Visualization Methods for Educational Timetabling Problems: A Systematic Review of Literature

Wanderley de Souza Alencar, Hugo Alexandre Dantas do Nascimento, Fabrizzio Alphonsus A. de M. Nunes Soares and Humberto J. Longo

Instituto de Informática, Universidade Federal de Goiás, Alameda Palmeiras, Campus Samambaia, Goiânia-GO, Brazil

1 INTRODUCTION

In educational institutions (schools, colleges, universities, etc.), regardless of the level at which they act, one of the most recurrent and important problems is the preparation of class-teacher schedules. The scientific literature dedicated to this problem generally calls it Educational Timetabling Problem (Ed-TTP), being a subcategory of the so-called Scheduling Optimization Problems, as exemplified by Fernandes et al. (2016) and Saviniec et al. (2018).

Solving an instance of an Ed-TTP is, typically, to find a way to associate a set of events to a set of available timeslots, and, in some cases, also to define the locations where these events will occur. An event is often called a class or a didactic session, and brings together a group of teachers and students who will carry it out using a room (a typical classroom, a laboratory, a theater, a studio, etc.). In real scenarios, another common fact is the need for an eligible solution that satisfies hard and soft constraints relating expected events and all the necessary resources, such as teachers or lectures, auxiliary equipment (blackboard, data projector, computer, etc.), specific rooms (laboratory, theater, a studio, etc.) and technical support teams. The hard constraints are of compulsory attendance, while soft constraint are optional, but convenient in favor of improving the quality of the obtained solution.

As a result of more than six decades of research, the scientific community devoted to timetabling demonstrates that most of its variations belong to the class of \(NP\)-Hard problems, as shown by Werra et al. (2002) and Elloumi et al. (2014), among others. There are currently a large number of computational methods available for its satisfactory resolution, dealing with different problem specifications and sizes of instances. Many of these methods are synthesized in (Schaerf, 1999; Lewis, 2008; Pillay, 2014; Babaei et al., 2015; Oude Vrielink et al., 2017).

In general, the identified solution for an instance of a timetabling problem is visually presented in the form of a 2D table or a set of them, showing the days of the week and the times of the scheduled activities and even the place where each activity will occur. User interaction with the table can be supported in order to provide extra textual information or to allow manipulation of the time scheduling. Such a configuration, very common in many timetabling systems, is referred here as a traditional 2D-table format/system.

Understanding all types of information using a table representation may be a challenging task. The Information Visualization (IV) area highlights that some other visual techniques can emphasize properties of an abstract object, thus allowing a better understanding of it and raising cognition, as argue Gershon...
and Eick (1997) and Card et al. (1999).

In this context, the current paper presents the “state of art” of IV methods applied to Ed-TTPs by an analysis of the scientific publications registered in the last twenty years (1998–2018) in seven solid scientific databases. In addition, we present some possibilities for future work that researchers, dedicating themselves to the intersection of the areas of timetabling and information visualization, may pursue.

The remainder of this work is organized as follows: Section 2 describes the methodology adopted for the literature review; Section 3 explains the steps conducted in the review and their quantitative results; Section 4 presents the contributions of the main papers found in the SRL, with focus to the proposed IV techniques and their forms of interaction; and Section 5 draws our conclusions about the relationship between IV and Ed-TTP.

2 RESEARCH METHODOLOGY

For the study, we performed a Systematic Review of the Literature (SRL), since it is a process that provides well-defined, solidly backed bases for the identification, evaluation, synthesis and interpretation of the relevant evidences to a particular research question. Following the SRL guideline defined in (Kitchenham and Charters, 2007), three main stages were taken: (I) defining a review protocol; (II) conducting the review of the studies; and (III) reporting the review.

The first stage involved, in turn, the definition of: (L1) research questions; (L2) keywords, synonyms and search strategy; (L3) inclusion and exclusion criteria for the selection of studies; and (L4) a checklist for study quality assessment.

Each of these elements is detailed next. To carry out the SRL in a collaborative way between the authors of current paper, two software platforms were used: Parsifal (https://parsif.al/) and SESRA (http://sesra.net/).

2.1 Research Questions

The main objective of the SRL was to answer the following question: “What is the state of the art in the application of IV methods to the presentation or/and the manipulation of solutions for Ed-TTP?”.

Some more specific questions that unfolded the previous one were formulated: (RQ.1) What are the IV methods currently available for presenting and/or manipulating solutions for Ed-TTP? (RQ.2) Do the visualizations help to identify different aspects of a timetabling solution such as: unassigned classes and teachers/lectures, the distribution of the classes over the week, time scheduling conflicts, the matching between the teachers’ preferences and what was effectively scheduled for them, etc.? (RQ.3) Has the effectiveness of the new visualizations been compared to the traditional 2D-table format? (RQ.4) If the answer for Question RQ.3 is “yes”, then what measurements were adopted for evaluating effectiveness? (RQ.5) Do the visualizations show a single solution or the “solution space” for the timetabling problem being solved? (RQ.6) Do the visualizations support human interaction? (RQ.7) If the answer for Question RQ.6 is “yes”, then what types of interaction (for example: applying data filters, drag-and-dropping elements of the visualization, selecting areas of interest for focused optimization, etc.) are supported? (RQ.8) If the answer for Question RQ.6 is “yes”, which are the aims of the interactive features designed for the visualizations (helping understand the timetabling; helping define the timetabling problem in a more precise way; or helping find a better solution for the timetabling problem)?

It is important to emphasize that the above questions RQ.6–RQ.8 are directly related to the use of interactive techniques, mediated by a visual user interface, as an approach to solve scheduling problems, in this case, a timetabling problem.

2.2 Keywords, Synonyms and Search Strategy

The following keywords, usually with a high incidence in the scientific literature dedicated to the topic, and their associated synonyms, were chosen for driving the SRL: (K.1) timetabling (timetable, time table, time-table); and (K.2) visualization (GUI, interface, presentation, view, visual, visualisation); (K.3) educational (classroom, college, education, school, university).

The general logic search string was defined as [[timetabling]∧[visualization∧educational]], with each pair of brackets a disjunction of synonyms.

This string was adjusted for every electronic database in order to meet its search syntax. We searched on seven databases: (1) ACM Digital Library; (2) DOAJ – Directory of Open Access Journals; (3) IEEE Digital Library; (4) Science Direct; (5) Scielo; (6) Scopus; and (7) Springer Link, since they gather most of the publications on the area of timetabling and IV.

In order to concentrate the search on the most relevant papers (and not potentially in thousands of unrelated publications), the search string was applied only
on the fields paper titles, keywords, and abstracts of the databases, usually referenced as TAK, except for the DOAJ and the Springer Link, in which the search was made throughout the full text of the papers. For the Scopus database, only the K.1 part of the search string was applied on TAK, while the remainder of it (K.2 and K.3) was verified on the full text. This broadened the search, but still kept the amount of results manageable. Another special case was the ACM Digital Library, where the search was done in the “The ACM Guide to Computing Literature” collection, since this contains the complete set of articles in the area of Computing and related fields.

In addition, we restricted the search in all databases to the period from 1998 to 2018 (until Oct/2018). Through an extensive search before the realization of the SRL, in the aforementioned databases and using the same search string, publications prior to the year 1998 were identified but disregarded, since they notably use traditional 2D-table to present solutions for the Ed-TTP. Furthermore, when some kind of interactivity was reported, it was performed by simple drag-and-drop or copy-and-paste operations and select/deselect of elements, such as cells and regions. Consequently, only works published in the last twenty years were considered relevant.

### 2.3 Inclusion/Exclusion Criteria

Inclusion and exclusion criteria were defined for guiding the selection of primary studies relevant to the SRL. A study was included if it was: (1) available in full-text; (2) published in a journal or in the annals of a conference; (3) a technical report, including surveys; or (4) a master dissertation or a doctoral thesis. On the other hand, it was excluded if it: (1) was not written in English; (2) did not present, or discuss, techniques of the IV area; or (3) did not propose and/or discuss an IV technique for the Ed-TTP; (4) focused only on traditional 2D-table visualizations with ordinary user interactions.

A set of Yes/No quality assessment questions was also defined based on a refinement of the main research questions (RQs). For instance, they asked if the visualization method in use could be changed in the graphical interface or if a particular type of interaction for some specific goal was supported.

### 3 CONDUCTING THE REVIEW

The proper search strings, with the corrected search conditions, were applied to the databases. Next, as a Study Selection step, the bibliographic references were evaluated according to the inclusion and the exclusion criteria (with the help of the SRL software platforms). When necessary, the full text of the studies was consulted.

Table 1 summarizes the number of papers recovered when searching in each database, and the amount of accepted (selected) and reject studies during the Selection step. In total, when removing duplicates, only 10 studies were accepted, with 8 of them being academic productions of a same research group.

<table>
<thead>
<tr>
<th>Database</th>
<th>Total</th>
<th>Accepted</th>
<th>Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM DL</td>
<td>346</td>
<td>4</td>
<td>342</td>
</tr>
<tr>
<td>DOAJ</td>
<td>11</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>IEEE DL</td>
<td>54</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>Scielo</td>
<td>1,734</td>
<td>0</td>
<td>1,734</td>
</tr>
<tr>
<td>ScienceDir</td>
<td>64</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>Scopus</td>
<td>1,362</td>
<td>7</td>
<td>1,355</td>
</tr>
<tr>
<td>Springer</td>
<td>2,314</td>
<td>2</td>
<td>2,312</td>
</tr>
</tbody>
</table>

Some limitations faced during the search process and the strategies adopted in order to overcome them are worth mentioning. The Scielo database did not allow a search with the complete search string. Therefore, three search activities had to be performed, each using a distinct pair of the components of the search string: timetabling (K.1), visualization (K.2) and educational (K.3). The employed strings were: (K.1 ∧ K.3), that returned six studies; (K.2 ∧ K.3), that resulted in 1,734 studies; and (K.1 ∧ K.2), that returned only one study. The union of these results had exactly 1,734 papers. After reading and analyzing titles and abstracts, only twelve papers had to be read completely and none of them were accepted.

Another particular, and more complex case was that of Springer Link database. The search in it can not be limited to Title-Abstract-Keywords (TAK) and had to be performed throughout the full text. As a consequence, 16,542 studies were returned: 2,545 published in conferences and 13,997 in journals. The results were then filtered to show only the ones in the fields of “Computer Science”, “Engineering” and “Mathematics”, what reduced the search to 2,314 papers.

Given the limited number of accepted papers, the quality criteria were not evaluated and, therefore, the Quality Assessment step was not performed.

Finally, all accepted papers were fully read and all relevant data for answering the search questions was extracted. Table 2 synthesizes the major aspects of identified studies in this SRL. The table columns contain the following information:
1. Bibliographic reference of the study;
2. IV method(s) that was applied;
3. Data elements displayed by the IV method;
4. Types of Interaction supported by the IV, considering five possibilities — [Graph] allow to change the objective function of the underlying optimization timetabling problem; [Constraints] allow to change constraints of the optimization problem; [Technique] allow to choose and run one of many optimization methods; [Manual Solution] a timetabling solution can be manually changed using the visualization; and [Selected Area] allow to select a region for deleting, including or changing data in group;
5. Available application that implements the work; and
6. Optimization technique(s) being used for the resolution of the underlying problem.

It is important to note that one of the difficulties encountered during the SRL is that some of the studies do not clearly present information that allows us to answer certain research questions in a safe way. In this scenario, some level of interpretation was necessary, but always with the largest effort to provide the right answer. In the worst case, the evaluation could not be performed in that context or the response was inconclusive, which is signaled by a dash (−) in Table 2.

Next, we describe the main concepts in some of the studied selected by the SRL.

4 VISUALIZATION METHODS FOR Ed-TTP

In one of the first studies published in the investigated period (Piechowiak and Kolski, 2004), the authors propose an interactive decision support system for the analysis of educational timetabling problems and the elaboration of solutions for them. The system has a visualization module that shows an interface customized to the user profile (designers, analyzers and consultors) and that employs both 2D-tables and a time chart (resources × time) for user interactions. A timetabling solution is built manually, but with a constraint-based reasoning engine to assist the user in obtaining a solution to the target problem, including hard and soft constraint violation detection in a semi-automatic way.

Another paper, by Thomas et al. (2008) focuses on solving Examination Timetabling Problems (Exam-TTP), a type of Ed-TTP. The authors propose a visual framework that operates on three interrelated phases: pre-processing (visualizing raw data inputted from the user), processing (solving the optimization problem and visualizing the generated timetabling solution) and post-processing (improving the timetabling solution) of a timetabling solution. A visual model is used by the authors as an instrument to clarify the problem complexity — a NP-hard problem — and to provide an integrated visualization of the phases that can contribute to its understanding and satisfactory resolution. The article details the use of IV techniques for the pre-processing phase, in which, for example, directed graph drawings indicate the relationship between enrolled candidates (students) with the courses for a particular semester. In another example, courses are represented by nodes and constraints are modelled by edges, as shown in Figure 1a. The node color indicates the period of the day when a course is taught. Nodes with the same color are being offered concurrently in the same academic period.

In an extension of the previous work, and by means of a pair of strongly related studies (Thomas et al., 2009b,a), more visualization techniques are proposed for the pre-processing phase. The new visualizations include: a tree view (aggregating many types of data), a directed graph drawing, a histogram combined with a circular graph layout (courses × enrolled students); and some standard histograms (rooms × subjects and rooms × timeslots). It is possible for the user to interact with the visualizations but for the aim of better understanding the data.

Still researching on Exam-TTP, further studies of the same group of authors (Thomas et al., 2010b,c) propose a tool, called VizSolution, for the processing phase of the visual framework they conceived. An interactive visualization approach is adopted, in which a user and a machine (a scheduler implemented as a constraint satisfaction program) operate in a symbiotic way for solving a timetabling problem, including the allocation of classrooms. The tool allows to define the problem by means of an element called Filter, which employs graph drawing to represent constraints, as shown in Figure 1b. Graph drawings (in 2D and projected 3D) are also used to indicate conflicts and/or preferences.

With the goal of helping the user to more easily identify and to solve time conflicts in teachers and courses’ schedules, some researchers (Fui et al., 2010) have proposes a system, named CORECTS, that models a timetabling solution as a graph. A modified version of a standard graph drawing algorithm is employed for visually presenting the solution and to highlight conflicts. Via “simple strokes gestures” made by user on the visualization (using a touch screen monitor), it is possible to do operations that
affect the conflicts.

Back to the Exam-TTP (Thomas et al., 2011), parallel coordinates were used for answering the question: “How hard is this problem to solve?”. The authors conceived the ParExaViz tool, which facilitates the exploration of raw data of an Exam-TTP instance and to highlight conflicts.

Addressing the problem of university timetabling, some other authors (Abdelraouf et al., 2011) introduce a visual graphic communication tool that lets the users to specify their problem in an abstract manner, involving human resources (people), events (lectures) and meta-events (courses). These elements are represented by nodes in a graph, while edges indicate their relationships. The edges have different interpretations according to the type of the elements involved. A visual interface integrates textual and graphical components (many tabs). The timetabling problem is solved by a module that implements a constraint satisfaction algorithm.

Finally, a system called ExamViz (Thomas et al., 2012) is conceived with an integrated problem solving environment (PSE) to the Exam-TTP. It works as a computational steering mechanism with automated steering interactions and/or with a user-driven process. Through the user interface, it is possible to perform conflict analysis in the timetable and to apply a reconciliation process based on evolutionary algorithms. The analysis can be done visually using parallel coordinates as well as 2D-tables and graph drawings.

5 CONCLUSIONS

Based on a SRL, we found that, in the last twenty years, only few studies have investigated the use of advanced information visualization methods for helping understand and/or solve educational timetabling problems. This contrasts to the majority of scientific papers in the timetabling area, which main concern is to provide a method capable of solving the problem close to optimality, with little or no attention to the visualization aspect. Some examples of such works are Nouri and Driss (2016), Babaei et al. (2018) and Lindahl et al. (2018). Nevertheless, in the latter article (Lindahl et al., 2018), the authors highlight that the work could be complemented by having a graphics user interface. Another recent work (Oude Vrielink et al., 2017) emphasizes that, while the academia tends to develop intelligent and profound methods to solve timetabling problems, the industry appears to develop and to design an easy to use interactive tool that aims at meeting the needs of the educational institutions.

Some ideas for future work were identified during this SRL. They include:

1. Developing new visualization approaches for helping to identify and understand the main factors responsible for infeasibility in large-scale timetabling problem instances;
2. Proposing new visualizations for supporting interactive optimization of educational timetabling problems and investigating their effectiveness and usability;
3. Extending the SRL for including the study of commercial and non-commercial software systems for Ed-TTP, focusing on the visualizations and on the interaction features; and
4. Adjusting visualization methods proposed to timetabling in other areas, such as for organizing hospitals’ agendas and transportation (bus, train and airplane) schedules, to the Ed-TTP scenario.
Table 2: Analysis of the Accepted Publications. Some features are marked as [Y]es, [N]o or [–] for inconclusive.

<table>
<thead>
<tr>
<th>Study Identification and Characterization</th>
<th>Interaction types by IV Technique(s)</th>
<th>Displayed Data Elements</th>
<th>Optimiz. Manual Select</th>
<th>Application and Solution Technique</th>
<th>Application</th>
<th>Manual or by any automatic scheduling</th>
<th>Solve the Ed-TTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thomas et al. (2008)</td>
<td>Oriented cluster graph drawing, classes and students enrolled</td>
<td>–</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Manual or by any automatic scheduler</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2009b)</td>
<td>Directed graph drawing, histogram, daisy chart, tree view</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>There is no attempt to solve the problem, just processing/visualizing raw input data</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2009a)</td>
<td>2D-table, oriented cluster graph drawing, histogram and tree representation</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Constraint Satisfaction Program (in a constraints network, with backtrack)</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2010b)</td>
<td>2D-table, graph drawing (2D, 3D), timetable (complete), constraints and conflicts</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Constraint Satisfaction Program (in a constraints network, with user collaboration)</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2010c)</td>
<td>2D-table, graph drawing (2D, 3D), timetable (complete), constraints and conflicts</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Constraint Satisfaction Program (in a constraints network, with user collaboration)</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2011)</td>
<td>Parallel coordinates (for multi-dimensional variables)</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>There is no resolution of the problem, just processing raw data</td>
<td></td>
</tr>
<tr>
<td>Thomas et al. (2012)</td>
<td>2D-table, graph drawing, tree representation</td>
<td>N</td>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>Manual and user-driven problem solving with constraints and conflicts</td>
<td></td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

This work was financially supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG).

REFERENCES


