CBR-Mining Approach to Improve Learning System Engineering in a Collaborative E-Learning Platform

Fatima Zahra Berriche[1], Besma Zeddini[1], Hubert Kadima[1] and Alain Riviere[2]
[1]*Ecole Int. des Sciences du Traitement de l'Information, Paris-Saclay University, Avenue du Parc, Cergy, France*
[2]*Mechanical Engineering Department, Institut Superieur de Mecanique de Paris, Saint Ouen, France*

Keywords: System Engineering, System Engineering Learning, Process Mining, Case based Reasoning, Collaborative E-Learning Environment.

Abstract: System engineering (SE) is an approach that involves customers and users in the development process and more particularly during the definition of requirements and system functionalities. In order to meet the challenges and increasing complexity of system engineering, the training of engineering students in this field is necessary. It enables learners to acquire sound theoretical and practical knowledge, and to adapt to the majority of profiles of the position related to system engineering field proposed by industrial companies. In this paper, we present a continuity of our research work (Berriche et al., 2015), we study the feasibility of the CBR-mining (case based reasoning and process mining) approach in the context of our platform dedicated to the learning of system engineering. First, we apply the CBR-mining approach to monitor student interactions from log files. Secondly, we propose clusters that bring together all the educational processes most performed by students. We have experimented this approach using the ProM Framework.

1 INTRODUCTION

The design tasks of complex systems follow an exponential growth curve and are increasingly hard. The successful development of complex systems depends on a real identification of the needs and the requirements that the system is supposed to satisfy. Before the emergence of the system engineering concept, complex systems were developed according to the vision of the designers and the engineers without the involvement of the users opinions. This resulted in the abandonment of many complex systems, that were very well built from a technology point of view, but which were considered failures because they didn't meet the requirements of the users.

To manage customer requirements and to cope the competition of the industrial sectors on the world scale. It is necessary to define a methodology and an approach that involves customers and users in the development process and more particularly during the definition of system functionalities. This is the main objective of system engineering.

According to INCOSE (International Council on system engineering), system engineering (SE) is an approach that involves customers and users in the development process and more particularly during the definition of requirements and system functionalities. To adapt to the complexity of today's industrial systems, the training of engineering students in this field is necessary. It allows learners to acquire sound theoretical and practical knowledge and to adapt to the majority of job profiles in the field of system engineering proposed by industrial companies.

Our research is carried out by the PLACIS project (Collaborative Platform for Systems Engineering) Financed by the Ministry of Higher Education and Research in the context of IDEFI projects (Alexis and Antoine, 2014). The general objective of the project is to ensure that the majority of students receive thorough preparation for system engineering.

The PLACIS project provides an environment for project-based learning of system engineering. It is an environment for building and sharing traces of learners. In this context, we proposed an approach for the analysis of massive e-learning traces.

The remainder of this paper is structured as follows, in section 2 we provide an overview of collaborative e-learning environments for learning system engineering and describe briefly the PLACIS platform. Section 3 presents our approach CBR-

Berriche, F., Zeddini, B., Kadima, H. and Riviere, A.
CBR-Mining Approach to Improve Learning System Engineering in a Collaborative E-Learning Platform.
DOI: 10.5220/0006693200940101
In *Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2018)*, pages 94-101
ISBN: 978-989-758-297-4

mining and its application in the system engineering educational do- main. In section 4 we discuss the applicability of our approach in the context of PLACIS. The final section 5 (Conclusion and future works) gives a summary of the main findings of this study according to the challenges mentioned above and outlines perspectives for future work.

2 COLLABORATIVE E-LEARNING FOR SYSTEM ENGINEERING

2.1 Collaborative E-Learning Platform

In the field of learning and community-based knowledge management, collaboration is based on a common goal, each member performing a part of the overall task, drawing on environmental resources (organizational memory), in its own resources (individual competence) and those of the group: this is called a collaborative learning.

Collaborative e-learning is not simply an association of a collaborative aspect with the e-learning environment phenomena. Indeed, collaborative e-learning environments can be described as a context where information technology facilitates interaction between learners for knowledge acquisition or sharing (Liaw and Huang, 2007) and access to e-learning content from different systems (Masud, 2016). It could be described as ”a variety of educational practices in which peer interactions are the most important factor in learning, although without excluding other factors such as the learning materials and interactions with teachers” (Dillenbourg et al., 2009). Collaborative learning is a key paradigm for informal learning (for example, sharing knowledge among communities of practices), but has been somewhat underutilized in a context of engineering systems.

The use of collaborative e-learning for learning system engineering continues to increase in exponential fashion. The expanding volume of system engineering knowledge imposes more challenges on the teacher and learner in system engineering field. To assist teachers with this information overload, there is an e-learning related solution for system engineering education. The e-learning solution allows teachers and educators to provide a shared workplace for students to interact and learn through cooperation (Li et al., 2008). The interaction could be formed in many ways like:

- Interactions with members of the groups they are assigned to (assigned group);
- Interactions with members of the larger community that is the class (class group);
- Interactions with peers from the discipline’s community (discipline community) (Doyle et al., 2015).

Therefore, among the collaborative e-learning for learning system engineering, the PLACIS platform that will be discussed in more detail in the next section.

2.2 PLACIS Platform

PLACIS is a project funded by the French National Agency for Research under Investments for the future program. PLACIS is the French acronym for this project and this is the official name of this project. It started in september 2012 and is managed by the IPGP l’Institut Polytechnique Grand Paris (composed of three engineering schools). This large-scale project promotes active learning and teaching through industrial, educational, international and at-a-distance collaborative projects, carried out by engineer students (Alexis and Antoine, 2014).

The work presented in this article is made in the context of PLACIS platform. The overall objective of the project is to ensure that the majority of students receive a thorough preparation in system engineering and a quasi-professional know-how based on the most recent tools, notably on the modules of collaborative systems, as well as about group work experience.

The PLACIS platform emphasizes on the technique PROJECT-BASED LEARNING (Bougaa et al., 2017), the technique allows to groups of students from several distant universities to participate in international projects as in the professional context. It therefore implements teaching methods based on participatory observation, learning by doing and sharing experience. Our experimental platform for learning and teaching system engineering can be controlled by several system engineering standards, for example (IEEE 1220, EIA 632, ISO/IEC 15288, ISO/CEI 29110).In the context of our research, we have chosen as a case of study the standard ISO/IEC 15288 which is a systems engineering standard covering processes and life cycle stages.

It defines a set of processes and associated terminology from an engineering viewpoint. These processes can be applied at any level in the hierarchy of a systems structure. Selected sets of these processes can be applied throughout the life cycle for managing and performing the stages of a systems life cycle (ISO, 2015). The standard provides a mapping of the lifecycle processes of the systems. It includes the design, development, production, use, resources and withdrawal of a system, as well as the improvement of life-cycle processes.

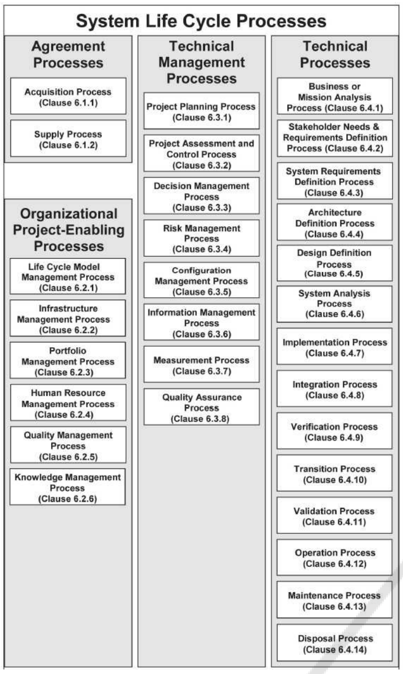

Figure 1: System Life cycle processes (ISO, 2015).

As shown in the figure 1, the standard defines processes divided into four categories: Technical, Project, Agreement, and Enterprise. Each process is defined by a purpose, outcomes, and activities. ISO 15288 comprises 25 processes which have 123 outcomes derived from 403 activities.

3 CASE BASED REASONING AND PROCESS MINING

3.1 Case Based Reasoning

The Case-Based Reasoning (CBR) was introduced in 1975 by Minsky. In 1994, Kolodner formalized the foundation of CBR in his book (Domeshek and Kolodner, 1993) and in the same time Aamodt and Plaza formalized the CBR cycle (Aamodt, 2002). The CBR is a process that involves the reuse of past experiences, and was used in expert systems and cognitive science. In this approach, the user trying to solve a new problem recognizes similarities with previously solved problems called cases. A case is a commonly specific problem that has been identified, resolved, stored and indexed in a memory with the solution, and optionally the process for obtaining it (Watson, 1999) (Robles, 2006). The CBR systems are applied in many fields such as: medicine, trade, industrial diagnosis, monitoring and financial analysis.

The lifecycle of CBR

The lifecycle of a CBR is composed of four steps: Retrieve, Reuse, Revise, and Retain:
Retrieve: Given a target problem, retrieve from memory cases that are relevant to solving it. A case consists of a problem, its solution, and, typically, annotations about how the solution was de- rived.
Reuse: Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation.
Revise: Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, re- vise.
Retain: After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory.

3.2 Process Mining

Process mining aims at discovering, monitoring and improving real processes by extracting knowledge from event logs available in todays information systems. Log files may contain data recording three perspectives: behavioral (tasks and their time of execution), informational (data used and produced by tasks) and organizational (actors which perform tasks and their relationships). Figure 2 exposes the three types of process mining.

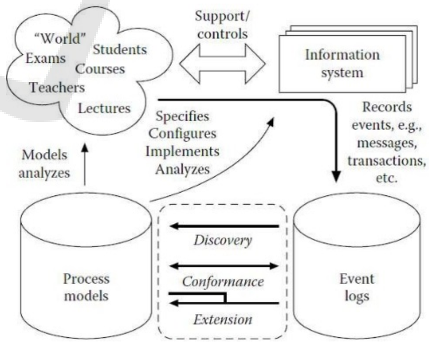

Figure 2: Three types of process mining: Discovery, Conformance, and Extension (Enhancement).

- Discovery: Produces a process model from event logs without using any a priori information. The discovered models can be represented by formal languages such as Petri nets or notations such as Business Process Model and Notation (BPMN).
- Conformance: Check if reality conforms to

the process model: compare an existing process model with an event log of the same process. It is also used to detect, locate and explain deviations, and to measure the severity of these deviations.

- Enhancement: Extends or improves an existing process model using information about the actual process recorded in some event log.

3.3 Combining Case Based Reasoning and Process Mining

The CBR solves problems by finding similar cases in its knowledge base and adapting to the particular case. This solving is made of a number of phases: case representation, indexing, similarity compare-son, retrieval and adaptation. In the literature, several contributions to the combination of data mining techniques with the CBR process were proposed Arshadi and Jurisica have combined CBR with data mining to improve case-based classification (Arshadi and Jurisica, 2005), Kim and Han proposed a new method of indexation and classification based cases, based on the use of artificial neural networks competitive (Kim and Han, 2001), Beddoe and Petrovic used genetic algorithms to facilitate selecting and weighting features to personnel rostering (Beddoe and Petrovic, 2006).

The character of system engineering process oriented us towards the use of process mining, because it is a new method that has been recently used in various domains, such as healthcare (Anyanwu et al., 2002), time prediction (Backus et al., 2006) and in various other areas for analyzing resource behavior (van der Aalst et al., 2008). But it has not yet been applied in the domain of system engineering education. In our work, we decide to adopt an approach that combines case based reasoning with process mining in order to improve learning system engineering.

4 APPLIYING CBR-MINING TO PLACIS PLATFORM

The context of application to which we are interested is the e-learning environments. PLACIS offers a platform for learning system engineering based on standards specifically dedicated to system engineering. It is considered as a medium for disseminating knowledge and interaction between the different ac- tors (teachers, learners, etc.).

The use of this platform by learners generates a large amount of data, which may contain information about user profiles, Information about the concepts of system engineering, problems treated by the projects deployed on the platform, decisions taken, and the constraints considered.

The purpose of this article is to make PLACIS an intelligent and adaptive e-learning platform able to analyze the profiles of each student in the platform and use them to: on one hand understand learners' learning habits as well as the skills they have acquired and on the other hand to adapt future projects integrated on the platform to the knowledge of the learner. Given that our platform is focused Project-based learning, we chose to experiment as a case study, a project realized under the platform by several groups of students. The aim is to enable learners to learn system engineering using the standard ISO 15288. In order to familiarize themselves with this standard, they applied some processes related to the ISO/IEC 15288 standard for the design of a surveillance drone. Among these processes, we cite: Stakeholder needs and requirements definition process; System requirements definition process; Architecture definition process.

In this context, we decided to apply our approach that combines CBR and process mining (as shown in Figure 3).

We assume that the database used is relevant and it covers a large number of system engineering projects. As the figure shows, we focus only on the retrieving step in the CBR process to extract similar projects; and on the reuse step in which we used process mining to monitor student interactions and to propose clusters that bring together all the educational processes most performed by students

A search algorithm for similar projects must be developed using the CBR approach to search for similar cases. In our case the data are structured in the form of system engineering projects, each project contains activities of standard 15288. And every activity encompasses several tasks. The database contains also profiles of the student attached to each project. This data are stored in various databases.

First, we build a consolidated journal (stored as a CSV file). Given that we use the ProM framework for process discovery and analysis, we transform this log into MXML (Mining eXtensible Markup Language) using the ProM Import plug-in.

To illustrate this we show an extract from our MXML file (Figure 4), it contains information about integrated projects in the platform. For each project a group of students were assigned and a set of activities and tasks were attributed. Students are supposed to returns the projects in time.

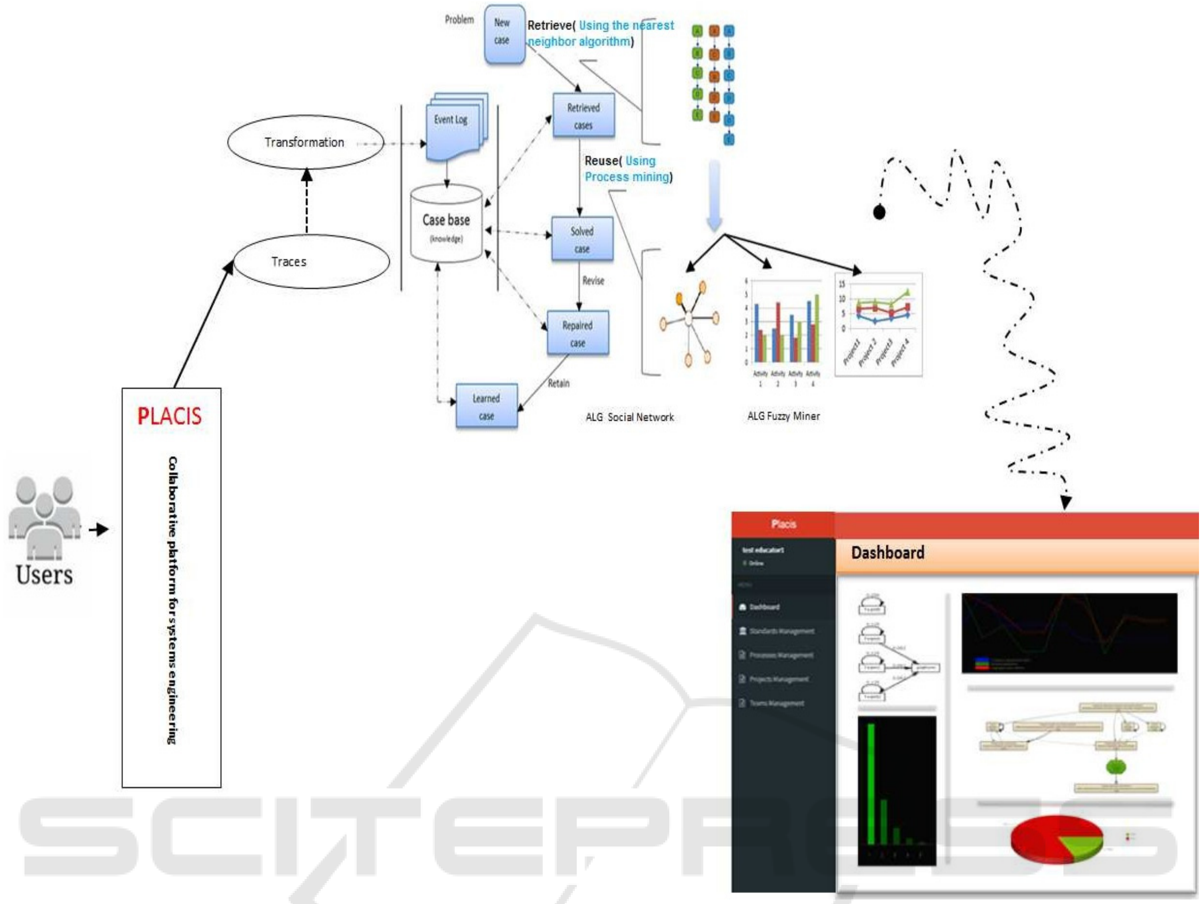

Figure 3: CBR-Mining approach.


```
<ProcessInstance id="5" description="Design process instance">
<AuditTrailEntry>
<caseId> http://placis.bel4.com/</caseId>
<WorkflowModelElement>Prepare for stakeholder needs and requirements definition</WorkflowModelElement>
<EventType>Identify the stakeholders who have an interest in the system throughout its life cycle. </EventType>
<Timestamp>2017-09-30T14:32:00.000+01:00</Timestamp>
<Originator>TeamC</Originator>
</AuditTrailEntry>
<AuditTrailEntry>
<WorkflowModelElement>TeamC</WorkflowModelElement>
<EventType>download</EventType>
<Timestamp>2017-09-30T16:00:00.000+01:00</Timestamp>
<Originator>platform</Originator>
</AuditTrailEntry>
```


Figure 4: Screenshot of the MXML event log.

In order to extract information from this file and to test the applicability of our approach, we apply the mining algorithm Social Network Miner. It is an algorithm of process mining techniques to analyze the social network from the event logs such as finding the pattern of work collaboration of people in the organization. The technique provides five kinds of metrics to generate social networks (van der Aalst et al., 2005) (as shown in Figure 5):

Handover of Work Metric: This metric determines who passes work to whom. This information can be extracted from an event log finding subsequent activities in the same case (i.e., process instance), where the first activity is completed by one individual and the second one is completed by another individual.

Subcontracting Metric: While in the previous metric relationship between two individuals is unidirectional, in this one it is bidirectional. Considering a single case of an event log and two individuals, we know that individual i subcontract individual j, when in-between two activities executed by individual i there is an activity executed by individual j.

Working Together Metric: metrics help us to identify the resources which are used to work together. They ignore causal dependencies and focuses on resources which work together for the same case.

Similar Task Metric: The main idea is to determine who performs the same type of activities. To do so, each individual has his own profile based on how frequently he conducts specific activities. Then the profiles are compared to determine the similarity.

Reassignment Metric: The basic idea of this metric is to detect the reassigning of activities from one individual to another : if i frequently delegates work to j but not vice versa it is likely that i is in a higher hierarchical than j.

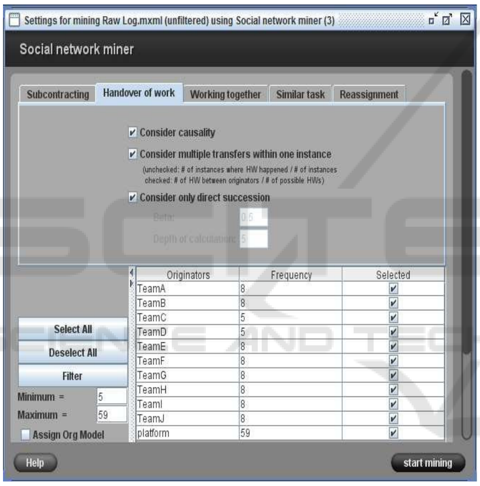

Figure 5: Social Network Miner.

Furthermore, we used Handover of work metric of the social network miner technique. The technique has enabled us to evaluate the interactions between the students and the platform, the aim is to examine the degree of involvement of students in projects. Figure 6 shows the resulting Social Network Miner graph with respect to the handover of work between platform (Teachers) and 10 groups of students, each group has 4students. As shown in figure 6, all of the team A, Team C,D,E,F,G,H,I,J have already passed at least one Activity/task to platform (Teachers), on the contrary Team B didn't return their work. The loop over each team represents the fact that in some situations, the Teams took care of the two different types of tasks (i.e., such as pre-test and posttest) by themselves without any collaboration and contribution with teacher. Nevertheless, Platform

	TeamA	TeamB	TeamC	TeamD	TeamE	TeamF	TeamG	TeamH	TeamI	TeamJ	platform
TeamA	0.0517241...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0344827...
TeamB	0.0	0.1034482...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TeamC	0.0	0.0	0.0517241...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0172413...
TeamD	0.0	0.0	0.0	0.0517241...	0.0	0.0	0.0	0.0	0.0	0.0	0.0172413...
TeamE	0.0	0.0	0.0	0.0	0.0517241...	0.0	0.0	0.0	0.0	0.0	0.0344827...
TeamF	0.0	0.0	0.0	0.0	0.0	0.0517241...	0.0	0.0	0.0	0.0	0.0344827...
TeamG	0.0	0.0	0.0	0.0	0.0	0.0	0.0517241...	0.0	0.0	0.0	0.0344827...
TeamH	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0517241...	0.0	0.0	0.0344827...
TeamI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0517241...	0.0	0.0344827...
TeamJ	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0517241...	0.0344827...
platform	0.0172413...	0.0	0.0172413...	0.0172413...	0.0172413...	0.0172413...	0.0172413...	0.0172413...	0.0172413...	0.0172413...	0.0

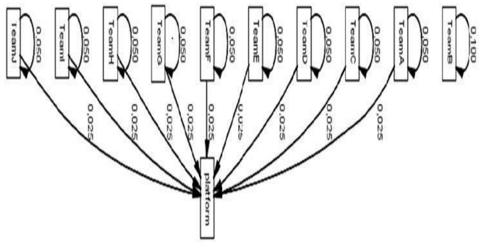

Figure 6: Social Network Miner graphs.

(Teacher) has already validated at least one activity/task for groups that have already submitted their work on the platform.

As we have already mentioned previously the approach that we have proposed allows also to cluster all activities/task the most realized by groups, and to adapt future projects integrated in the platform to the level knowledge of the learner. For that reason, we decided to use the Fuzzy Miner algorithm (Gnther and van der Aalst, 2007). Fuzzy Miner is a configurable discovery algorithm that allows, through its parameters, the generation of multiple models at different levels of detail, helping to deal with unstructured processes.

It clusters cohere activities with log significance but high correlation. Furthermore, the fuzzy miner enables to determine which activities and arcs need to be included based on an extensible set of configurable parameters and is therefore considered to be hierarchical. The figure 7 shows the resulting process model from our log file. The square nodes rep- resent the event classes, their importance (the maximum value is 1.0) is provided as the event class in each node, and the thickness of an arc indicates the frequency. This model shows that activities related to the design cycle, and more particularly the activities related to the stakeholder needs and requirements definition process and system requirements definition process (Prepare Stakeholder needs and requirement definition; Define stakeholder needs; analyze stakeholder requirements; prepare for system requirements definition; analyze system requirements) are the most frequent events fulfilled by student groups. However, fuzzy miner clusters have also less frequent activities, highly correlated and are represented in aggregated form as clusters. The others activities of the project with less correlated behavior are eliminated from the model process.

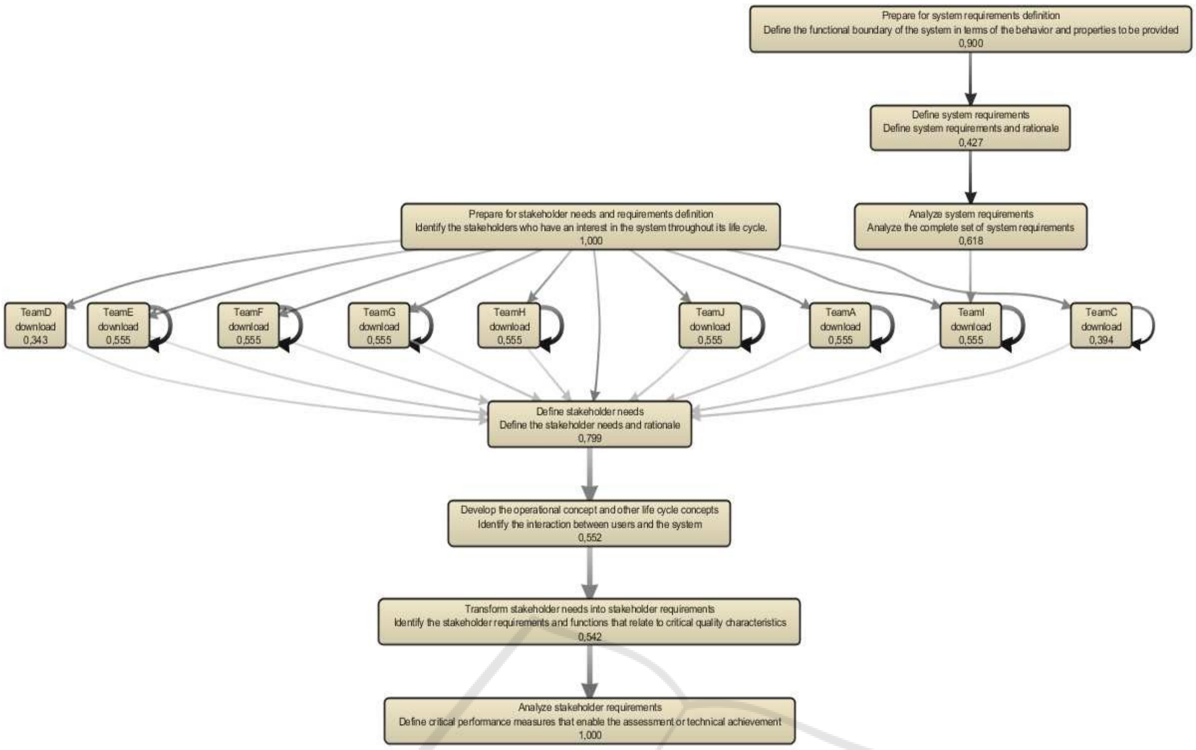

Figure 7: Screenshot of the Fuzzy Miner.

5 CONCLUSIONS

In this paper, we have studied the potential of CBR-Mining approach in the educational domain, especially the education of system engineering. The first step consists of using social mining techniques (implemented in ProM Framework) to examine and assess interactions between platform (teachers) and the group of students. The technique allowed us to better visualize the handover of work from Originator A to Originator B. The second step consists of creating clusters that bring together all the educational processes most performed by students. This study has been implemented using the Fuzzy Miner plugin of ProM framework.

Our future work will continue in several directions. We intend to test the applicability of our approach CBR-Mining with other process mining techniques. Among the objectives of our future work is to focus on applying clustering techniques of learners, in order to propose to the teacher a possible distribution of the students in different groups. And also to propose techniques which allow discovering interaction patterns from email datasets(van der Aalst and Nikolov, 2008) in order to discover interactions patterns between students in their collaborative learning tasks, communication actions and online discussions. Moreover, we also intend to develop ways of evaluating students in order to restore the knowledge that can be used by teachers. To conclude, we plan to design indicators and interfaces in order to test the traces of learners and display them in dashboards. The idea is to allow learners and teachers to visualize all of the platform's projects (current projects, closed projects, successful projects) and to reflect on how they have accomplished their work.

ACKNOWLEDGEMENTS

This paper reflects only the authors' views. The work reported is carried out by the PLACIS project, which has received research funding from French Ministry for Higher Education and Research.

REFERENCES

Aamodt, A. (2002). A knowledge representation system for integration of general and case-specific knowledge. In *Proceedings Sixth International Conference on Tools with Artificial Intelligence*, New Orleans, LA, USA,

USA. IEEE.Alexis, F. and Antoine, L. (2014). Work-in-progress: Col- laborative platform for systems engineering. In *In- ternational Conference on Interactive Collaborative Learning (ICL)*, Dubai. IEEE.

Anyanwu, K., Sheth, A., Cardoso, J., Miller, J., and Kochut, J. (2002). Healthcare enterprise process development and integration. *Journal of Research and Practice in Information Technology*.

Arshadi, N. and Jurisica, I. (2005). Data mining for case-based reasoning in high-dimensional biological domains. *IEEE Transactions on Knowledge and Data Engineering*, 17:1127 – 1137.

Backus, P., Janakiram, M., Mowzoon, S., Runger, C., and Bhargava, A. (2006). Factory cycle time prediction with a data-mining approach. *IEEE Transactions on Semiconductor Manufacturing*, 19:252 – 258.

Beddoe, G. and Petrovic, S. (2006). Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering. *European Journal of Operational Research*, 175:649–671.

Berriche, F., Zeddini, B., Kadima, H., and Riviere, A. (2015). Combining case-based reasoning and pro- cess mining to improve collaborative decision-making in products design. In *12th ACS/IEEE International Conference on Computer Systems and Application AICCSA*, Marrakech. IEEE.

Bougaa, M., Bornhofen, S., Riviere, A., and Tucoulou, J. (2017). Toward a process-centered approach for systems engineering education. In *ASEE Annual Conference and Exposition*, Colombus, Ohio, United States.

Dillenbourg, P., Jarvela, S., and Fischer, F. (2009). The evolution of research on computer supported collaborative learning: From design to orchestration. *Technology enhanced learning: Principles and Products*.

Domeshek, E. and Kolodner, J. (1993). Using the points of large cases. *Artificial intelligence for engineering design analysis and manufacturing (AI EDAM)*, 7:87–96.

Doyle, C., Sammon, D., and Neville, K. (2015). Building an evaluation framework for social media-enabled collaborative learning environments (smecles). *Journal of Decision Systems*, 24(3):298–317.

Gnther, C. and van der Aalst, W. P. (2007). Fuzzy mining: Adaptive process simplification based on multi perspective metrics. In *BPM: International Conference on Business Process Management*, Berlin, Heidelberg. Springer.

ISO (2015). Systems and software engineering system life cycle processes. *International Organization for Standardization*, pages 1–108.

Kim, K. and Han, I. (2001). The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases. *Expert Systems with Applications*, 21:147–156.

Li, Q., Lau, R., Shih, T., and Li, F. (2008). Technology supports for distributed and collaborative learning over the internet. *ACM Transactions on Internet Technology*, 8(2):5:1–5:24.

Liaw, S. and Huang, H. (2007). Developing a collaborative e-learning system based on users perceptions. In *International Conference on Computer Supported Cooperative Work in Design*. Springer -Verlag Berlin Heidelberg.

Masud, M. (2016). Collaborative e-learning systems using semantic data interoperability. *Computers in Human Behavior*, 61:127–135.

Robles, G. C. (2006). *Management de linnovation technologique et des connaissances: Synergie entre la thorie TRIZ et le Raisonnement Partir de Cas. Application en gnie des procds et systmes industriels*. PhD thesis, Universite de Toulouse.

van der Aalst, W., Nakatumba, J., Rozinat, A., and Russell, N. (2008). Business process simulation: how to get it right? page 25.

van der Aalst, W. M. P., Reijers, H. A., and Song, M. (2005). Discovering social networks from event logs. *Computer Supported Cooperative Work (CSCW)*, 14:549593.

van der Aalst, W. P. and Nikolov, A. (2008). Emailanalyzer: An e-mail mining plug-in for the prom framework. In *International Conference on Business Process Management*.

Watson, I. (1999). Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems*, 12:303–308.