

PROBLEMA DE ALOCAÇÃO DE HORÁRIOS E SUAS VARIAÇÕES: UMA REVISÃO SISTEMÁTICA

Mathews Monoo Pereira de Bellis

Instituto Federal de São Paulo - IFSP
Rua Monsenhor José Vita, 280, Campos do Jordão - SP
m.monoo@aluno.ifsp.edu.br

Marques Moreira de Sousa

Instituto Federal de São Paulo - IFSP
Rua Monsenhor José Vita, 280, Campos do Jordão - SP
marques.sousa@ifsp.edu.br

Marlon da Silva

Instituto Federal de São Paulo - IFSP
Rua Monsenhor José Vita, 280, Campos do Jordão - SP
marlon.silva@ifsp.edu.br

RESUMO

O Problema de Geração de Horários Escolares é um recorrente desafio para instituições de ensino brasileiras, uma vez que a renovação de dados e restrições é periódica, seja semestral ou anual. Este trabalho apresenta uma Revisão Sistemática de Literatura realizada em artigos completos publicados em português e que estejam relacionados à instituições de ensino brasileiras, a fim de identificar as principais metodologias utilizadas para a solução deste problema. Os artigos foram obtidos a partir do motor de busca Google Acadêmico, um dos mais utilizados atualmente. A partir dos critérios utilizados, foram selecionados 31 artigos em um total de 1280 resultados da busca, além de serem classificados por escores qualitativos. Questões de pesquisa acerca das características estudadas nos trabalhos são definidas e discutidas.

PALAVRAS CHAVE. Otimização Combinatória, Alocação de Horários, Revisão sistemática.

Tópicos: AdP&ED - PO na Administração Pública e Educação

ABSTRACT

The School Timetabling Problem is a recurring challenge for Brazilian teaching institutes, once that their data and restrictions renewal is periodic, either semi-annually or annually. This work presents a Systematic Review of the Literature executed on complete papers published in Portuguese and that also are related to Brazilian Education Institutes, to identify the most used methodologies to solve this problem. The papers were obtained from Google Scholar search engine, one of the most used nowadays. From the criteria utilized, 31 articles were selected out of 1280 search results, in addition to being classified by qualitative scores. Research questions about the characteristics studied individually in the works are defined and discussed.

KEYWORDS. Combinatorial Optimization. Timetabling problem. Systematic review.

Paper topics: AdP&ED - OR in government and Education

1. Introdução

Nas instituições de ensino, a criação e a manutenção de grades horárias é uma atividade necessária e importante. Porém, é uma atividade complexa, pois é necessário alocar indivíduos em locais com capacidade restrita e horários conflitantes. Tal tarefa é facilitada através da modelagem do problema de horário escolar.

O Problema de Geração de Horários Escolares (PGHE) (do inglês *timetabling problem*) consiste em obter uma grade horária para as aulas em uma determinada instituição de ensino, dado um conjunto limitado de horários de modo a satisfazer várias restrições. Uma das maiores dificuldades em problemas de alocação de horários escolares está em sua complexidade. Tais problemas são considerados NP-completos [Tassopoulos e Beligiannis, 2012] e exigem a utilização de alternativas para encontrar soluções próximas do ótimo que sejam satisfatórias.

Na literatura especializada, o problema de horário escolar é classificado em três categorias, de acordo com o cenário de aplicação [Arenales et al., 2015]: Ensino Médio (*school timetabling problem*), em que diferentes turmas compartilham o mesmo currículo e devem ter a grade horária completamente preenchida; Universidades (*university course timetabling problem*), em que alunos possuem liberdade na seleção de disciplinas; Exames no final do período letivo (*examination timetabling problem*), em que inexiste conflitos entre horários por parte dos alunos.

Em todas as categorias de problemas de horário escolar, existem dois tipos de restrições que são classificadas de acordo com a relevância. Neste caso, podem existir restrições altamente prioritárias (*hard constraints*) e restrições que não sejam essenciais (*soft constraints*) mas que acomodam de uma melhor forma determinadas grades [Babaei et al., 2015].

Dentre as restrições prioritárias (*hard constraints*), existem restrições em que um professor não pode lecionar em duas salas simultaneamente; uma disciplina não pode ser disponível para duas salas simultaneamente; em cada espaço de tempo, deve ser atribuído exatamente um professor, uma turma e uma sala de aula; algumas disciplinas necessitam ser lecionadas em determinados horários; a capacidade das salas de aula devem ser respeitadas, entre outras. O não cumprimento de restrições prioritárias torna a solução do modelo inviável.

Por outro lado, existem restrições que podem ser consideradas na obtenção de uma solução que não torna o modelo inviável, conhecidas como *soft constraints*. Dentre essas restrições, estão inclusos exemplos como o professor poder escolher horários preferenciais; o professor pode solicitar uma determinada sala de aula ou laboratório; evitar ou reduzir a quantidade de aulas vagas por parte de alunos e professores; evitar que uma turma tenha aula em determinados períodos ao longo da semana, assim como outras que são específicas a cada instituição.

Tanto restrições prioritárias ou não seguem alguns critérios para serem formuladas. Segundo Pillay [2014], a confecção destas restrições são orientadas de acordo com o objetivo do modelo. Ao criar estas restrições, podem ser considerados:

- *Requisitos do problema*: específicos das regras da escola ou do sistema educacional adotado;
- *Prevenção de conflitos*: evita que recursos (professores, salas de aula, laboratórios) sejam utilizados ao mesmo tempo;
- *Utilização de recursos*: previne que recursos sejam pouco ou muito utilizados ao longo da grade horária;
- *Carga de trabalho*: limita o quanto um lugar ou um professor poderá ser aproveitado;
- *Distribuição do período*: especifica a disposição de recursos para o período de aulas, de modo a evitar aulas vagas ou um desequilíbrio na distribuição das aulas ao longo da semana;

- *Preferências*: descreve as preferências de trabalho de algum recurso;
- *Aulas*: determina quais recursos devem ser utilizados para determinadas aulas, tais como laboratórios, ginásios ou salas de videoconferência.

Esta lista de fatores é altamente orientada na construção de grades horárias para cursos superiores. No entanto, modelos de otimização de horário escolar para universidades são considerados NP-difíceis [Babaei et al., 2015] e devem sempre ser aplicados antes do início de um período letivo – em geral, semestralmente.

De acordo com os dados do último Censo da Educação Superior¹, o Brasil conta com 2537 instituições de ensino superior (IES), totalizando 34785 cursos presenciais, 384474 docentes e 6394244 matriculados presencialmente. Nota-se, com estes dados, que uma grande variedade de particularidades são levadas em conta na construção de modelos para problemas de horário escolar, uma vez que nem todas as instituições contam com docentes exclusivos, uns cursos são noturnos, enquanto outros são em período integral, entre outras características que dificultam a padronização de um modelo genérico para solucionar o problema de alocação da grade horária nas IES.

Diante da diversidade de formas que o problema de alocação de horário escolar pode ser modelado e da variedade de métodos empregados para resolvê-los, surge a necessidade de explorar horizontalmente o tema, de modo a conhecer mais profundamente como ele está sendo tratado no Brasil e quais estratégias são mais utilizadas. Tal estudo permitirá um melhor embasamento na proposição de novos modelos que possibilitem o aperfeiçoamento das soluções, uma vez que se trata de um problema NP-difícil.

O objetivo deste trabalho é apresentar uma revisão sistemática da literatura (RSL) acerca do problema de geração de horário escolar e as principais abordagens para a resolução de tais problemas em instituições de ensino brasileiras, subsidiando assim uma escolha pertinente de estratégia para o tratamento do problema no contexto do Instituto Federal de São Paulo, Campus Campos do Jordão.

Este trabalho é estruturado da seguinte forma: a Seção 2 apresenta a metodologia da RSL utilizada. Os resultados obtidos e a respectiva discussão são apresentados na Seção 3. Conclusões e sugestões para trabalhos futuros são apresentados na Seção 4.

2. Metodologia

Por se tratar de um estudo de cunho exploratório, tecnicamente baseado em trabalhos científicos já publicados na literatura, uma RSL caracteriza-se por ser um estudo secundário que visa sintetizar informações importantes que foram devidamente publicadas e validadas por meio de trabalhos primários [Sousa et al., 2019]. Complementarmente, uma RSL diferencia-se de outros tipos de trabalhos pela característica determinística, em que outros pesquisadores podem reproduzir a metodologia utilizada por meio de um método de execução bem definido, conferindo confiabilidade às informações relatadas.

Segundo Kitchenham [2012], para definir os métodos necessários para se atingir os objetivos de uma RSL, são necessárias três etapas básicas: (i) planejamento, (ii) condução e (iii) relatório. Na primeira etapa é definido o protocolo a ser utilizado, responsável por garantir a futura reprodutibilidade do trabalho. Na segunda etapa são consultados os critérios definidos no protocolo visando a seleção criteriosa dos trabalhos a serem analisados, os quais possuem potencial de resposta às perguntas específicas da pesquisa. Nesta fase, ocorrem a busca e a filtragem de trabalhos, necessárias para que, na fase (iii), os dados relevantes sejam analisados e condensados em formato de relatório.

¹<http://portal.inep.gov.br/web/guest/censo-da-educacao-superior>

2.1. Planejamento

2.1.1. Questões de pesquisa

Assim como abordado em outros trabalhos que conduziram pesquisas de caráter exploratório [Alves et al., 2018; Sousa et al., 2019], faz-se necessário definir questões gerais acerca da condução dos trabalhos analisados. As questões de pesquisa definidas para aplicação são:

Q-1 Quais técnicas são utilizadas para resolver o problema?

Q-2 Quais as características dos problemas resolvidos?

Q-3 Quais as características dos experimentos realizados?

Q-3.1 Qual a quantidade de casos de testes?

Q-3.2 Foram utilizadas instâncias definidas em problemas anteriores ou foram criadas novas?

Q-3.3 Os testes são realizados com dados reais ou artificiais?

Q-3.4 Com quais outras técnicas previamente aplicadas foi comparada a abordagem proposta?

A Questão **Q-1** visa identificar quais as técnicas são efetivamente aplicadas para solução de problemas relacionados à definição do quadro de horários em instituições de ensino e, a partir de sua resposta, espera ser possível identificar o porquê de tais escolhas. O objetivo de **Q-2** é classificar os trabalhos de acordo com características pertinentes a cada variante do problema. Por fim, **Q-3** e suas subquestões buscam investigar as características dos testes que foram conduzidos em cada trabalho.

2.1.2. Seleção de fontes e *string* de busca

A escolha de fontes diversas propicia maior abrangência da busca, possibilitando o levantamento de trabalhos que estejam publicados em repositórios específicos. Entretanto, não necessariamente o acréscimo de novos motores de busca significa maior abrangência, visto que há a disponibilidade de motores que apresentam completude satisfatória de acordo com os limites estabelecidos para pesquisas de caráter exploratório a nível nacional, como é o caso do presente trabalho.

Após testes com os motores de busca Google Acadêmico², Periódicos CAPES³ e SciELO⁴, optou-se por utilizar unicamente o primeiro motor de busca citado. Tal motor foi escolhido por agregar uma série de repositórios e ter retornado resultados mais relevantes do que os outros, além de ser gratuito. Três *strings* de busca foram formatadas e concatenadas para a realização da pesquisa (ver Tabela 1). As *strings* foram dimensionadas para identificar trabalhos que trazem no título, palavras-chave ou corpo, referências ao problema de alocação de horários em instituições de ensino brasileiras. Devido à restrição de escopo do presente trabalho, foram consideradas palavras em português que possuam correlação com o problema abordado. Variações de nomenclatura para o nome do problema podem ser comumente identificadas. Assim, o uso de *strings* simplificadas de forma individualizada nem sempre garante a correta exploração do espaço de busca. Desta forma, foi utilizada a *string final* que compõe de forma integral as *substrings* previamente mencionadas (*strings* 1 a 3).

²<https://scholar.google.com.br>

³<https://www.periodicos.capes.gov.br/>

⁴<https://scielo.org/>

Tabela 1: *Strings* de busca utilizadas no motor de busca

<i>String 1</i>	Com, no mínimo, uma das palavras: “QUADRO DE HORÁRIOS” “PROGRAMAÇÃO DE HORÁRIOS”
<i>String 2</i>	Com, no mínimo, uma das palavras: “ELABORAÇÃO DE GRADE HORÁRIA” “OTIMIZAÇÃO DO QUADRO DE HORÁRIOS”
<i>String 3</i>	Com, no mínimo, uma das palavras: “ALOCACÃO DE SALAS DE AULA” “GERAÇÃO AUTOMÁTICA DE HORÁRIOS”
<i>String final</i>	Com, no mínimo, uma das palavras: “QUADRO DE HORÁRIOS” “PROGRAMAÇÃO DE HORÁRIOS” “ELABORAÇÃO DE GRADE HORÁRIA” “OTIMIZAÇÃO DO QUADRO DE HORÁRIOS” “ALOCACÃO DE SALAS DE AULA” “GERAÇÃO AUTOMÁTICA DE HORÁRIOS”

2.1.3. Critérios de inclusão e exclusão

Os critérios de inclusão (I) e exclusão (E) utilizados nesta RSL têm como objetivo filtrar estudos primários que estejam alinhados com o tema tratado e apresentem nível mínimo de qualidade esperado para uso em estudos secundários:

- I-1** Os artigos devem estar disponíveis de forma integral e gratuita online, ou os autores devem ter acesso aos artigos através da sua instituição via Periódicos Capes;
- E-1** Artigos que não abordem como tema principal a definição de quadro de horários em instituições de ensino brasileiras;
- E-2** Artigos que não foram submetidos a uma revisão por pares;
- E-3** Estudos incompletos ou que não apresentam detalhes suficientes sobre a abordagem utilizada;
- E-4** Estudos secundários ou terciários;
- E-5** Trabalhos que não possuam *abstract*;
- E-6** Artigos não publicados em *workshops*, conferências ou periódicos;
- E-7** Trabalhos não apresentados em formato de artigo científico completo (6 ou mais páginas);
- E-8** Artigos que tenham sido publicados antes de 2010;
- E-9** Artigos que não estão escritos em língua portuguesa.

2.1.4. Critérios de qualidade

Para a realização da análise qualitativa dos trabalhos selecionados, os seguintes critérios foram estabelecidos para medição e comparação, utilizando escores com base em critérios estabelecidos por Alves et al. [2018] e Sousa et al. [2019]:

- C-1** O método proposto foi descrito de maneira clara e abrangente, permitindo que seja possível a replicação dos resultados? Não (escore = 0), Parcialmente (escore = 0,5), Sim (escore = 1);
- C-2** O método proposto alcança soluções melhores para qual porcentagem das instâncias testadas? melhora < 30% (escore = 0,3), 30% ≤ melhora < 90% (escore = 0,6) e melhora ≥ 90% (escore = 1);

- C-3** Os testes computacionais foram aplicados a instâncias artificialmente criadas ou são baseadas em casos reais? Nenhum (escore = 0), A (escore = 0,5), R (escore = 1);
- C-4** Foram realizadas comparações com outras abordagens contidas na literatura? Não (escore = 0), Parcialmente (escore = 0,5), Sim (escore = 1);
- C-5** Foram realizados testes estatísticos para a calibração de parâmetros das heurísticas? Não (escore = 0), Sim ou em caso de solução puramente exata ou similar (escore = 1);
- C-6** Foram realizados testes estatísticos para validar os resultados encontrados pela solução proposta? Não (escore = 0), Parcialmente (escore = 0,5), Sim (escore = 1).

2.2. Condução

A pesquisa foi realizada em 22 de abril de 2020, retornando 1280 trabalhos pelo motor de busca Google Acadêmico. Inicialmente, foram identificados 338 documentos duplicados, os quais foram devidamente excluídos. Ao realizar uma leitura transversal do título e do resumo (quando necessário), foi possível identificar que um total de 854 documentos deveriam ser excluídos de acordo com o critério E-1, pois, conforme explicitado na Tabela 2, ainda que abordassem problemas de otimização, não eram direcionados à definição de quadro de horários em instituições de ensino brasileiras.

Após realizadas as duas primeiras filtragens, 88 documentos foram selecionados para avaliação detalhada segundo os critérios de exclusão definidos. Para um melhor gerenciamento dos documentos, foi utilizado o software StArt (*State of the Art through Systematic Review*)⁵.

Tabela 2: Critérios e Contagem de artigos excluídos.

Área de estudo relacionada	Quantidade	Porcentagem (%)
Agronegócio	12	1,41
Saúde	21	2,46
Matemática (exclusivamente)	37	4,33
Arquivos de Grades Horárias	51	5,97
Arquivos de Projetos Pedagógicos	54	6,32
Transporte	117	13,70
Outros / Sem relação	562	62,81
Total	854	100

O problema de alocação de horários é um clássico problema de otimização combinatória que possui grande aplicabilidade prática, uma vez que instituições de ensino públicas e privadas podem utilizá-lo para modelar seus quadros de horários com base em suas instalações físicas, serviços oferecidos e profissionais disponíveis. Na literatura, tal problema é abordado utilizando-se diferentes nomenclaturas. Entretanto, a ideia básica consiste na definição de horários a serem aplicados em instituições de propósito variado. Neste trabalho, estudos que utilizam o problema de alocação de horários em outras áreas que não sejam o ensino são descartados. Na Tabela 2 é possível encontrar as principais áreas de estudos não consideradas e que foram retornadas pelo motor de busca.

A partir disso, para realização da filtragem dos trabalhos, os critérios foram aplicados manualmente em cada um dos 88 trabalhos restantes, analisando-se o título, resumo, palavras-chave, formato e local de publicação. Após o filtro, 57 trabalhos foram excluídos por se encaixarem em

⁵http://lapes.dc.ufscar.br/tools/start_tool

um ou mais critérios de exclusão, sendo os critérios E-7, E-1 e E-5 os que apresentaram maiores porcentagens de aplicação, respectivamente. Ao final, foram selecionados 31 trabalhos para o desenvolvimento desta RSL. Devido a limitação de espaço, as referências para os trabalhos foram citadas em um apêndice publicamente disponível <https://bit.ly/327ZaIX>.

3. Resultados e Discussão

Nesta seção, são apresentados os resultados e discussões acerca das questões de pesquisa consideradas na Seção 2. Após aplicação do protocolo de busca e seleção, os estudos filtrados foram identificados com IDs únicos para facilitar sua referência no decorrer do texto. Além da identificação, para cada estudo foi destacado sua origem, utilizando (C) para estudos publicados em conferências e (R) para estudos em revistas. Ainda são apresentados os títulos de cada trabalho e o ano de publicação. Todas essas informações se encontram detalhadas na Tabela 3.

Tabela 3: Estudos selecionados de acordo com a *string* de busca e os critérios de inclusão e exclusão.

ID	Origem	Título	Ano
T-1	C	Heurísticas baseadas no algoritmo de coloração de grafos para o problema de alocação de salas em uma instituição de ensino superior	2010
T-2	C	Investigação de Resolução do Problema de Programação de Horários Escolares usando FPGAs e VHDL	2010
T-3	R	Aplicação da Metaheurística Busca Tabu ao Problema de Alocação de Aulas a Salas em um Instituição Universitária	2011
T-4	C	Formulação para o Problema de Alocação de Salas de Aula com Minimização de Deslocamentos	2011
T-5	R	Heurísticas Eficientes para o Problema de Geração de Grade Escolar Automatizada	2011
T-6	C	Resolução do Problema de Programação de Cursos Universitários Baseada em Currículos via uma Meta-Heurística Híbrida GRASP-ILS-Relaxado	2011
T-7	R	Sistema de Alocação de Horários de Cursos Universitários: um estudo de Caso no Departamento de Computação da Universidade Federal de Sergipe	2011
T-8	C	Uma Abordagem Híbrida de SAT e Busca Tabu para o Problema da Programação de Horários Escolares	2011
T-9	C	Programação Linear Binária na Geração da Grade Horária do Curso de Engenharia de Produção da UFPR	2012
T-10	C	Técnicas de Busca Local para o Problema da Programação de Horários Escolares	2012
T-11	R	Um Método Heurístico Construtivo para o Problema da Grade Horária Escolar	2012
T-12	R	Uma Heurística GRASP+VND para o Problema de Programação de Horário Escolar	2012
T-13	C	Geração de Grade de Horários para Disciplinas de uma Instituição de Nível Superior Utilizando Programação Linear Inteira Multiobjetivo	2013
T-14	C	Planejamento de Grade de Horário em uma Universidade Brasileira usando Algoritmos Genéticos	2013
T-15	C	Aplicação de Algoritmos Genéticos e Métodos Evolucionários na Resolução do Problema de Alocação de Turmas: caso UFPR	2014
T-16	C	Desenvolvimento de um Modelo Híbrido Baseado em Algoritmo Genético e Busca Tabu para Resolução do Problema de Quadro de Horários Escolar	2014
T-17	C	Problema de Alocação de Salas em Cursos Universitários: um estudo de caso	2014
T-18	C	Um Estudo de Caso sobre a Geração de Quadros de Horários no Departamento de Ciência da Computação da UFRGS	2014
T-19	R	Aplicação de Algoritmos Genéticos na Construção de Grades de Horários	2015
T-20	C	Multithreading Iterated Local Search aplicado ao Problema de Horários Escolares	2015
T-21	C	Quadro de Horários Escolar: uso de Algoritmo Memético para Acelerar a Convergência do Método Evolutivo	2015
T-22	C	Solução do Problema de Alocação de Salas Utilizando um Modelo Matemático Multi-Índice	2015
T-23	C	Alocação de Horários de Professores e Turmas em um Curso Universitário	2016
T-24	C	Simulated Annealing Aplicado ao Problema de Programação de Horários do CCA-UFES	2016
T-25	C	Uma Heurística para o Problema de Programação de Horários: um estudo de caso	2016
T-26	R	Uma Investigação do Problema de Elaboração de Grade Horária (timetabling problem)	2016
T-27	C	Algoritmo Híbrido Iterated Local Search e Simulated Annealing para o Problema de Tabela-Horário de Universidades	2017
T-28	R	Aplicação e Comparação das Meta-Heurísticas Arrefecimento Simulado e Algoritmos Genéticos para Alocação de Horários Universitários	2017
T-29	R	Inserção de Parâmetros Controladores da Aleatoriedade no Método GRASP Aplicado a um Problema de Programação de Horários em Escolas	2017
T-30	C	Um Modelo em Prog. Linear Inteira para Alocação de Disciplinas: um estudo de caso no curso de Ciência da Computação da Universidade de Fortaleza	2017
T-31	R	Modelo de Programação Matemática na Elaboração de Quadros de Horários para Cursos de Graduação do IFRR	2019

Dentre os estudos elencados, pouco mais de 2/3 são provenientes de publicações em conferências (congressos, simpósios e fóruns), enquanto o restante está disponível em revistas científicas. Outro fato relevante é que dentre os artigos publicados em congressos, grande parte foi apresentada em edições do Simpósio Brasileiro de Pesquisa Operacional (SBPO) que é o principal evento nacional de divulgação de pesquisas realizadas na área de otimização combinatória.

3.1. Técnicas utilizadas: Q-1

Considerando a abrangência da área de otimização combinatória e a dinamicidade de técnicas utilizadas para resolução de problemas nesse campo, a primeira questão de pesquisa que o presente trabalho busca responder está relacionada às técnicas aplicadas na criação dos quadros de horários.

Para tanto, a Figura 1 ilustra de forma generalizada a representatividade das técnicas utilizadas nos estudos selecionados. Analisando a figura, observa-se que a maioria dos estudos fazem uso de técnicas heurísticas para resolução do problema de alocação de quadros de horários, correspondendo a um percentual próximo a 71% (22) dos estudos. Em seguida, 22,6% (7) dos demais

estudos utilizam métodos exatos baseados em modelos de programação linear e, por fim, outros 6,5% (2) utilizam técnicas de coloração de grafos para encontrar soluções viáveis ao problema.

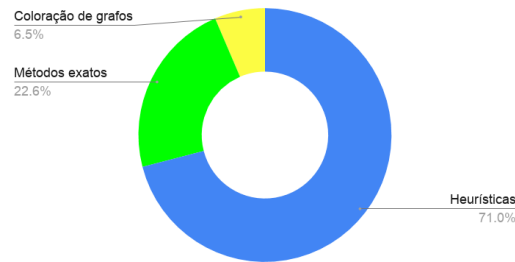


Figura 1: Quantificação de técnicas utilizadas nos trabalhos por tipo de abordagem.

Na área de otimização combinatória, existe uma diversidade de heurísticas que podem ser utilizadas para resolver os mais variados problemas e suas variantes. Para o PGHE, isso não é diferente, pois dentre os 22 estudos que utilizam explicitamente heurísticas, os autores fazem uso de ao menos seis heurísticas tradicionalmente conhecidas: Algoritmos Genéticos (AG) e suas variantes como Algoritmo Memético, *Iterated Local Search* (ILS), *Simulated Annealing* (SA), *Greedy Randomized Adaptive Search Procedure* (GRASP), *Tabu Search* (TS) e *Variable Neighborhood Descent* (VND). Dentre essas, as heurísticas AG, ILS e SA são as que mais aparecem nas soluções propostas, sendo referenciadas em 8, 6 e 6 estudos, respectivamente. É importante ressaltar que em alguns estudos, como T-5, T-6, T-10, T-12, T-16, T-27 e T-28, formas de hibridização de heurísticas são definidas para aumentar a capacidade de aprimoramento das soluções por meio da estratégia proposta.

3.2. Características dos problemas resolvidos: Q-2

A segunda questão de pesquisa tratada nesta RSL está relacionada com a característica do problema de alocação de horário que cada estudo busca resolver. Neste ponto, é comum verificar que os estudos filtrados quase em sua totalidade buscam resolver problemas de instituições de ensino vinculadas aos autores, os quais possuem uma base comum geralmente envolvendo professores, turmas, salas e disciplinas. Em seguida, cada estudo extrapola a problemática para realidades individuais onde são acrescentadas diferentes restrições ao problema.

Os cenários que descrevem os problemas resolvidos nem sempre são apresentados com riqueza de detalhes, podendo em alguns casos apresentar dados parciais, ou ainda em casos extremos, apenas descrever de qual instituição os dados foram coletados (estudos T-5, T-19, T-24 e T-28), o que dificulta a apresentação de informações nesta RSL.

Buscando descrever as características dos problemas resolvidos nos estudos, a Figura 2 ilustra graficamente o comportamento observado em relação a cinco variáveis que estão comumente presentes na definição de grades de horários. São referentes a número: de salas; de alunos; de turmas, de disciplinas e de professores. Observa-se que a figura traz quantitativamente esses números separados de acordo com a observação do valor mínimo, máximo e médio encontrados nos estudos. Conforme esperado, o maior valor relatado, tanto individualmente, quanto na média é em relação ao número de alunos. Porém, esse dado é o menos utilizado nos estudos analisados, aparecendo de forma explícita apenas em 5 dos 31 estudos (ver Figura 3). Ainda na Figura 3 é possível verificar que as variáveis mais utilizadas pelos autores se referem ao número de professores, disciplinas e turmas. Para maioria dos estudos, ao menos duas ou três dessas variáveis são utilizadas em paralelo para caracterizar o problema tratado.

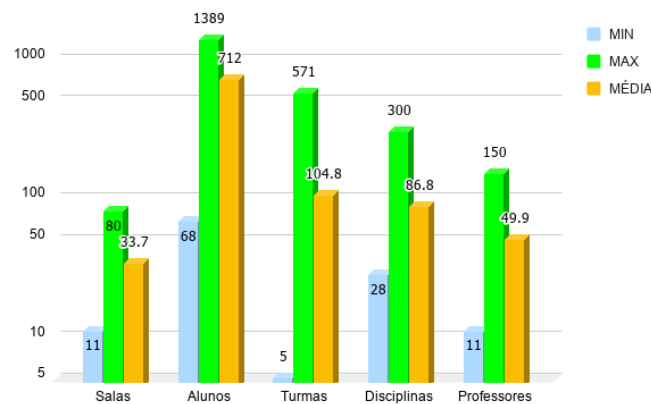


Figura 2: Composições mínima, máxima e média de características dos problemas resolvidos.

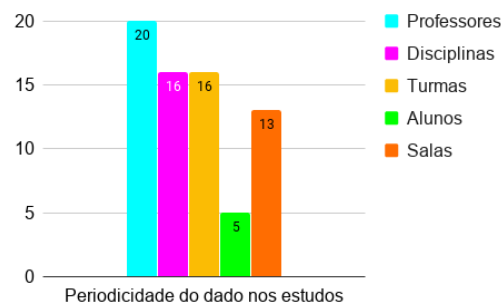


Figura 3: Quantidade de vezes que determinado dado é explicitamente reportado nos estudos analisados.

3.3. Características dos experimentos: Q-3

Nesta última questão de pesquisa analisada, são compiladas quatro subquestões que abordam características dos experimentos que foram realizados nos estudos. Na Figura 4 é ilustrado um gráfico que apresenta em intervalos a quantidade de testes realizados nos estudos selecionados. Assim, respondendo à questão Q-3.1, aproximadamente 55% (17) dos estudos realizam no máximo 10 testes distintos com as abordagens propostas. Esses testes são geralmente realizados por meio de alterações na soluções heurísticas propostas ou utilizando diferentes instâncias teste que são capazes de abordar variações em relação ao número de professores, salas, turmas, disciplinas ou alunos. Ressalta-se aqui um ponto positivo encontrado em três trabalhos (T-6, T-10 e T-27), os quais fazem uso de instâncias propostas na *International Timetabling Competition* (ITC) nos anos de 2007 e 2011, possibilitando assim que comparações sejam feitas de forma mais justa.

A segunda subquestão de pesquisa (Q-3.2) busca responder se as instâncias teste propostas são provenientes de trabalhos previamente publicados, de instituições de ensino ou criadas artificialmente pelos autores. Após levantamento, constatou-se que em grande parte dos estudos (30) são utilizadas instâncias teste provenientes de instituições de ensino ou de competições como a ITC. Apenas em três estudos os autores criaram novas instâncias para realizar os testes computacionais, sendo que, dois desses estudos testam suas abordagens utilizando também instâncias pré-definidas.

A questão de pesquisa Q-3.3 tem por objetivo responder se os testes realizados são baseados em casos reais de instituições de ensino ou se foram criadas instâncias artificiais que somente simulam demandas de instituições fictícias. Para esta questão, concluiu-se que por ser um problema

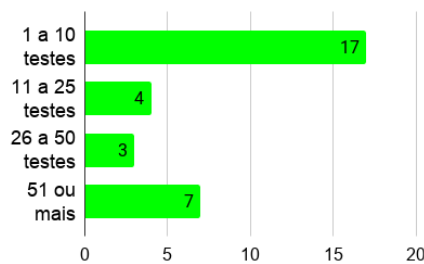


Figura 4: Quantidade de testes distintos realizados para validação do estudo.

vivenciado ao menos uma vez ao ano pelas instituições de ensino, a maioria dos autores opta por utilizar dados que refletem a realidade vivenciada pelos gestores dessas instituições ou coordenadores de cursos. Na Figura 5 essa informação fica mais explícita, pois em apenas um dos estudos (T-5) os autores não expuseram de forma clara a composição ou procedência das instâncias, levando-nos a considerar a mesma como artificial. Para outros 5 estudos, os autores primeiro testam suas abordagens em instâncias reais e, posteriormente, aplicam em instâncias com características artificiais.

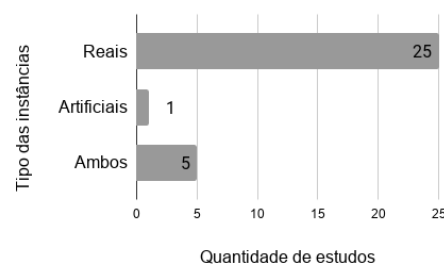


Figura 5: Quantidade de estudos utilizando tipos distintos de instâncias.

Por fim, a última questão de pesquisa Q-3.4 busca verificar se as técnicas propostas individualmente por cada estudo são comparadas a resultados preliminares, podendo ser provenientes de soluções manuais ou baseadas no uso de softwares. Foi identificado que as soluções de software propostas em 20 estudos são comparadas às soluções obtidas de forma manual pelas instituições, que em outros 15 estudos, suas soluções propostas são comparadas a soluções previamente obtidas por meio de softwares pagos ou publicados em pesquisas científicas e, que em alguns estudos ocorre dupla comparação, com soluções manuais e de software. Para o estudo T-19, foi identificada somente uma comparação entre duas abordagens de software distintas propostas no mesmo trabalho.

3.4. Avaliação da qualidade dos trabalhos

Em grande parte dos trabalhos, o critério de qualidade C-2 é difícil de ser quantificado, pois cada estudo segue uma padronização de apresentação dos resultados, os quais muitas vezes são apresentados de forma subjetiva (resultados bons, melhores, maioria dos casos foi superior, bastante satisfatório etc.).

Outros fatores relevantes observados em relação aos critérios C-5 e C-6 são a extensa utilização de métodos empíricos para definição de parâmetros utilizados nas heurísticas e a baixa quantidade de trabalhos que fazem uso de técnicas estatísticas para calibração de parâmetros e validação de qualidade das soluções obtidas pelas abordagens propostas.

Após a verificação individual dos estudos, foi realizada a atribuição de pontuações de acordo com os critérios de qualidade definidos na Seção 2. O estudo T-20 obteve a maior pontuação em relação aos demais, acumulando a soma de 5 pontos de 6 possíveis. Em seguida, os estudos T-1 e T-27 obtiveram pontuações idênticas (4,6), completando assim o ranking dos três melhores estudos.

4. Conclusões

A presente RSL realizou um levantamento sobre os estudos científicos publicados na literatura acerca do Problema de Geração de Horários Escolares durante a última década. Após aplicação da *string* de busca definida, foram retornados 1280 documentos pelo motor de busca. Desses, por meio de uma análise transversal do título e do resumo, foi possível identificar 338 documentos duplicados e outros 854 que se encaixaram no primeiro critério de exclusão definido. Após essa filtragem, 88 documentos foram analisados e novamente filtrados, restando assim 31 estudos que se encaixavam nos critérios definidos para aceitação.

Durante as pesquisas realizadas por meio do motor de busca Google Acadêmico, considerando os critérios de exclusão definidos, foi constatado que apenas o estudo T-11 provê uma espécie de revisão da literatura acerca do PGHE, apesar do estudo ter sido inicialmente submetido em 2011 e publicado em 2012. Como a busca foi limitada a estudos publicados a partir de 2010, a RSL proposta neste artigo possui pequena intersecção com o estudo T-11, não caracterizando assim um retrabalho. Adicionalmente, apesar do estudo T-11 apresentar características de estudos secundários, o mesmo não foi excluído, pois também apresentava uma abordagem algorítmica para resolução do problema aqui tratado.

Conclui-se que em relação à questão Q-1, a técnica mais utilizada para resolução do PGHE é baseada em heurísticas populacionais que utilizam Algoritmo Genético e suas variações. Quanto à questão de pesquisa Q-2, os problemas apresentam características bem diversas, seja em relação ao número de turmas, professores, disciplinas ou outras restrições específicas à instituição de ensino abordada.

Quanto à Q-3, não existe um padrão que consiga caracterizar as instâncias teste tratadas na maioria dos estudos, pois a estrutura do PGHE pode sofrer grandes variações de uma instituição de ensino para outra. Mesmo que não seja possível realizar uma comparação par a par, essa característica de aplicação prática inerente à instituição deve ser vista como positiva, pois pode promover avanços na forma com que as grades de horários são confeccionadas, poupando muito tempo e esforço dos profissionais que são responsáveis pela alocação muitas vezes totalmente manual ou parcialmente automatizada desses horários.

Ainda que vários trabalhos acerca do PGHE tenham sido anualmente publicados em congressos e revistas nacionais e internacionais nas últimas décadas, é representativa a quantidade de instituições de diferentes tamanhos que ainda utilizam a definição de horários de forma manual. Algumas razões podem motivar os responsáveis pela definição de horários a optarem pelo formato manual, tais como o alto custo de softwares comercializados especificamente para tal propósito, a quantidade de restrições particulares a cada instituição e o desconhecimento das várias pesquisas científicas que são publicadas acerca do problema. Essa última pode ser a mais preocupante, pois sugere que a população em geral não está sendo alcançada e, conseqüentemente, beneficiada pelas pesquisas que são realizadas por estudantes e pesquisadores brasileiros.

Por fim, foi identificado que ainda existe uma lacuna a ser preenchida no campo de pesquisas acerca do PGHE, que trata da generalização das abordagens propostas, de forma a contemplar a definição de horários desde a pequena instituição a nível municipal, até grandes instituições públicas e privadas que precisam muitas vezes gerir uma grande quantidade de recursos.

As ameaças e limitações à validade do presente trabalho foram minimizadas por meio do acompanhamento direto dos orientadores da pesquisa. Apesar de ter sido utilizada uma base única para a pesquisa, o viés de publicação foi evitado verificando se outras bases traziam trabalhos distintos, o que não foi constatado. De forma complementar, a *string* de busca final foi composta por termos que consideram sinônimos empregados na literatura para referenciar o PGHE.

Para trabalhos futuros, pretende-se formular um modelo para alocação de horário escolar para um curso superior noturno, cuja busca pela solução ótima é mais restritiva. Isto implica em explorar duas linhas distintas: a formulação de programação matemática que minimize as idas de docentes à instituição, uma vez que muitos deles residem em outros municípios, e a proposição de um método heurístico para encontrar uma solução inicial e, subsequentemente, obter uma solução próxima ao ótimo. Pretende-se que a heurística seja modelada para implementação em plataforma Web, de modo que seja de fácil acessibilidade aos gestores educacionais.

Agradecimentos

Os autores gentilmente agradecem o apoio financeiro recebido via Edital 20/2019 do Instituto Federal de São Paulo e a colaboração dos membros do Grupo de Pesquisa de Aplicações em Matemática e Computação (GPAMC) do Instituto Federal de São Paulo Campus Campos do Jordão.

Referências

- Alves, P. Y. A., Delgado, K. V., e da Silva, V. F. (2018). Problema de roteamento e inventário com janelas de tempo: Uma revisão sistemática da literatura. In *Anais do XIV Simpósio Brasileiro de Sistemas de Informação*, p. 222–215. SBC.
- Arenales, M., Armentano, V., Morabito, R., e Yanasse, H. (2015). *Pesquisa Operacional*. Elsevier, Rio de Janeiro.
- Babaei, H., Karimpour, J., e Hadidi, A. (2015). A survey of approaches for university course timetabling problem. *Computers & Industrial Engineering*, 86:43–59. ISSN 03608352. URL <https://linkinghub.elsevier.com/retrieve/pii/S0360835214003714>.
- Kitchenham, B. A. (2012). Systematic review in software engineering: where we are and where we should be going. In *Proceedings of the 2nd international workshop on Evidential assessment of software technologies*, p. 1–2.
- Pillay, N. (2014). A survey of school timetabling research. *Annals of Operations Research*, 218(1):261–293. ISSN 0254-5330. URL <http://link.springer.com/10.1007/s10479-013-1321-8>.
- Sousa, M. M., Ochi, L. S., e de Lima Martins, S. (2019). Traveling salesperson problem with hotel selection: A systematic review of the literature. In *Proceedings of the XV Brazilian Symposium on Information Systems*, p. 1–8, Aracaju.
- Tassopoulos, I. X. e Beligiannis, G. N. (2012). Solving effectively the school timetabling problem using particle swarm optimization. *Expert Systems with Applications*, 39(5):6029–6040. ISSN 09574174. URL <http://dx.doi.org/10.1016/j.eswa.2011.12.013>.