Proceedings of the 34th European Modeling & Simulation Symposium (EMSS), 038
19th International Multidisciplinary Modeling & Simulation Multiconference

ISSN 2724-0029 © 2022 The Authors.
doi: 10.46354/i3m.2022.emss.038

Assignment model for courses online

Alejandro Felipe Zárate Pérez[1], Idalia Flores de la Mota[1]

[1] Universidad Nacional Autónoma de México, Av. Universidad 3004, Col Copilco Universidad, Coyoacán, Ciudad de México,04510, México

[*]Corresponding author. Email address: alejandro.zarate@unam.mx

Abstract

One of the biggest problems in education, particularly in e-learning, is the high dropout rate that occurs in courses, so this paper presents a mathematical optimization model to maximize the completion efficiency of all scheduled courses from the assignment of teachers in each course. This will be achieved using historical data to the teachers in previous courses and observations of courses opening rate in previous years. With this data and the administrative requirements of the institution, a linear programming assignment model will be used to maximize the student completion rate for the entire scheduled period.

Keywords: optimization, teaching analytics, assignment model, e-learning analytics

1. Introduction

In the last thirty years, technology has been advancing faster and faster and this has also been reflected in education, an example of this is the modality called "Distance Education", which has gained more strength, as it is more accessible thanks to innovative technologies. This modality has the purpose to give education access to diverse sectors that have not been able to be attended, due to situations such as geographic, employee, time, among others.

Nowadays, with the incorporation of ICT (Information and Communication Technologies), it is possible to glimpse the scope that these represent for distance education, thus playing an essential role, because of the application of these recent technologies to the educational and training field, what is called "e-learning".

E-learning is a way of using ICT as a means of distribution for educational materials and other services, in which there is also an interrelation between teachers and students. Thus, in this new teaching-learning environment, web technology is used through the Internet.

Within education we find two types of education: academic and continuing education. In this paper we will focus only on the second: continuing education.

There are several definitions of continuing education, some of which vary according to the country to which we refer. However, for this paper we will take the UNAM definition of continuing education[1]:

> It is an educational modality designed, organized, systematized, and programmed that complements the curricular formation and deepens and broadens knowledge in all fields of knowledge; it trains and updates professionally and is aimed at the university community and the general public.

The Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) part of Universidad Nacional Autónoma de México, offers online continuing education courses, mainly in computing, through the Coordination of Continuing Distance Training, which are aimed to the public, the university community and institutions and companies that request them.

[1] http://abogadogeneral.unam.mx/legislacion/abogen/documento.html?doc_id=87

Online education, particularly online continuing education, has serious problems with the dropout of students who enroll in courses. The causes of this dropout can be diverse: lack of technological knowledge, deficiencies in didactic content, inadequate technological platforms, unsuitable teachers, etc.

In this paper the problem approach is from the teachers' perspective, focusing mainly on the work they do during a course.

2. State of the art

At a technical level, we can affirm that all the systems necessary to teach online courses use a database for their correct operation, in which the information of the lessons, activities and even the participants' grades are stored.

However, it not only stores didactic information, but also collects information on the interactions of the participants within the platform, this information can be useful because through it we can identify how the teaching-learning process is developing.

2.1. Learning analytics

Learning analytics is a relatively new concept, so there are several definitions about it, for this paper we will use the definition given by Durall et all in 2012 in "The Horizon Report of The New Media Consortium and the Universitat Oberta de Catalunya", which says the following:

> Learning analytics consists of the interpretation of a wide range of data produced and collected about students to guide their academic progression, predict future performance, and identify problematic elements. The purpose of collecting, recording, analyzing, and presenting this data is to enable teachers to quickly and effectively adapt educational strategies to the level of need and ability of each student. Even in their early stages of development, learning analytics respond to the need to monitor and control campus activity for strategic decision making. On the other hand, they aim to take advantage of the large amount of data produced by students in academic activities.

With this definition we will make the necessary considerations for the definition, collection, analysis, and prediction of the data of this paper.

The delivery of online courses is done through a Moodle LMS, the entire course is conducted within this platform, so all activities are recorded within the system database, among these activities we have review of materials, delivery of assignments, participation in forums, etc.

A teacher is designated for each course, who oversees answering questions, encouraging student participation, and evaluating assignments.

All activities within the course must be completed within the period determined for the course, therefore, each activity is planned to be delivered according to the planning of each course. Based on the instructional design of each course, they are planned in such a way that one topic per week is reviewed, which implies that the weeks of each course correspond to the topics contained in each one of them.

Therefore, each course taught must follow a schedule for delivery and review of activities by the course participants: students and advisor. This schedule must be followed for the course to end smoothly, since a delay on the part of students or advisor can result in the course being abandoned, which means that only some students complete their course satisfactorily.

This is one of the biggest problems of education, and online education is not the exception, the high levels of dropouts that occur during each course, and there are cases where it reaches alarming levels.

Therefore, we have designed our own methodology to monitor these courses, defining the concept of Terminal Efficiency (TE) (Zarate Perez & Flores de la Mota, 2021), which is defined as:

(1)

$$TE = \frac{CAMS}{TCA}$$

Where:

CAMS: course activities made by the students.

TCA: total course activities

Therefore, at the end of the course we calculate the terminal efficiency of the course by the teacher in charge of the course.

Before continuing, we show an example of the Terminal Efficiency.

Let's suppose a problem with 5 students in a course with 4 topics, por each topic a student must do 1 forum participation, 1 homework and 1 questionnaire. Observe that for each student must do 12 activities (4 themes y 3 activities by topic).

So, the TCA result is:

TCA = (5 students) * (4 topics) (3 activities)

TCA = 60

At the end of this course the results of the students are:

- 3 students made all activities.
- 1 student made only 3 topics completely.

- 1 student did not make anything

So, the CAMS result is:

CAMS = 36 (3 students with all activities) + 9 (1 student with only 3 topics) + 0 (1 student with nothing)

CAMS = 45

With the TCA and CAMS results we can calculate the TE's course:

TE = 45 / 60 = 0.75

This TE's course implies that only the 75% of the scheduling activities were made it.

It's not the objective of this paper but we can demonstrate that the teacher and the TE's course are correlated.

Since the teacher influences the terminal efficiency of a course then the teacher who is appointed to teach a course may have a higher or lower terminal efficiency for the assigned course.

Therefore, if we need the courses to have the highest terminal efficiency possible then we must appoint professors with the highest terminal efficiencies for the courses taught.

3. Methodology

To the teaching of courses, there is a planning of them to be scheduled on a regular basis. This course planning is shown monthly to site visitors, but it is a process that is done on a semester basis, i.e., all courses for the semester have already been scheduled to start in a specific period.

Up to now, this planning has only followed criteria such as: courses without repeating them twice in a row, courses of one type of each period, etc. In other words, only administrative criteria.

The assignment model we propose incorporates the terminal efficiency (TE) of the advisors and a course open rate (COR).

This will allow us to assign teachers with the best terminal efficiency to courses with the highest probability of opening, to improve online course completion rates from the planning stage.

The terminal efficiency of the assessors is defined in (Zarate Perez & Flores de la Mota, 2021) but the CAR is not, so we must define it below.

We will define this Course Opening Rate (COR) as the probability that a course X has been opened once it has been programmed, calculated using the following equation:

(2)

$$COR = \frac{Number\ of\ times\ \ the\ course\ is\ opened}{Number\ of\ times\ the\ course\ was\ scheduled}$$

Therefore, the value of this is between zero and one. Where one indicates that the course is opened every time it is programmed and zero indicates that it has never been opened even though it has been programmed once.

3.1. Optimization model

With the data on teacher terminal efficiency and course openness rate we will model course planning and incorporate the terminal efficiency and openness rate into the objective function.

As mentioned before, the purpose of the above is not only to conduct an adequate planning of the courses according to the proposed conditions, but also to directly assign the best advisors (in terms of terminal efficiency) in a course and period to achieve the best terminal efficiency in the proposed periods.

3.2. Variables

To solve this problem, we will use a linear programming model, in particular a binary integer linear programming model. For which we define the following variables:

(3)

x_{kij}: Binary variable for the course i in the period j with the teacher k

$\forall\ i \in \{1,2,\ldots,n\_i\ \}$

$\forall\ j \in \{1,2,\ldots,n\_j\ \}$

$\forall\ k \in \{1,2,\ldots,n\_k\ \}$

In addition, we define the following variables:

$$a_k = Terminal\ efficiency\ for\ the\ teacher\ k$$

$$p_i = \text{Course openning rate for the course } i$$

3.3. Objective function

(4)

$$Max\ z\ = \sum_{n=1}^{n_j} \sum_{i=1}^{n_i} \sum_{k=1}^{n_k} p_i a_k x_{ijk}$$

3.4. Constraints

Exactly one teacher by course and period

(5)

	Period			
Course	$\sum_{k=1}^{n_k} x_{11k} \leq 1$	$\sum_{k=1}^{n_k} x_{12k} \leq 1$	....	$\sum_{k=1}^{n_k} x_{1n_jk} \leq 1$
	$\sum_{k=1}^{n_k} x_{21k} \leq 1$	$\sum_{k=1}^{n_k} x_{22k} \leq 1$	....	$\sum_{k=1}^{n_k} x_{2n_jk} \leq 1$
	....	....	....	....
	$\sum_{k=1}^{n_k} x_{n_i1k} \leq 1$	$\sum_{k=1}^{n_k} x_{n_i2k} \leq 1$	....	$\sum_{k=1}^{n_k} x_{n_in_jk} \leq 1$

The following restrictions have to do with category of courses, to define it in a general way we suppose that we have n_j courses, which are grouped in m categories, that is, we have the categories $c_1, c_2, \ldots, c_m$.

Therefore, the size of each category will be:

$$n_{c_1}, n_{c_2}, \ldots, n_{c_m}$$

With these definitions we proceed to define the restrictions of the courses by category.

Up to y_1 courses in category 1 per period

(6)

$$\sum_{k=1}^{n_k}\sum_{i=1}^{n_{c_1}} x_{i1k} \leq y_1$$

$$\sum_{k=1}^{n_k}\sum_{i=1}^{n_{c_1}} x_{i2k} \leq y_1$$

...

$$\sum_{k=1}^{n_k}\sum_{i=1}^{n_{c_1}} x_{in_jk} \leq y_1$$

At least y2 courses of category 2 per period

(7)

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_1}+1}^{n_{c_2}} x_{i1k} \geq y_2$$

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_1}+1}^{n_{c_2}} x_{i2k} \geq y_2$$

...

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_1}+1}^{n_{c_2}} x_{in_jk} \geq y_2$$

Therefore, in general we can have as many constraints for categories as necessary, the term "Up to" and "At least" will define the sense of equality.

Then, in general, we can define the set of constraints with the term "Until" for any category as follows:

(8)

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_{m-1}}+1}^{n_{c_m}} x_{i1k} \leq y_m$$

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_{m-1}}+1}^{n_{c_m}} x_{i2k} \leq y_m$$

...

$$\sum_{k=1}^{n_k}\sum_{i=n_{c_{m-1}}+1}^{n_{c_m}} x_{in_jk} \leq y_m$$

Or with the inequality inverted if the term is "At least".

The following are the constraints that a course may not repeat during h periods.

For the general definition of the constraints, we define d as the result of n_j % h, so h must always be less than n_j.

A course doesn't repeat per h periods.

(9)

	Period			
Couses	$\sum_{j=1}^{h}\sum_{k=1}^{n_k} x_{1jk} = 1$	$\sum_{j=h}^{2h}\sum_{k=1}^{n_k} x_{1jk} = 1$	....	$\sum_{j=n_j-d-1}^{n_j}\sum_{k=1}^{n_k} x_{1jk} = 1$
	$\sum_{j=1}^{h}\sum_{k=1}^{n_k} x_{2jk} = 1$	$\sum_{j=h}^{2h}\sum_{k=1}^{n_k} x_{2jk} = 1$	....	$\sum_{j=n_j-d-1}^{n_j}\sum_{k=1}^{n_k} x_{2jk} = 1$
	....	....	....	....
	$\sum_{j=1}^{h}\sum_{k=1}^{n_k} x_{n_ijk} = 1$	$\sum_{j=h}^{2h}\sum_{k=1}^{n_k} x_{n_ijk} = 1$	....	$\sum_{j=n_j-d-1}^{n_j}\sum_{k=1}^{n_k} x_{n_ijk} = 1$

Finally, we have the constraints about to the teacher, these constraints limit the number of courses (n_c) that an advisor can have for the scheduling plan.

Maximum number of course schedule to a teacher

(10)

$$\sum_{j=1}^{n_j}\sum_{i=1}^{n_i} x_{ij1} \leq n_c$$

$$\sum_{j=1}^{n_j}\sum_{i=1}^{n_i} x_{ij2} \leq n_c$$

...

$$\sum_{j=1}^{n_j}\sum_{i=1}^{n_i} x_{ijn_k} \leq n_c$$

In case a maximum number of courses is not required, but a minimum number of courses, it is only necessary to invert the inequality.

As can be seen, the number of restrictions and variables depends entirely on the number of courses, periods, and advisors we have, in general we can say that:

Variables quantity= $n_i * n_j * n_k$

4. Results

To exemplify how the model works, we will use a problem with the following characteristics:

- Five courses and five consultants are available.
- The courses are grouped into two categories as follows:

Category 1 = {Course 1, Course 2}

Category 2 = {Course 3, Course 4, Course 5}

- There are four periods of courses
- For category 1, there is just one course per period.
- For category 2 there must be at least one course per period.
- If a course is schedule in actual period, so this course must not be schedule for the next period.

In the cases of terminal efficiency and openness rate we have the following data:

TE = {0.8, 0.9, 0.88, 0.97, 0.91}

COR = {0.9, 1, 0.79, 0.95, 0.85}

The linear programming model was proposed following the equations of the theoretical model; it is not included here due to its length, since it has 100 decision variables. The result of this model is presented in the following table:

	Period 1	Period 2	Period 3	Period 4
Course 1	Adviser 2		Adviser2	
Course 2		Adviser 4		Adviser 4
Course 3	Adviserr 1		Adviser 1	
Course 4	Adviserr 5		Adviser 5	
Course 5		Adviserr 3		Adviser 3

With this resulting assignment, the terminal efficiency for the possible open courses would be maximized, which would increase the number of students completing a course.

At the end of the programmed period, it will be necessary to recalculate the TE and the COR, since these will change according to the results obtained, so it will be advisable to use the latest available data for the next course planning.

5. Conclusions

The model proposed solves all the conditions proposed from the beginning: administrative, efficiency and probability. However, it presents difficulties at the time of solving it, since with few advisors, courses, or periods it has many variables.

This considerable number of variables presents serious difficulties to solve it with the usual software tools, from the approach of the model to transfer it to the software to the restrictions of variables that can be solved with any specific software.

For a complete integration of the model in the daily operations of the institution it is necessary to find an efficient way to solve it, since courses are usually scheduled for 8 periods within which 50 advisors are used for the same number of courses. With these numbers we would have 20,000 variables to schedule a full semester.

One of the solutions that can be used is to divide this programming into smaller periods: monthly, bi-monthly, quarterly, etc. This would imply that the number of courses can also be reduced and therefore the variables used would be considerably reduced.

To this end, we consider that the main solution strategies could be the following: dividing the problem into smaller periods and/or creating a customized solution. These two options are not mutually exclusive; on the contrary, solving smaller problems with an explicitly created solution would reduce the complexity of the model and the solution would be found more quickly.

One of the advantages of dividing the problem is not only to reduce the complexity of the problem, but also that we could have the results of the terminal efficiencies for shorter periods, that is, we would not have to wait a full semester to review the performance of the teachers but would have to perform the measurements monthly or bimonthly, so that based on these new data we could make a new optimal assignment.

Regarding a customized solution, this could be a complete system that obtains the data directly from the Data Warehouse to perform the efficiency calculations and prepare them for use in the model.

With the data prepared, the system itself would just wait for the courses and consultants required for the next period and proceed to find the solution for it.

And at this point, to find the solution, it would be necessary to program it using exact methods, heuristics, evolutionary algorithms, or simulation. we even consider that the ideal scenario would be a mixture of all three.

The mixture is possible because the data for terminal efficiencies and opening rates are stochastic, which gives us the opportunity to simulate them when we have enough observations to associate them with a distribution function. Then to use these simulated data in the proposed allocation model, for which we can start solving it using heuristic methods (taboo search, genetic, or ant colony) and after a reasonable time pass the solution found by the heuristic method to one of the exact algorithms to speed up the search for the corresponding optimum.

Finally, as we have been able to observe, the work proposes a method to maximize the terminal efficiency of the courses, which implies that if this efficiency is increased, student desertion is also reduced. Although the model was developed according to the needs of the institution, we consider that the idea can be extended to any online course, because in most courses students must submit activities to evaluate the knowledge acquired during the topics.

References

1. Zarate Perez, A. F., & Flores de la Mota, I. (2021). Un modelo para el seguimiento de cursos de capacitación. Undécima Conferencia de Directores de Tecnología de Información y Comunicación en Instituciones de Educación Superior, TICAL2021 y 5° Encuentro Latinoamericano de e-Ciencia “Repensando La Universidad Impulsada por las Tecnologías Digitales” (págs. 208-221). RedClara.

2. Syakinah Faudzi, Syariza Abdul-Rahman, Rosshairy Abd Rahman, "An Assignment Problem and Its Application in Education Domain: A Review and Potential Path", *Advances in Operations Research*, vol. 2018, Article ID 8958393, 19 pages, 2018. https://doi.org/10.1155/2018/8958393

3. Gaspars-Wieloch, Helena. 2021. The Assignment Problem in Human Resource Project Management under Uncertainty. Risks 9: 25. https://doi.org/10.3390/risks9010025

4. Valery Gordon, Jean-Marie Proth, Chengbin Chu, A survey of the state-of-the-art of common due date assignment and scheduling research, European Journal of Operational Research, Volume 139, Issue 1, 2002, Pages 1-25, ISSN 0377-2217, https://doi.org/10.1016/S0377-2217(01)00181-3.

5. P.I. Tillett, An operations research approach to the assignment of teachers to courses, Socio-Economic Planning Sciences, Volume 9, Issues 3–4, 1975, Pages 101-104, ISSN 0038-0121, https://doi.org/10.1016/0038-0121(75)90018-X.

6. John J. Dinkel, John Mote, M. A. Venkataramanan, (1989) OR Practice—An Efficient Decision Support System for Academic Course Scheduling. Operations Research 37(6):853-864. https://doi.org/10.1287/opre.37.6.853

7. John M. Mulvey, A classroom/time assignment model, European Journal of Operational Research, Volume 9, Issue 1, 1982, Pages 64-70, ISSN 0377-2217, https://doi.org/10.1016/0377-2217(82)90012-1.

8. Janice K. Winch, Jack Yurkiewicz (2014) Case Article—Class Scheduling with Linear Programming. INFORMS Transactions on Education 15(1):143-147. https://doi.org/10.1287/ited.2014.0128ca

9. M. König, C. Koch, I. Habenicht and S. Spieckermann, "Intelligent BIM-based construction scheduling using discrete event simulation," *Proceedings of the 2012 Winter Simulation Conference (WSC)*, 2012, pp. 1-12, doi: 10.1109/WSC.2012.6465232.