

Formulações e Algoritmos para Problemas de Programação de Horários Escolares Modelados em XHSTT

Plano de Curso de Doutorado apresentado ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais (PPGEE/UFMG) como parte das exigências para ingresso no programa em nível de doutorado.

Área de Concentração: Sistemas de Computação e Telecomunicações. Linha de Pesquisa: Otimização.

Belo Horizonte, Julho, 2014.

Aluno - George Henrique Godim da Fonseca

Orientador - Eduardo Gontijo

1 Resumo

O Problema da Programação de Horários Escolares consiste em alocar horários e recursos a aulas respeitando a diversas restrições dadas *a priori*. Um bom quadro de horários melhora a satisfação das entidades envolvidas, melhora o desempenho dos estudantes e a gestão dos recursos da instituição. Além de sua importância prática, o problema é da classe $\mathcal{NP}$ -Difícil e avanços na solução de problemas dessa categoria são um desafio para a comunidade científica de otimização. Devido à relevância do problema em questão, a *Third International Timetabling Competition* (ITC2011) foi proposta. A competição estimulou o desenvolvimento de diversos softwares de qualidade para a solução do problema bem como instaurou um grande conjunto de instâncias *benchmark* [10]. Apesar do surgimento de bons softwares para a solução do problema, é dito que há muito espaço para melhorias nos resolvedores desenvolvidos [3, 8]. Ademais, pouco esforço foi feito no sentido de prover limites inferiores e superiores para as instâncias e para provar a otimalidade de soluções. Este projeto então propõe o estudo e aprimoramento de formulações e algoritmos para o modelo do Problema de Programação de Horários proposto pela ITC2011. Em específico, pretende-se (1) desenvolver novos movimentos de vizinhança junto à abordagem de busca local existente; (2) incrementar a formulação recentemente proposta através de técnicas como cortes, geração de colunas e decomposição; (3) desenvolver uma heurística matemática com a formulação e algoritmo desenvolvidos.

2 Objetivos

O principal objetivo do presente trabalho é desenvolver, aprimorar e avaliar abordagens exatas e heurísticas para o Problema da Programação de Horários Escolares expressos no formato XHSTT [13], formato adotado pela ITC2011. Como objetivos específicos, cita-se:

- Melhorar os limites inferiores conhecidos para as instâncias do problema;
- Melhorar as melhores soluções conhecidas do problema através dos algoritmos e formulações a serem desenvolvidos;
- Propor a primeira heurística de programação matemática ao problema;
- Avaliar técnicas de programação matemática aplicadas ao problema, como cortes e geração de colunas.
- Verificar a viabilidade da modelagem de outros problemas de agendamento fora do contexto educacional no formato XHSTT.

3 Relevância da Contribuição

A programação automática de horários escolares tem sido alvo de diversas pesquisas nas áreas de Pesquisa Operacional e Inteligência Artificial. Em Santos e Souza [14] são apresentadas algumas das razões para este interesse:

- dificuldade de resolução: encontrar um quadro de horários que satisfaça todos os interesses envolvidos é uma tarefa difícil, ademais, frequentemente, a simples construção de um quadro de horários válido já é uma tarefa complicada;
- importância prática: a confecção de um bom quadro de horários pode melhorar a satisfação do corpo docente e permitir que a instituição de ensino seja mais eficiente na gestão de seus recursos, além do mais, a programação adequada das atividades letivas permite um melhor desempenho dos alunos;
- importância teórica: o problema apresentado neste trabalho é classificado como $\mathcal{NP}$ -Difícil [5] e progressos na solução de problemas desse tipo são um dos grandes objetivos das pesquisas correntes em computação, matemática e pesquisa operacional.

O estudo da literatura nos leva a concluir que os métodos heurísticos de busca local vêm obtendo as melhores soluções para essa categoria de problemas [17, 11]. Apesar de abordagens heurísticas geralmente atingirem soluções satisfatórias em um tempo de processamento viável, a obtenção de limites inferiores/superiores fortes e de soluções provadamente ótimas para o problema continuam um desafio em aberto.

Recentemente, devido ao grande avanço nos resolvedores comerciais de Programação Linear Inteira Mista [1], abordagens exatas vem sendo aplicadas com sucesso a diversos problemas de agendamento, como o Problema de Escalonamento de Equipes de Trabalho [15], o Problema de Sequenciamento de Tarefas em Máquinas [7] e o Problema de Planejamento de Projetos [18].

Apesar do avanço no uso de abordagens exatas em diversos problemas de agendamento, essa abordagem foi pouco explorada na competição dado que não era autorizado o uso de software comercial de otimização na mesma. Após a competição, Kristiansen *et al.* [8] propuseram a primeira formulação matemática para o problema considerado na competição. Apesar de ser uma formulação simples, foi possível com a mesma obter melhores (ótimas) soluções para duas das instâncias da competição e provar a otimalidade para outras quatro soluções previamente conhecidas.

4 Metodologia

O desenvolvimento do projeto se inicia com o curso de disciplinas pertinentes à pesquisa. Essas disciplinas, junto com o conhecimento obtido pelo candidato na graduação e no mestrado constituirão o embasamento teórico para a realização da pesquisa. As etapas principais para o desenvolvimento deste trabalho são assim definidas ¹:

1. Revisão da literatura - Concomitantemente com o curso de disciplinas e ao longo de todo o projeto, o candidato se manterá atualizado sobre os principais avanços na área através de constante pesquisa bibliográfica.
2. Desenvolvimento de formulação matemática - Conforme mencionado, recentemente foi desenvolvida a primeira formulação matemática para o problema em questão. Essa etapa compreende da reprodução da formulação desenvolvida por Kristiansen *et al.* [8] e de desenvolvimento de eventuais formulações alternativas ao problema.
3. Aprimoramento da formulação matemática - Diversos relatos na literatura apontam para grandes ganhos no desempenho advindos do aprimoramento de formulações matemáticas [2, 15, 16]. Duas técnicas importantes nesse contexto são a geração de cortes e a geração de colunas. A primeira consiste da inclusão de restrições que removem (cortam) uma ou mais soluções fracionárias mas não removem nenhuma solução inteira válida do espaço de busca, tornando a formulação mais eficiente; a segunda consiste em gerar padrões de pedaços de solução e encontrar a combinação de padrões ótima (*pricing*) [6].
4. Aprimoramento de heurísticas - Atualmente uma das melhores heurísticas conhecidas para o problema abordado é de autoria do candidato em conjunto com o GOAL (Grupo de Otimização e ALgoritmos) da Universidade Federal de Ouro Preto [3]. Apesar da heurística proposta ter vencido a ITC2011, os autores ressaltam que diversas extensões do trabalho podem ser desenvolvidas, como novos movimentos de vizinhança e a avaliação de outras metaheurísticas para o problema [4].
5. Desenvolvimento de heurística matemática - Apesar de heurísticas matemáticas estarem sendo aplicadas com sucesso a vários problemas difíceis em otimização [9, 18, 2], nenhuma heurística matemática foi proposta para o problema considerado na ITC2011. Uma heurística matemática é uma heurística que se utiliza iterativamente da solução de modelos matemáticos para alcançar soluções. A cada iteração, o modelo matemático está livre para otimizar um subconjunto de variáveis do problema. Apesar de não serem capazes de provar a otimalidade das soluções, heurísticas matemáticas são capazes de explorar eficientemente o espaço de soluções.

¹Note que etapas podem ser adicionadas ou removidas de acordo com o andamento do trabalho.

Essa etapa consiste do desenvolvimento da primeira heurística matemática para o problema proposto pela ITC2011.

6. Avaliação computacional das técnicas desenvolvidas - Esta etapa consiste da realização de experimentos computacionais sistemáticos sobre as técnicas desenvolvidas no presente trabalho. Para a avaliação das mesmas será considerado o conjunto de instâncias do problema provido pela competição [12]. Esses resultados serão posicionados com relação aos resultados já conhecidos na literatura.

Os principais resultados obtidos na presente pesquisa serão submetidos, em forma de artigo científico a periódicos de reconhecida qualidade e alto fator de impacto na área de otimização, tais como *Computers & Operations Research*, *Annals of Operations Research* e *European Journal of Operations Research*.

5 Cronograma

Semestre Atividade	1	2	3	4	5	6	7	8
Curso de disciplinas	X	X						
Desenvolvimento do proj. de qualificação		X	X	X				
Redação do projeto de qualificação				X				
Exame de qualificação				X				
Desenvolvimento do projeto de tese					X	X	X	
Redação da tese							X	X
Defesa da tese								X

6 Disciplinas a serem Cursadas

A obtenção de créditos para almejar o título de Doutor em Engenharia Elétrica se dará da seguinte forma:

- Computação Evolucionária - 2014/1
- Técnicas Clássicas de Reconhecimento de Padrões 2014/1 ²
- Otimização Multiobjetivo - 2014/2
- Otimização em Engenharia Elétrica - 2015/1
- Projeto Orientado I 2015/1
- Projeto Orientado II 2015/2

²As disciplinas Computação Evolucionária e Técnicas Clássicas de Reconhecimento de Padrões foram cursadas como disciplina isolada pelo candidato.

Referências

- [1] Robert E Bixby. A brief history of linear and mixed-integer programming computation. *Documenta Mathematica*, pages 107–121, 2012.
- [2] Ártón P. Dorneles, Olinto C.B. de Araújo, and Luciana S. Buriol. A fix-and-optimize heuristic for the high school timetabling problem. *Computers & Operations Research*, (0):–, 2014.
- [3] George Fonseca, Haroldo Santos, Tulio Toffolo, Samuel Brito, and Marcone Souza. *A SA-ILS approach for the High School Timetabling Problem*. PATAT '12 Proceedings of the 9th International Conference on the Practice and Theory of Automated Timetabling. 2012.
- [4] George H.G. Fonseca and Haroldo G. Santos. Variable neighborhood search based algorithms for high school timetabling. *Computers & Operations Research*, (0):–, 2014.
- [5] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco, CA, USA, 1979.
- [6] M.C. Goldbarg and H.P.L. Luna. *Otimização combinatória e programação linear: modelos e algoritmos*. Campus, 2000.
- [7] Olivier Guyon, Pierre Lemaire, Eric Pinson, and David Rivreau. Solving an integrated job-shop problem with human resource constraints. *Annals of Operations Research*, 213(1):147–171, 2014.
- [8] Simon Kristiansen, Matias Sorensen, and Thomas R. Stidsen. Integer programming for the generalized (high) school timetabling problem. *Journal of Scheduling (to appear)*, (0):–, 2014.
- [9] Vittorio Maniezzo, Thomas Stutzle, and Stefan Vob, editors. *Matheuristics - Hybridizing Metaheuristics and Mathematical Programming*, volume 10 of *Annals of Information Systems*. Springer, 2010.
- [10] University of Twente. International Timetabling Competition 2012, 2012. Available at <http://www.utwente.nl/ctit/hstt/itc2011/welcome/>, Acessado em dezembro / 2012.
- [11] Nelishia Pillay. A survey of school timetabling research. *Annals of Operations Research*, pages 1–33, 2013.
- [12] G. Post, L. di Gaspero, J. H. Kingston, B. McCollum, and A. Schaerf. The third international timetabling competition. In *Proceedings of the 9th International Conference on Practice and Theory of Automated Timetabling*, Son, Norway, 2012.

- [13] Gerhard Post, Jeffrey Kingston, Samad Ahmadi, Sophia Daskalaki, Christos Gogos, Jari Kyngas, Cimmo Nurmi, Nysret Musliu, Nelishia Pillay, Haroldo Santos, and Andrea Schaerf. XHSTT: an XML archive for high school timetabling problems in different countries. *Annals of Operations Research*, page 1–7, 2011. 10.1007/s10479-011-1012-2.
- [14] H. G. Santos and M. J. F. Souza. Programação de horários em instituições educacionais: formulações e algoritmos. In *XXXIX Simpósio Brasileiro de Pesquisa Operacional, Fortaleza. Anais do XXXIX SBPO. Rio de Janeiro : SOBRAPO*, pages 1 : 2827–2882. 2007.
- [15] Haroldo G. Santos, Túlio A.M. Toffolo, Rafael A.M. Gomes, and Sabir Ribas. Integer programming techniques for the nurse rostering problem. *Annals of Operations Research*, pages 1–27, 2014.
- [16] Haroldo Gambini Santos, Eduardo Uchoa, Luiz Satoru Ochi, and Nelson Maculan. Strong bounds with cut and column generation for class-teacher timetabling. *Annals of Operations Research*, 194(1):399–412, 2012.
- [17] A. Schaerf. A survey of automated timetabling. *Artificial Intelligence Review*, 13(2):87–127, 1999.
- [18] Túlio A. M. Toffolo, Haroldo G. Santos, Marco A. M. Carvalho, and Janiele A. Soares. An integer programming approach for the multi-mode resource-constrained multi-project scheduling problem. In *MISTA 2013 - 6th Multidisciplinary International Scheduling Conference: Theory & Applications, 2013, Gent, Belgium*, pages 840–848. 2013.