded by an information system). EPM [7] aims at (i) constructing complete and compact educational process models that are able to reproduce all observed behaviour (process model discovery), (ii) checking whether the modelled behaviour

(either pre-authored or discovered from data) matches the observed behaviour (conformance checking), and (iii) projecting extracted information from the logs onto the model, to make the tacit knowledge explicit and facilitate better understanding of the process (process model extension).

The results of EPM can be used to get a better understanding of the underlying educational processes, to generate recommendations and advice to students, to provide feedback to either students, teachers or/and researchers, to detect learning difficulties early, to help students with specific learning disabilities, to improve management of learning objects, etc.; However, the obtained model cannot fit well to the general student's behaviour and can be too large and complex for use or analysis by an instructor. These models are usually called *spaghetti models*.

In order to cope with such a situation, in this article we propose a new clustering approach, by partitioning the above complexe models, to obtain a simple and comprehensive process models. Our approach consists of two steps, which is a combination of two recent works [1] and [5]. The first step is based on the clustering approach proposed in [1] in the field of educational process mining. It aims to decompose educational processes following key performance indicators. The second step is based on the work of Ariouat et al in [5] which consist of clustering business log events by using a clustering algorithm called *AXOR Algorithm*. This algorithm uses traces profiles in order to identify and distinguish the most performant learning paths So, the combination of these two works can be seen as an extension of the approach of [1] by the approach of [5], (which was primarily proposed to cluster activities based on the notion of activity profile [10]), to improve and optimize the results of clustering techniques in the field of educational process mining.

The paper is organized as follow: In section 2 we introduce different concepts of educational process mining, in section 3 we present our two-steps clustering technique. and finally, in section 4 we conclude this article.

## II. EDUCATIONAL PROCESS MINING

### A. Definition:

Process mining is a relatively new technology which emerged from business community [12]. It focuses on the development of a set of intelligent tools and techniques aimed at extracting process-related knowledge from event logs. The complete overview of process mining application in the educational field (known as educational process mining [7], [8]) is illustrated in Figure 1. An event log corresponds to a set of process instances (i.e. traces) following a business process. Each recorded event refers to an activity and is related to a particular process instance. An event can have a timestamp and a performer (i.e. a person or a device executing or initiating an activity). Typical examples of event logs in education may include student's registration procedures and attended courses, student's examination traces, use of pedagogical resources and activity logs in e-learning environments.

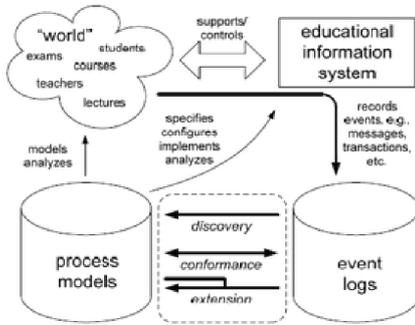


Fig. 1. Process mining concepts [7]

To discover a suitable process model, it is assumed that the event log contains a representative sample of behaviour. However, the application of process discovery techniques presents some challenges given the huge volume and the traces' heterogeneity often encountered in educational datasets [1].

### III. PROCESS MODEL DISCOVERY USING A TWO-STEP CLUSTERING TECHNIQUE

In order to handle the complexity and heterogeneity of the training paths encountered in the educational domain and to cope with the issue mentioned in [13], we propose a two-step clustering technique as a pre-processing step. Our goal is to identify the best training paths by dividing a training event log into homogenous subsets of cases following both their structural similarity and an employability indicator indicating the effectiveness of a training path. In our two-step clustering approach, training paths are firstly partitioned following performance indicators (employability factor and period of unemployment) then training path in each obtained cluster are partitioned further following their structural similarity. Figure 2 illustrates the overall methodology of our clustering approach.

Our motivating example is based on real-world professional training databases from a worldwide consulting company. This company has around 6 000 employees that are free to choose different training courses aligned with their profiles,

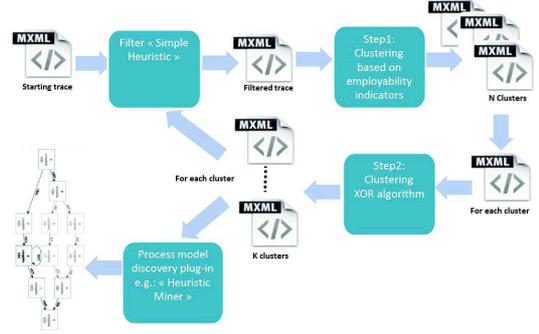


Fig. 2. Methodology of decision's criteria extraction

during their careers. These training courses are provided by internal or external training organizations. The data collected for analysis reports all the 16 260 training courses followed by 3440 employees, during the last three years, performed by 494 training organisations. This data includes the employees profiles (identifier, function, and number of years of service), their careers (i.e., the jobs/missions they did) and their training paths (the set of training courses taken during the past three years) (See Table 1). In this section, we show how process mining techniques can be used to analyze the training processes underlying this dataset.

Matricul	Profil	Training Id	Training Label	Training Orga Id	StartDate	EndDate
7	CONSULTANT	Tr 830	EXCEL BLEARNING	Org 13	11/07/2011	31/12/2011
8	CONSULTANT	Tr 769	QF TEST	Org 135	26/04/2011	28/04/2011
9	CONSULTANT	Tr 252	INTERCULTURAL WORKING RELATONS : INDIA SELENIUM	Org 135	01/07/2011	01/07/2011
10	CONSULTANT	Tr 260	UML FUNCTIONAL ANALYSIS	Org 135	25/10/2011	26/10/2011
11	CONSULTANT	Tr 812	DESIGN PATTERNS AND APPLICATION C++	Org 135	24/10/2011	27/10/2011
12	CONSULTANT	Tr 774	SQL BASIC	Org 135	08/12/2011	09/12/2011
13	CONSULTANT	Tr 1923	C++ ADVANCED	Org 135	03/04/2012	05/04/2012
14	CONSULTANT	Tr 813	XML BASIC AND XPATH	Org 135	04/04/2012	06/04/2012
15	CONSULTANT	Tr 2014	DESIGN PATTERNS AND APPLICATION IN C++	Org 135	10/04/2012	11/04/2012
14	CONSULTANT	Tr 1282	DESIGN PATTERNS AND APPLICATION IN C++	Org 135	13/09/2012	14/09/2012
...	...	...	...	...	...	...

TABLE I. EXAMPLE OF AN EDUCATIONAL EVENT LOG

#### A. First step

This step consists of creating clusters of similar trainees' profiles based on a training path performance indicator expressed via two criteria. The first one, called employability, concerns the matching between the obtained skills after a training course and those required by a mission. The second criterion represents the time period between a training course followed by an employee and a new mission on which the employee is staffed after it. So, it is important to know, before using these data, we applied the Preprocessing phase.

a) **Matching criterion:** The criteria that models the matching between skills acquired during a training course and the ones required for a given job/placement, is considered as a real number included between 0 and 1. Hence, this criteria does the matching between a training course followed by an employee, with an identifier  $\ll i \in \{1, \dots, 3340\} \gg$ , and a job/ = placement will be noted  $\ll A_i \gg$  with  $\ll A_i \in \{0, 1\} \gg$ . The set of skills obtained by, an employee, identified by  $i$ , during his/her trainings is expressed as follows:  $\ll F^i = \{F_1^i, F_2^i, \dots, F_{n_i}^i\} \gg$ .

Where  $n_i$  is an integer greater than or equal to 1. Generally  $n_i$  is at least equal to 3 and less than 10. We note also that for all  $\ll j \in \{1, \dots, n_i\} \gg$ ,  $F_j^i$  indicates that the training course  $j$  is followed by the employee  $i$ . For example,  $F_1^{10} = \text{Anglais}$  means that the employee 10 has followed the English training course. In the same way, the set of skills required by a given job/placement on which the employee  $i$  has been staffed is noted as follows:  $\ll M^i = \{M_1^i, M_2^i, \dots, M_{m_i}^i\} \gg$  Where  $m_i$  is an integer greater than or equal to 1. Generally it is equal to 4 or 5. Also, for all  $K \in \{1, \dots, m_i\}$ ,  $M_k^i$  indicates that the skill number  $k$  is required for the job under consideration. For instance,  $M_1^{10} = \text{Anglais}$  means that the found job/placement for the employee number 10 requires English language skill. In addition, the required skills by a given job/placement are weighted according to their importance for the success of this job. This weighting is modelled as follow:  $\ll P^i = \{P_1^i, P_2^i, \dots, P_{m_i}^i\} \gg$ . Where for all  $j \in \{1, \dots, m_i\}$ ,  $0 < P_j^i < 1$  is the weight associated to the competence  $\ll M_j^i \gg$  and  $\sum_{k=1}^{m_i} P_j^i = 1$ .

Therefore, the matching criteria between skills obtained by training courses and skills required for a given job/placement is calculated by the following formula:

$\ll A^i = \sum_{k=1}^{m_i} P_j^i \times \Pi_{\{M_k^i \in F^i\}} \gg$ . With  $\Pi_{\{M_k^i \in F^i\}}$  is an indicator computed by the following rule:

$$\Pi_{\{M_k^i \in F^i\}} = \begin{cases} 1 & \text{si } M_k^i \in F^i \\ 0 & \text{si } M_k^i \notin F^i \end{cases} \quad (1)$$

Hence, the distribution characterizing this matching criterion, using our training catalogue and employee information recorded in our example training courses' dataset, is given Figure 3.

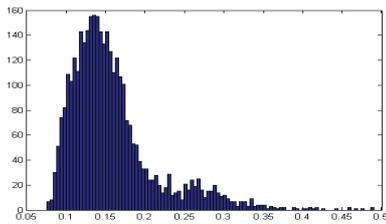


Fig. 3. Matching distribution between training courses and jobs for the employees of our training dataset example.

b) **Time period between training courses and jobs/placements:** This criterion represents, for an employee  $i$ , the time periode between the end of a given training course and the start of his/her next job/placement. This criterion follows the normal log probability law. This law is widely used in the modelling of survivor duration. In fact, using the durations, expressed in working days, we obtain the estimated parameters for the used normal log law as follow:

$$\hat{\mu} = 3.16445 [3.11872, 3.21018]$$

$$\hat{\sigma} = 1.12863 [1.09721, 1.1619]$$

The graphic representation of the fit of this law is given in Figure 4. Let us note that we normalize the durations according to the maximum one in order to have a criterion value comprised between 0 and 1. The goal of this normalization is to homogenize the duration criterion with the matching one.

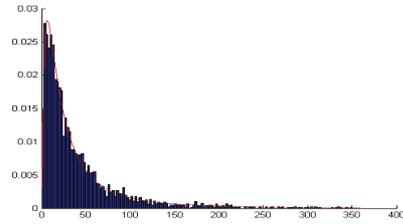


Fig. 4. Density probability of the log normal law describing the time between training courses' end and the beginning of new jobs for the employees of the training courses' dataset example.

### c) Classification according to duration and matching criterion:

In these experiments, we do classification based on the matching and duration criteria defined below. This classification will help us identify class of training paths for employees that allow them to be staffed on jobs shortly after a training course.

**Definition of the cluster number:** To get these classes we use the "K-means" technique [2][3], where the optimal number of clusters is determined using a method based on the average silhouette of many clusters where the number of the clusters is varied (the number of clusters  $K$  is varied between 2 and 5). For more details on this silhouette method, interested readers may refer to [14]. The obtained results are presented in Figure 5. When analyzing this figure, we identify a breaking down of the progression of the average silhouette when  $K=3$ , this means that the clustering is optimal when we do a classification with 3 partitions (i.e. clusters).

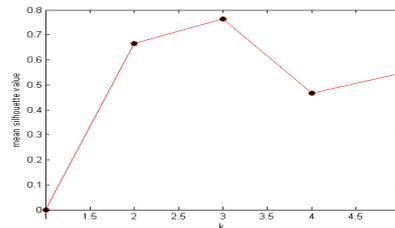


Fig. 5. Silhouette Graphical analysis is used to determine the optimal number of clusters. X-axis represents number of clusters and Y-axis indicates associated Silhouette scores.

Clustering for  $K=3$ : According to the results obtained in the previous analysis, we apply the "K-Means" method, based on the matching and duration criteria, with  $K=3$ , on our training courses' dataset example. The obtained results are given in Figure 6.

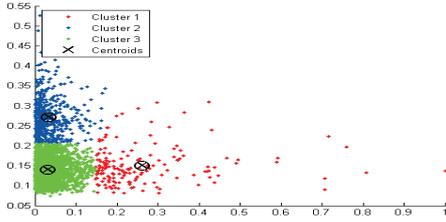


Fig. 6. Results of the K-means clustering method applied on our training courses dataset example using the matching and duration criteria. X-axis represents time period (normalized) between training and the next job. Y-axis corresponds to employability score in (0,1).

Let us note that the first cluster (cluster 1) contains trainees with the worst employability indicator and the longest unemployment duration (time period between a training course and the following next job/placement). We applied Heuristic Miner [9], one of the robust algorithms to investigate the processes in users' behaviour. Heuristic Miner can be used to express the main behaviour registered in an event log. It focuses on the control flow perspective and generates a process model in the form of a Heuristics Net for the given event log. Therefore, the Heuristic Miner algorithm was designed to make use of a frequency based metric and so it is less sensitive to noise and the incompleteness of logs. We used the default threshold parameters of Heuristic Miner algorithm provided by ProM to discover the process model from the training traces of the trainees grouped in the first cluster. We obtain clearly identifiable training paths, as illustrated in Figure 7.

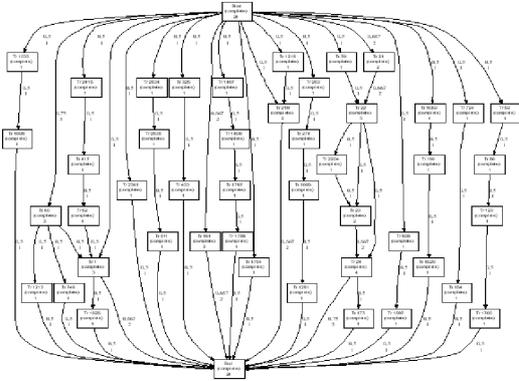


Fig. 7. A fragment of the process model showing all the training patterns of cluster 1.

Cluster 2 contains trainees with the best employability factor and the shortest unemployment duration (time period between a training course and the following next job/placement). Let us note that the training paths underlying this cluster correspond to the highest performing ones regarding employability factor and period of unemployment. Figure 8 illustrates the process model discovered from the training traces grouped in the second cluster (using the Heuristic Miner). Clearly, it's a spaghetti process. Cluster 3 contains average training paths regarding employability factor and period of unemployment. The process model discovered from Cluster 3 is even more complex.

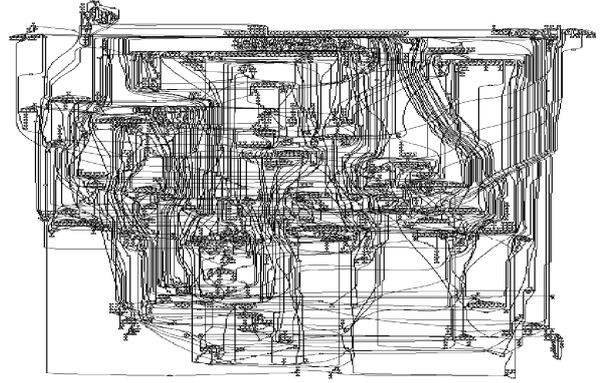


Fig. 8. Fragment of the process model (spaghetti-like) underlying cluster 2.

In order to obtain simpler training process models, the clusters two and three will be analyzed separately in the second step. The second step of our approach, we simplify further the process models -discovered in the first step, by grouping training paths from each cluster following their structural similarity. Let us note that the first step facilitates the detection of the process patterns and enhances analysis performance because it reduces the searching scope from the whole trainees' information to limit ones for each inferred clusters.

### B. Second step

The second step consists of grouping training paths (traces in event log) from each of the last two complex clusters discovered in the first step, using our clustering technique propose in [5]. Instead of extracting features from traces, our approach is based on activity profiles to distinguish the most performant learning paths from the less performant ones, and we use the hamming distance [15] measure between two traces using the logical XOR operator (hamming distance).

a) **Activity Profiles:** Traces in an event log are characterized by profiles. A profile is a vector  $Profile\_Vector$  composed of  $n$  items (which support binary values) as follow :

$$Profile\_Vector_i = (a_1, a_2, \dots, a_n).$$

Where  $a_{j=1..n}$  are binary values representing if an activity  $a_j$  is present in the trace or not. 1 means the activity occurs at least once, and 0 means that the activity does not occur in the trace. These resulting vectors will be used to calculate the distance between two cases.

b) **Distance Measure:** To calculate the distance between two cases, we use the XOR operator as follow : let us, consider two traces  $a$  and  $b$  :  $a = (0001111)$  and  $b = (1101011)$  We have  $a \text{ XOR } b = (1100100)$ . To calculate the distance between  $a$  and  $b$ , we use this measure  $D$  = the sum of the elements of the resulting vector :  $D = 1 + 1 + 0 + 0 + 1 + 0 + 0 = 3$ . The value 3 can be seen as the degree of difference between the trace  $a$  and  $b$ .

The clustering algorithm (*AXOR Algorithm*) that we use in this step is summarize in (Algorithm 1). The user can choose the minimum distance  $D$  as a criteria to select traces to put in the same cluster. In (Algorithm 1) we have specified that all

traces having a distance ( $D < 3$ ) between them, will be put in the same cluster.

### Algorithm 1 AXOR Algorithm

```

Require: Given a cluster C consisting of N traces,
Ensure: Partition the N traces into M sub-clusters,
1: int MinDistance[i][j] = 0;
2: list Clusters;
3: for i:= trace1 to traceN-1 do
4:   for j:= tracei+1 to traceN do
5:     distance(i, j) = tracei xor tracej;
6:     if (MinDistance[i][j] > distance(i, j) and distance(i, j) < 3) then
7:       MinDistance[i][j] = distance(i, j);
8:     end if
9:   end for
10: end for
11: int ClusterNumber = 0;
12: for i:= 1 to N do
13:   if NotClustered(tracei) then
14:     ClusterNumber++;
15:     Clusters[ClusterNumber] = CollectClusterTraces(tracei);
16:   end if
17: end for

```

With our clustering algorithm, the user doesn't need to specify the number of clusters to find. In fact, we use the fitness as a quality measure (see Table III). Fitness is a quality measure that indicates the gap between the behaviour actually observed in the log and the behaviour described by the process model. It gives the extent to which the log traces can be associated with execution paths specified by the process model. If a model has a poor fitness value, this indicates that the mined process model does not successfully parse most of the log traces as mentioned in [16]. In our example, when we apply our clustering technique on the first group of trainees (i.e. the first cluster of the first step) we obtain three clusters (cluster 1.1, cluster 1.2 and cluster 1.3). The figure 9 illustrates the training process models obtained from the cluster 1.1.

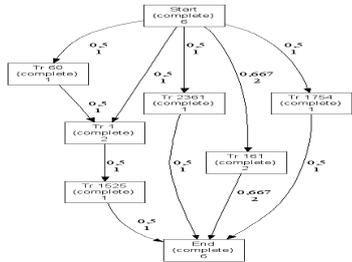
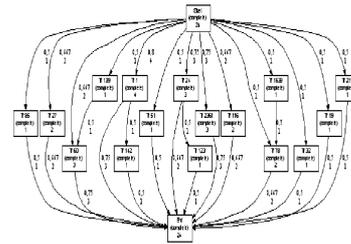


Fig. 9. Training process models obtained from the first cluster of the first step, using the clustering approach (second step)

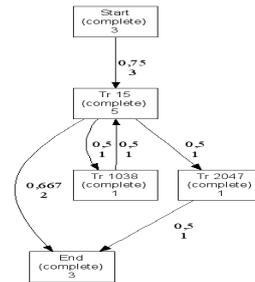
Data sets	Cluster 1.1	Cluster 1.2	Cluster 1.3
Fitness	0.96	1.0	1.0

TABLE II. FITNESS OF THE OBTAINED MODELS FOR CLUSTER 1

In our example, when we apply our clustering technique on the second group of trainees with the best employability (i.e. the second cluster of the first step), we obtain two clusters (cluster 2.1 and cluster 2.2). Figure 10 illustrates the training process models obtained from the two clusters obtained above.



(a) Cluster 2.1



(b) Cluster 2.2

Fig. 10. Training process models obtained from the second cluster of the first step, using the clustering approach (second step)

Data sets	Fitness
Cluster 2.1	1.0
Cluster 2.2	0.94

TABLE III. FITNESS OF THE OBTAINED MODELS FOR CLUSTER 2

When we apply for the third group of trainees with the average employability indicators, we obtain seven clusters, Figure 11 shows the training process model underlying the cluster 3.1.

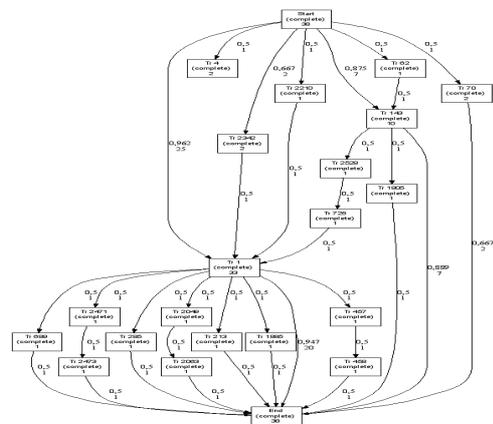


Fig. 11. One of the training process model obtained from the third cluster of the first step, using the clustering approach (second step of our approach)

Data sets	Fitness
Cluster3.1	0.95
Cluster3.2	0.85
Cluster3.3	0.97
Cluster3.4	0.95
Cluster3.5	0.92
Cluster3.6	1.0
Cluster3.7	1.0

TABLE IV. FITNESS OF THE OBTAINED MODELS FOR CLUSTER 2

In summary, figures 9, 10 and 11, show the most typical behaviour of the students in each case/dataset. And we note that the obtained models are simpler and understandable comparing with those obtained in the first step. Also, we obtained a good fitness for each cluster.

#### IV. CONCLUSION

Most of the traditional data mining techniques focus on data dependencies or simple patterns and do not focus on the process as a whole and do not provide visual representation of the complete educational process ready to be analyzed [4]. Process mining is a relatively new technology which emerged from business community. Educational process mining (EPM) allows to get a better understanding of the underlying educational processes.

In this paper, we have proposed a two-step clustering technique. The first step consists of creating clusters of similar trainees' profiles based on a training path performance indicator expressed via two criteria (employability, time period). The second step consists of creating clusters of similar trainees' profiles based on the AXOR Algorithm for each case, in order to improve the mining process and, at the same time, optimize both the performance/fitness and comprehensibility/size of the obtained model. In particular, the comprehensibility of the model is a core goal in education due to the transferral of basic knowledge that it entails.

We found out that, in our work, the second step seems to be the most adequate to partition efficiently training event logs. However there are some questions we have to investigate when partitioning the process mining problem into smaller problems such as how to combine the results of the individual sub-problems into solutions for the original problems. An important point to discuss when using decomposed process discovery, is how to assess the quality of a decomposition before starting the time-consuming actual discovery algorithm. In [11], the authors defined three quality notions (cohesion, coupling and balance) that can be used to assess a decomposition, before using it to discover a model or check conformance with.

Another important step in our works is develop new clustering and classification techniques taking into account semantic annotations on event logs. For instance, trace clustering techniques can be extended to partition event logs depending on trace similarities at the conceptual level. We also intend to develop classification techniques to split semantically annotated event logs based on traces' distance from a set of process models or templates, defined at the conceptual level.

#### REFERENCES

- [1] Hicheur Cairns A., Gueni B., Fhima M., Caiens A., Davis S. and Khelifa N.: *Custom-Designed Professional Training Contents and Curriculums through Educational Process Mining*. In The Fourth International Conference on Advances in Information Mining and Management, pp 53-58 (2014).
- [2] Seber, G. A. F: *Multivariate Observations* Hoboken, NJ: John Wiley & Sons, Inc., 1984.
- [3] Spath, H.: *Cluster Dissection and Analysis: Theory, FORTRAN Programs*, Examples. Translated by J. Goldschmidt. New York: Halsted Press, 1985.
- [4] Pechenizkiy, M., Truka, N., Vasilyeva, E., van der Aalst, W. and De Bra, P.: "*Process Mining Online Assessment Data*", In Proc. EDM'09, pp. 279-288,(2009).
- [5] Ariouat H., Barkaoui K., and Akoka J.: *Improving process models discovery using AXOR clustering algorithm*, International Conference on Information Science and Applications, February 2015, Vol. 339(Springer Berlin Heidelberg), pp.pp 623-629, Series Lecture Notes in Electrical Engineering, Berlin.
- [6] Siemens, G., Baker, R.S.J.d.: *Learning Analytics and Educational Data Mining: Towards Communication and Collaboration*. International Conference on Learning Analytics and Knowledge. 1-3. 2012.
- [7] Trcka, N., Pechenizkiy, M.: *From Local Patterns to Global Models: Towards Domain Driven Educational Process Mining*. International Conference on Intelligent Systems Design and Applications, Milan, Italy, 1114-1119.2009.
- [8] Trcka, N., Pechenizkiy, M. and van der Aalst, W.: *Process Mining from Educational Data (Chapter 9)*", In Handbook of Educational Data Mining. pp. 123-142. CRC Press(2010).
- [9] Ayutaya, N. S. N., Palungsuntikul, P., and Premchaiswadi, W.: *Heuristic mining: Adaptive process simplification in education*. International Conference on ICT and Knowledge Engineering. 221-227.2012
- [10] Song M. , and Gunther C.W., and van der Aalst W.M.P. : *Trace Clustering in Process Mining.*, In: Ardagna, D., Mecella, M., Yang, J. (eds.) BPM 2008. LNBIP, vol. 17, pp. 109-120. Springer, Heidelberg (2009)
- [11] Calders,T., Pechenizkiy,M.: *Introduction to The Special Section on Educational Data Mining*, ACM SIGKDD Explorations Newsletter. (2012)
- [12] Van der Aalst, W. M. P., Andriansyah, A., Alves de Medeiros, A. K., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., Van den Brand, P., Brandtjen, R., and Buijs: "*Process mining manifesto*," In BPM 2011 Workshops Proceedings, pp. 169-194.
- [13] Jagadeesh Chandra J. C. Bose, R. S. Mans and Wil M. P. van der Aalst: *Wanna improve process mining results*, CIDM 2013: 127-134
- [14] Kaufman L., and P. J. Rousseeuw: *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: John Wiley & Sons, Inc., 1990
- [15] Forney G: *Generalized minimum distance decoding*. IEEE Transactions on Information Theory (Volume:12 , Issue: 2 ), 1966
- [16] Rozinat A. and vanderAalst W.M.P. : *Conformance Checking of Processes Based on Monitoring Real Behavior*. Information Systems, Volume 33, Issue 1, Pages 64-95.