Uma abordagem exata para o problema da distribuição de disciplinas

Leopoldo Santos Silva

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO DA FACOM-UFMS

Data da Defesa:

Assinatura:______________________________

Uma abordagem exata para o problema da distribuição de disciplinas

Leopoldo Santos Silva

Orientador: *Edna Ayako Hoshino*

Monografia apresentada como requisito parcial para aprovação na Componente Curricular Não-Disciplinar Trabalho de Conclusão de Curso do curso de Graduação em Ciência da Computação, da Universidade Federal de Mato Grosso do Sul.

UFMS - Campo Grande
julho/2024

Resumo

Este trabalho propõe uma abordagem de programação linear inteira para resolver o problema da distribuição de disciplinas. São considerados como fatores relevantes para o modelo desenvolvido, o limite de carga horária semestral de cada docente, a aptidão, que relaciona cada disciplina às áreas de atuação dos docentes, e a satisfação individual dos professores, a qual se busca maximizar através do modelo desenvolvido. Ainda que se trate de um problema NP-difícil, foi possível obter soluções ótimas em um período de tempo relativamente curto para instâncias de complexidade média. Testes computacionais realizadas em instâncias geradas de forma pseudo-aleatória e em uma instância real mostram que a aplicação desta abordagem no âmbito da Facom é viável.

Sumário

CAPÍTULO

1

Introdução

Para o funcionamento de uma instituição de ensino é vital que sejam definidos os responsáveis por ministrar os cursos em cada uma das turmas existentes. Essa definição deve respeitar uma série de restrições, sendo que algumas são comuns entre diversas instituições e outras podem advir de particularidades, sejam elas advindas de questões contratuais ou de necessidades levantadas pelos tópicos abordados.

1.1 Justificativa

O processo de distribuição de turmas em uma universidade tem forte impacto no corpo docente. Devido as demandas legais (como os limites mínimo e máximo para a carga horária de cada professor) e pedagógicos (como aptidão para ministrar ou não uma disciplina), é comum que não seja possível atribuir as turmas desejadas aos professores.

Além do mais, a complexidade do problema cresce consideravelmente com o aumento do número de turmas e docentes, eventualmente tornando inviável o uso de uma solução manual (como ainda é feito em muitas instituições).

Atualmente a FACOM já utiliza um programa para resolver o processo de distribuição de disciplinas, o qual é baseado em algoritmos para o problema de emparelhamentos em grafos. Ainda assim, existem vantagens em utilizar uma abordagem do ponto de vista da programação linear inteira, como por exemplo a facilidade de remover ou adicionar restrições ao algoritmo, além de fornecer uma solução ótima, ou suficientemente próxima dela.

1.2 Objetivos

O objetivo deste trabalho é propor e analisar o uso de algoritmos exatos baseados em programação linear inteira para resolver o problema da distribuição de disciplinas.

1.3 Organização do texto

O primeiro capítulo contextualiza o problema inicial que levou ao desenvolvimento da pesquisa, bem como os objetivos desta. O segundo capítulo apresenta uma descrição formal do problema e faz uma revisão dos trabalhos já desenvolvidos no tema, e o terceiro apresenta os conceitos necessários para entendimento do projeto desenvolvido, descrito no Capítulo 4. Os testes realizados e os resultados obtidos a partir deste estão dispostos no Capítulo 5, enquanto o Capítulo 6 descreve as conclusões obtidas e relaciona possíveis trabalhos a serem desenvolvidos futuramente.

CAPÍTULO

2

Descrição do Problema e Trabalhos Relacionados

Este capítulo descreve formalmente o problema tendo em vista as necessidades da FACOM e apresenta alguns dos trabalhos já realizados sobre o tema, ou assuntos relacionados, e suas principais contribuições.

2.1 Descrição formal do problema

Sejam $\mathcal{P}$ e $\mathcal{T}$ os conjuntos que contêm, respectivamente, os professores e turmas a serem ministradas. Considere um peso w_{ij}, associado a cada par $(i, j) \in \mathcal{P} \times \mathcal{T}$, representando quão desejável uma turma j é ao professor i. Cada professor $i \in \mathcal{P}$ tem um valor mínimo de carga horária anual e um valor máximo para a carga horária a ser atribuída no primeiro e no segundo semestre, sendo eles denotados por $CHmin_i$, $CHmax_i^1$ e $CHmax_i^2$, respectivamente. Além disso, toda turma $j \in \mathcal{T}$ está relacionada a um único semestre e possui um valor correspondente a sua carga horária CH_j. Os subconjuntos $\mathcal{T}_1$ e $\mathcal{T}_2$ são os subconjuntos maximais que contém, respectivamente, todas as turmas relacionadas ao primeiro e segundo semestre.

O **problema da distribuição de disciplina** (DPD) consiste em atribuir as turmas aos professores de modo que:

- cada turma $j \in \mathcal{T}$ seja atribuída a exatamente um professor $i \in \mathcal{P}$;
- a soma da carga horária anual de cada professor $i \in \mathcal{P}$ seja no mínimo $CHmin_i$;

- a soma da carga horária semestral de cada professor $i \in \mathcal{P}$ não ultrapasse $CHmax_i^1$ e $CHmax_i^2$ no primeiro e no segundo semestre, respectivamente;

- a soma da satisfação dos professores, dada pela soma dos pesos w_{ij} seja máxima.

2.2 Trabalhos relacionados

O problema da atribuição de turmas a professores é amplamente estudado. Carter e Laporte (1998) o caracteriza como um subproblema do "Problema do horário", enquanto define seus principais componentes, sendo eles:

1. Curso: Uma disciplina ensinada uma ou mais vezes durante a semana, em uma ou mais turmas, que podem ser conduzidas por um ou mais professores.

2. Turma: Um grupo de estudantes cursando uma mesma disciplina

3. Programa: Uma série de cursos que juntos formam uma grade.

A este trabalho interessa somente a ideia de que uma disciplina é ensinada um determinado número de vezes na semana, resultando em uma carga horária, e que professores também podem ministrar diferentes instâncias de uma disciplina, as quais serão referidas daqui em diante utilizando o termo turma.

Além disso, Avella e Vasilyev (2005) compara diversas abordagens para a solução do problema do horário, demonstrando a eficiência de diferentes estrategias de geração de cortes do espaço de solução.

Hosny (2012) elabora algumas heurísticas, ainda que direcionadas ao uso no problema do horário, enquanto Szwarc et al. (2018) levantam a atenção sobre aptidão e afinidade do corpo docente sobre as turmas atribuídas, visando melhoria na qualidade educacional.

Al-Yakoob e Sherali (2006) segue uma abordagem mais próxima da utilizada neste trabalho, considerando que as turmas já têm seus horários pré-definidos, ainda que permita uma variação de até 15% posterior à atribuição. Já Domenech e Lusa (2016) considera que estes horários são dados e fixos. Ambos os trabalhos se utilizam de uma linearização do problema, e chegam a conclusão de que o uso do método gera resultados mais satisfatórios e de forma menos trabalhosa do que o uso de uma atribuição manual.

Em um trabalho subsequente, Al-Yakoob e Sherali (2013) trata o problema em duas fases, a partir da ideia de um conjunto de turmas pré-existente. Primeiramente as turmas devem ser distribuídas dentro do horários disponíveis de acordo com as preferências e restrições da universidade. Em seguida, com os horários fixados, os docentes devem ser atribuídos as turmas semanais de acordo com suas preferências. O trabalho utiliza o método de geração de colunas para resolver o problema, que se torna eficiente em cenários onde um grande número de variáveis do problema terão valor zero. A abordagem se mostra efetiva, atingindo resultados mais satisfatórios em um tempo milhares de vezes menor do que os atingidos em Al-Yakoob e Sherali (2006).

Além disso, Gunawan et al. (2008) aborda o problema considerando os créditos em cada turma, além de dividi-las em seções que podem ser ensinadas por mais de um professor. O método desenvolvido busca manter um equilíbrio entre as cargas atribuídas a cada professor.

Ressalta-se também que o problema da atribuição com base em um fator de aptidão mencionado anteriormente pode ser facilmente modificada para medir competência ou desempenho em diversas áreas, como feito em Szwarc et al. (2019), buscando distribuições de recursos, de forma que sejam robustas, eficientes e adaptáveis. Nota-se que parte significativa da literatura se volta para a solução do problema do horário, mas parte considerável das ideias podem ser utilizadas ou adaptadas para se encaixar nas especifidades do problema abordado neste trabalho.

Capítulo 3

Metodologia

Este capítulo descreve as ferramentas e métodos utilizados para atingir os objetivos propostos. Assim, para oferecer uma solução eficaz para o problema de distribuição de disciplinas, foi utilizada uma abordagem do ponto de vista da programação linear inteira (PLI). Os conceitos e definições apresentadas neste capítulo são baseadas nos livros de Vanderbei (2014) e Wolsey (2020).

3.1 Programação linear

Programação linear (PL) consiste em um método de otimização matemática, que visa maximizar ou minimizar o valor de uma função linear, referida como **função objetiva**, definida a partir de um conjunto de variáveis chamadas **variáveis de decisão**. Além disso, o modelo também pode estar sujeito a um conjunto de restrições, também dependentes dos valores das variáveis de decisão. Um modelo padrão, com n variáveis de decisão x_i com pesos c_i na função objetivo e sujeitas a um conjunto de m restrições é da forma:

$$\max \quad \sum_{i=1}^{n} c_i x_i \tag{3.1}$$

$$\text{s.a.} \sum_{i=1}^{n} a_{ij} x_i \leq b_j \qquad \forall j \in \{1, \ldots, m\} \tag{3.2}$$

$$x_i \geq 0 \qquad \forall i \in \{1, \ldots, n\} \tag{3.3}$$

3.1.1 Programação linear inteira

Ainda no contexto da PL, o termo **programação linear inteira** (PLI) se refere a um subconjunto da PL, no qual as variáveis de decisão estão limitadas a valores inteiros. Além disso, é importante notar que esta variação não necessariamente torna o problema mais simples. Pelo contrário, problemas de PLI muitas vezes são NP-difíceis. Para resolvê-los, utilizamos técnicas específicas, como por exemplo *Branch-and-Bound*. Um modelo padrão de maximização para um problema de PLI é descrito em seguida:

$$\max \quad \sum_{i=1}^{n} c_i x_i \tag{3.4}$$

$$\text{s.a.} \sum_{i=1}^{n} a_{ij} x_i \leq b_j \qquad \forall j \in \{1, \ldots, m\} \tag{3.5}$$

$$x_i \in \{0, 1\} \qquad \forall i \in \{1, \ldots, n\} \tag{3.6}$$

3.2 Relaxação linear

Durante a resolução de problemas de PLI um passo importante é limitar o espaço de busca. Entre as técnicas utilizadas para isso está a relaxação linear, que consiste em simplificar o problema, assumindo que os valores das variáveis de decisão pertencem a todo o conjunto dos reais, e não somente dos inteiros, simplificando o problema para um problema de PL. Com isto é possível obter uma solução aproximada, chamada de solução relaxada. Esta técnica pode ser utilizada para obter uma solução inicial a ser explorada em PLI, ou para podar o espaço de busca no algoritmo de *Branch-and-Bound*.

3.3 Branch-and-bound

Branch-and-Bound é uma técnica que utiliza o paradigma de divisão e conquista, sendo especialmente útil em se tratando de um problema de otimização combinatória. Um algoritmo básico pode ser dividido em algumas etapas:

Inicialização. É utilizada uma árvore de decisões, com o nó inicial correspondendo ao problema original.

Branching. O espaço de busca é particionado em subproblemas, criando novos nós com bases em escolhas sobre o nó pai.

Bounding. São calculados os limitantes superior, que representa o valor máximo possível de ser alcançado (geralmente obtido pela relaxação linear) e inferior, representando a melhor solução encontrada (a qual pode ser obtida usando-se heurísticas).

Poda. Nesta fase o algoritmo checa se os nós podem ou não conter soluções melhores do que as já encontradas, utilizando os limitantes calculados na etapa de *bounding*. Caso isso não seja possível, corta-se a subárvore, removendo-a do espaço de busca.

CAPÍTULO 4

Trabalho Desenvolvido

Este capítulo apresenta o modelo desenvolvido para resolver o problema proposto, além de descrever melhorias desenvolvidas posteriormente, de forma a atender melhor os requisitos de uso, obtendo resultados mais desejáveis ao corpo docente e trazer resoluções a uma maior quantidade de instâncias.

4.1 Modelo inicial

Considere variáveis de decisão x_{ij} que relacionam um professor i a uma turma j, tendo valor 1 caso ela pertença à solução e 0 caso contrário.

Dessa forma, o modelo inicial de PLI desenvolvido para o DPD é dado por:

$$
\begin{aligned}
\max \quad & \sum_{i \in \mathcal{P}} \sum_{j \in \mathcal{T}} w_{ij} x_{ij} & & & (4.1)\\
\text{s.a.} \quad & \sum_{i \in \mathcal{P}} x_{ij} = 1 & \forall j \in \mathcal{T} & & (4.2)\\
& \sum_{j \in \mathcal{T}} CH_j x_{ij} \geq CHmin_i & \forall i \in \mathcal{P} & & (4.3)\\
& \sum_{j \in \mathcal{T}_1} CH_j x_{ij} \leq CHmax_i^1 & \forall i \in \mathcal{P} & & (4.4)\\
& \sum_{j \in \mathcal{T}_2} CH_j x_{ij} \leq CHmax_i^2 & \forall i \in \mathcal{P} & & (4.5)\\
& x_{ij} \in \{0,1\} & \forall i \in \mathcal{P} & \forall j \in \mathcal{T} & (4.6)
\end{aligned}
$$

A função objetivo (4.1) visa maximização do peso, de forma a satisfazer, dentro dos limites estabelecidos os desejos de todos os docentes. As restrições (4.2) garantem que cada turma seja atribuída a exatamente um professor. As restrições (4.3) à (4.5) fazem com que todos os professores recebam uma carga horária que satisfaça os valores mínimos e máximos correspondentes a ele em cada semestre, enquanto as restrições (4.6) são restrições de integralidade.

4.2 Áreas de atuação

A primeira melhoria implementada se baseia em levar em consideração as áreas das disciplinas e as áreas de atuação dos professores. Inicialmente a restrição (4.2) foi ajustada para que x_{ij} seja multiplicado por um coeficiente inteiro a. Caso o professor esteja apto a é tido como 1, e caso não pertença à área, a assume um valor maior que 1 (efetivamente removendo-o da solução, pois caso contrário a restrição seria violada):

$$\sum_{i \in \mathcal{P}^k} x_{ij} + \sum_{i \notin \mathcal{P}^k} 2x_{ij} = 1, \forall j \in T, \text{ sendo } k \text{ a área da disciplina de } j.$$

no qual P^k é o conjunto dos professores da área k.

Foi constatado que esta abordagem gerava muitas soluções inviáveis, e portanto foi utilizada uma nova abordagem. Ao invés de alterar o conjunto de restrições, foram efetuadas mudanças na função objetiva (4.1), partindo do modelo inicial. Nos casos em que o professor i não pertence à área de atuação necessária para a turma j, ela ainda poderá ser alocada, porém com o peso w_{ij} assumindo um valor negativo recebido como parâmetro de entrada do algoritmo, fazendo com que ainda seja possível um professor atuar fora de sua área caso seja necessário para a solução, embora penalizando na função objetivo.

4.3 Medida de satisfação

Para analisar os resultados obtidos, e da possibilidade de levar em consideração as satisfações individuais dos docentes nas distribuições posteriores, como instrumento de balanceamento a longo prazo, se faz necessário estabelecer um coeficiente de satisfação, que leve em consideração os pesos atribuídos pelo professor e o resultado fornecido pelo sistema.
Desta forma, após análise foi definido C_i como o coeficiente que relaciona a satisfação do professor i através da Equação 4.7.

$$C_i = \frac{Wa_i}{We_i} \tag{4.7}$$

no qual, We_i representa o peso médio que o professor i atribuiu as disciplinas e Wa_i representa o peso médio atribuído pelo sistema.

4.4 Geração das instâncias de teste

A fim de realizar os testes necessários para medir a efetividade do método, se faz necessário um conjunto diverso de entradas de teste. Assim, foi desenvolvido um algoritmo capaz de gerar uma lista de preferências de forma randômica a partir de 2 arquivos de entrada, com um deles contendo os dados relativos aos professores (como nome, carga horária mínima e máxima por semestre, e posteriormente áreas de atuação) e outro com um conjunto inicial de disciplinas (nome, cursos relacionados, semestre e carga horária).
Com estas informações, o programa replica as disciplinas aleatoriamente, formando um conjunto de turmas, para as quais são atribuídos (também de forma aleatória, respeitando as áreas de atuação) valores para as preferências dos professores, gerando um numero de entradas de teste com densidade de escolhas baseada em um fator variável.

CAPÍTULO
5

Resultados Computacionais

Este capítulo apresenta o ambiente e os resultados computacionais obtidos através do modelo desenvolvido, além de analisar o que representam em termos de eficiência e viabilidade de aplicação.

5.1 Ambiente computacional

A implementação do algoritmo para a resolução do PLI e efetuação dos testes realizados foi feita utilizando a linguagem C e o compilador GCC 8.3.0, com auxilio do *framework* SCIP Gamrath et al. (2020). O modelo foi refinado conforme novos requisitos foram levantados durante os testes (conforme discutido no capítulo 3), e as seções aqui dispostas se referem a diferentes versões do algoritmo e do modelo de PLI.

Os testes foram executados em um contêiner, rodando em uma máquina com as seguintes especificações:

- Processador 12th Gen Intel(R) Core(TM) i7-1255U 1.70 GHz
- RAM 16Gb
- Windows 10
- SSD com 256Gb de armazenamento

5.2 Validação do modelo

As instâncias utilizadas nos testes são geradas com base em um conjunto de 20 disciplinas, idêntico para ambos os semestres, considerando 4 áreas de atuação, de forma que uma disciplina se relaciona a exatamente uma área, enquanto o professor pode estar relacionada a uma ou mais.

O conjunto de disciplinas é processado no gerador para que sejam criadas novas turmas com base nas existentes, de forma aleatória, até atingir o número desejado de turmas. Para os testes realizados, foram geradas instâncias com 50 e 100 turmas.

Em seguida, um conjunto de professores é utilizado para gerar suas preferências, também de maneira aleatória. Essa geração leva em consideração as áreas de atuação das disciplinas, e a quantidade de preferências atribuídas a cada professor é determinada por um fator que indica a probabilidade de selecionar uma turma quando as áreas de atuação coincidem, chamado de densidade e representado por ρ.

Os testes a seguir utilizam um mesmo conjunto de 20 instâncias, variando apenas parâmetros utilizados na execução. As instâncias consideram a existência de 4 áreas de atuação, com o corpo docente de $|\mathcal{P}|$ professores estando capacitado para atuar em 2 áreas em média, e com uma boa flexibilidade na carga horária mínima e máxima por semestre.

Em todas as tabelas dos experimentos, o número de disciplinas será representado por $|\mathcal{T}|$, a *Densidade* se refere à probabilidade, entre 0 e 1, de existir uma preferência relacionando um professor a cada disciplina dentro de sua área de atuação. *Fo* representa o valor atingido pela função objetiva.

Além disso, foi utilizado um tempo limite de 15 minutos na varredura do espaço de solução. Algumas das instâncias foram capazes de alcançar solução ótima no tempo estabelecido, e para estas será o tempo em segundos necessário para alcançar uma solução ótima. Para as instâncias que alcançaram o tempo limite, será mostrado o valor do gap atingido. Por fim, C_m é o coeficiente médio da satisfação entre todos os docentes na entrada.

O teste a seguir se refere à versão do código, no qual as áreas de atuação e histórico não eram levadas em consideração. A verificação de aptidão sobre as áreas de atuação foi desabilitada, produzindo os resultados dispostos na Tabela 5.1.

instância	$\vert\mathcal{P}\vert$	$\vert\mathcal{T}\vert$	ρ	Fo	Tempo / Gap	C_m
1	10	50	0.10	153	0.04s	0.453
2	10	50	0.15	223	0.97s	0.754
4	10	50	0.20	277	1.17s	0.841
5	10	50	0.25	301	0.54s	0.989
6	10	50	0.30	299	1.37s	1.000
6	10	50	0.35	348	0.69s	1.166
7	10	50	0.40	364	1.07s	1.219
8	10	50	0.45	380	1.27s	1.281
9	10	50	0.50	410	0.80s	1.368
10	10	50	0.55	387	1.04s	1.310
11	20	100	0.10	614	8.07s	1.036
12	20	100	0.15	678	154.59s	1.161
13	20	100	0.20	741	2.63s	1.333
14	20	100	0.25	782	3.30%	1.358
15	20	100	0.30	780	4.70%	1.353
16	20	100	0.35	865	788.06s	1.553
17	20	100	0.40	867	693.32s	1.574
18	20	100	0.45	885	761.56s	1.538
19	20	100	0.50	905	5.90s	1.578
20	20	100	0.55	660	746.70s	1.110

Tabela 5.1: Resultados sem considerar a aptidão sobre as áreas de atuação, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Analisando os resultados, existem duas constatações imediatas. A primeira é que quanto maior o número de docentes, turmas, e consequentemente de preferências, mais difícil o problema se torna. A segunda é que um maior número de preferências resulta em um valor maior para a função objetiva e também para a satisfação média dos professores, o que é esperado, pois nem sempre é possível atribuir turmas desejadas pelos professores, e um maior número de preferências aumenta a probabilidade de que a turma escolhida esteja na lista, mesmo que com um peso baixo. A satisfação média obtida por este método foi de 1.252.

Em seguida o mesmo conjunto de instâncias foi utilizado considerando um valor de -5 como penalidade na função objetiva caso um professor não pertença a área, resultando nos dados apresentados na Tabela 5.2.

instância	$\vert\mathcal{P}\vert$	$\vert\mathcal{T}\vert$	ρ	Fo	Tempo / Gap	C_m
1	10	50	0.10	132	0.43s	0.387
2	10	50	0.15	162	0.58s	0.557
4	10	50	0.20	188	0.32s	0.556
5	10	50	0.25	215	0.80s	0.686
6	10	50	0.30	222	0.71s	0.752
6	10	50	0.35	298	0.33s	0.969
7	10	50	0.40	295	0.51s	0.937
8	10	50	0.45	340	0.74s	1.171
9	10	50	0.50	370	1.00s	1.259
10	10	50	0.55	327	5.38s	1.159
11	20	100	0.10	459	386.52s	0.793
12	20	100	0.15	564	159.79s	0.970
13	20	100	0.20	593	63.74s	1.059
14	20	100	0.25	630	1.10%	1.134
15	20	100	0.30	547	27.36%	1.035
16	20	100	0.35	745	160.25s	1.304
17	20	100	0.40	797	245.48s	1.440
18	20	100	0.45	751	8.97%	1.322
19	20	100	0.50	767	9.73%	1.344
20	20	100	0.55	538	10.28s	0.866

Tabela 5.2: Resultados com penalização -5 para atuação fora da área de aptidão, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Comparando os resultados obtidos, percebe-se um declínio significativo no valor atingido. Isto indica que a distribuição anterior atribui turmas fora das áreas de atuação dos professores, mesmo que essa atribuição não acrescentasse a função objetiva (pois nãos constam nas preferências disciplinas fora da area de atuação, dando a elas peso 0 mesmo quando estas são ignoradas na resolução do PLI). O valor médio da satisfação decaiu para 1.032.

Definindo o valor da penalidade como -10, obtemos os resultados apresentados na Tabela 5.3:

instância	$\lvert\mathcal{P}\rvert$	$\lvert\mathcal{T}\rvert$	ρ	Fo	Tempo / Gap	C_m
1	10	50	0.10	132	0.55s	0.383
2	10	50	0.15	162	0.56s	0.557
4	10	50	0.20	188	0.60s	0.559
5	10	50	0.25	215	0.79s	0.672
6	10	50	0.30	222	0.84s	0.752
6	10	50	0.35	298	0.58s	0.969
7	10	50	0.40	295	1.19s	0.93
8	10	50	0.45	340	0.90s	1.119
9	10	50	0.50	370	1.49s	1.258
10	10	50	0.55	327	4.69s	1.010
11	20	100	0.10	459	512.66s	0.798
12	20	100	0.15	562	158.00s	0.962
13	20	100	0.20	593	50.99s	1.049
14	20	100	0.25	630	0.97%	1.123
15	20	100	0.30	543	27.76%	0.985
16	20	100	0.35	745	152.16s	1.312
17	20	100	0.40	797	301.99s	1.450
18	20	100	0.45	809	1.26%	1.403
19	20	100	0.50	826	1.79%	1.435
20	20	100	0.55	538	8.51s	0.859

Tabela 5.3: Resultados com penalização -10 para atuação fora da área de aptidão, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Comparando com os resultados anteriores, é possível notar pouca diferença nos valores atingidos. Na entrada 15 o valor foi de 547 para 543. Observa-se também que o gap é alto para os testes nessa entrada, revelando a possibilidade de não termos atingido solução ótima. Já para as entradas 18 e 19 houve diminuição considerável do gap, acompanhada de um aumento do valor atingido. Como a penalidade é maior neste teste, isso indica que nos testes anteriores existiam soluções melhores que não foram exploradas no tempo estipulado. As entradas restantes obtiveram resultados semelhantes em ambos os testes, um indicador de que foi possível encontrar distribuições onde a penalidade não foi utilizada, ou seja, todos os professores estarão atuando dentro de suas respectivas áreas. Também não houve grande variação na satisfação média, que foi de 1.035

A fim de comparação, o mesmo conjunto foi utilizado restringindo professores a serem atribuídos somente a turmas em áreas correspondentes. Os resultados desse teste estão apresentados na Tabela 5.4.

instância	$\|\mathcal{P}\|$	$\|\mathcal{T}\|$	ρ	Fo	Tempo / Gap	C_m
1	10	50	0.10	132	0.08s	0.349
2	10	50	0.15	162	0.27s	0.539
4	10	50	0.20	188	0.03s	0.551
5	10	50	0.25	215	0.49s	0.636
6	10	50	0.30	222	0.23s	0.752
6	10	50	0.35	298	0.55s	0.997
7	10	50	0.40	295	0.17s	0.937
8	10	50	0.45	340	0.21s	1.118
9	10	50	0.50	370	0.78s	1.263
10	10	50	0.55	327	4.95s	1.110
11	20	100	0.10	459	244.85s	0.794
12	20	100	0.15	562	44.67s	0.952
13	20	100	0.20	593	47.71s	1.033
14	20	100	0.25	630	1.13%	1.139
15	20	100	0.30	—	—	—
16	20	100	0.35	745	155.06s	1.304
17	20	100	0.40	797	275.66s	1.449
18	20	100	0.45	784	4.48%	1.354
19	20	100	0.50	809	4.06%	1.400
20	20	100	0.55	538	9.55s	0.859

Tabela 5.4: Resultados com proibição de atuação fora da área de aptidão.

Comparando com o teste anterior, se confirma a hipótese levantada de que as soluções que atingiram o mesmo valor na função objetiva consistiam em soluções onde a penalidade não foi utilizada. É interessante notar como este método acabou por obter resultados mais rapidamente.

Também se destaca que a existência de uma solução onde todos os professores atuem dentro de suas respectivas áreas não necessariamente é ótima. Isto acontece em casos onde optar por atribuir uma turma que gere penalidade possibilite novas alocações, que somadas resultem em um valor de satisfação igual ou maior à penalidade escolhida. Novamente, a satisfação média não teve grande variação, sendo de 1.029.

Como comentado em capítulos anteriores, esta abordagem gera maior inviabilidade de soluções, especialmente em casos não ideais, onde existe a falta de profissionais devidamente capacitados a lecionar certas áreas. Vale ressaltar que no caso de um corpo acadêmico selecionado para atender as necessidades impostas pelas diferentes áreas do conhecimento, a probabilidade de existência de uma solução viável é significativamente maior. O conjunto de disciplinas e professores utilizados nesta bateria de testes apresenta certo equilíbrio, mesmo com aleatoriedade na distribuição das turmas, viabilizando o uso desta abordagem, de forma que apenas a entrada 15 não obteve uma solução.

5.3 Caso complexo

A fim de obter maior variabilidade e levar o algoritmo a situações de estresse, foi gerado um novo conjunto de instâncias. Foi mantido o número de 4 áreas de atuação, porém agora cada professor atua em somente uma área, além de possuírem restrições mais rígidas de carga horária. Além disso, existe uma aleatoriedade maior na distribuição de turmas entre cada semestre. Todos os testes a seguir utilizam um total de 20 professores e 100 turmas, com densidade de escolhas variável.

Utilizando a versão do algoritmo onde professores são restritos as suas áreas de atuação não foi possível encontrar soluções dentro de todo o espaço de busca. Assim, a tabela a seguir descreve, como anteriormente, os resultados obtidos ao ignorar as áreas de atuação:

instância	$\lvert\mathcal{P}\rvert$	$\lvert\mathcal{T}\rvert$	ρ	Fo	Tempo / Gap	C_m
1	20	100	0.10	436	127.44s	0.804
2	20	100	0.15	503	566.65s	0.984
4	20	100	0.20	611	1.31%	1.082
5	20	100	0.25	653	662.99s	1.134
6	20	100	0.30	692	2.47%	1.222
6	20	100	0.35	—	—	—
7	20	100	0.40	—	—	—
8	20	100	0.45	777	5.05%	1.309
9	20	100	0.50	—	—	—
10	20	100	0.55	507	26.38s	0.923

Tabela 5.5: Resultados em instâncias mais complexas ignorando a área de aptidão, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Para algumas instâncias não foi possível encontrar uma solução dentro do tempo estabelecido. Isso se deve a esta abordagem ampliar o espaço de busca em alguns casos. A existência de uma penalidade acaba por eliminar rapidamente um grande conjunto de soluções ineficazes. Além disso, com as restrições mais rígidas na variação da carga horária mínima e máxima, a satisfação media de 1.066 se mostra inferior a obtida ao realizar o mesmo teste no conjunto anterior de instâncias, principalmente quando se consideram somente as entradas com 20 professores, que no primeiro teste obtiveram soluções acima de 1 para todos os casos.

Em seguida, como no conjunto anterior, foi utilizada uma penalidade de -5 para professores atuando fora da área, com os resultados sendo:

instância	$\|\mathcal{P}\|$	$\|\mathcal{T}\|$	ρ	Fo	Tempo / Gap	C_m
1	20	100	0.10	204	774.23s	0.439
2	20	100	0.15	276	11.00%	0.670
4	20	100	0.20	235	71.37%	0.534
5	20	100	0.25	311	62.65%	0.655
6	20	100	0.30	327	52.75%	0.745
6	20	100	0.35	—	—	—
7	20	100	0.40	460	20.80%	0.927
8	20	100	0.45	530	18.94%	0.994
9	20	100	0.50	429	27.46%	0.897
10	20	100	0.55	230	34.50%	0.506

Tabela 5.6: Resultados com penalização -5 para atuação fora da área de aptidão nas instâncias complexas, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Houve um alto valor de gap na maioria das entradas, sendo que mesmo ao obter uma solução ótima, o valor da função objetiva decresceu consideravelmente. A satisfação media obtida foi de 0.708
Por fim, a penalidade foi definida como -10 pontos, obtendo os resultados dispostos na tabela abaixo

instância	$\|\mathcal{P}\|$	$\|\mathcal{T}\|$	ρ	Fo	Tempo / Gap	C_m
1	20	100	0.10	141	13.85%	0.427
2	20	100	0.15	66	277.33%	0.402
4	20	100	0.20	302	19.04%	0.708
5	20	100	0.25	254	86.18%	0.704
6	20	100	0.30	292	50.71%	0.735
6	20	100	0.35	347	55.68%	0.805
7	20	100	0.40	407	27.17%	0.944
8	20	100	0.45	428	37.32%	1.010
9	20	100	0.50	367	31.48%	0.934
10	20	100	0.55	114	134.45%	0.442

Tabela 5.7: Resultados com penalização -10 para atuação fora da área de aptidão nas instâncias complexas, na qual ρ indica a densidade de preferências e C_m o coeficiente médio da satisfação.

Novamente, os valores do gap foram altos. Em um caso desses fica evidente que utilizar um limite de 15 minutos é insuficiente para obter resultados satisfatórios. Vale ressaltar que a penalidade atribuída a função objetiva não se aplica no cálculo da satisfação individual (disciplinas fora da área de atuação contam como peso 0), de forma que mesmo com um valor encontrado menor, a satisfação media foi de 0.750, sendo superior à encontrada no teste anterior.

5.4 Análise em situação real

Por fim, foram utilizados os dados de professores, turmas e preferências para a distribuição de 2022. A instância conta com um total de 61 professores e 224 turmas, com 14 áreas de atuação definidas. O algoritmo foi executado considerando uma penalidade de -10 pontos para professores atuando fora de suas áreas. Não foi definido um tempo limite para este caso.

O programa foi capaz de encontrar uma solução ótima para esta entrada em 58,51 segundos, atingindo um valor total de 1981 pontos na função objetivo, com uma satisfação média de 1.386. Este valor de satisfação desconsidera da equação professores sem turmas atribuídas ou que não registraram preferências na lista.

Comparando com a atribuição realizada no mesmo ano, usando o algoritmo de emparelhamentos, isso representa um aumento de 275 pontos, ou seja, um aumento de aproximadamente 16.12% em relação ao valor original de 1706 pontos obtido pelo resultado final em 2022. Vale lembrar que a distribuição obtida pelo algoritmo de emparelhamentos não considera uma penalidade para professores que atuem fora de suas áreas, podendo este ganho na função objetivo obtida pelo algoritmo exato ser ainda maior caso esse fator seja levado em consideração. Além disso, é importante notar que os resultados da distribuição de 2022 obtidos pelo algoritmo de emparelhamentos foram concluídos após várias tentativas de se resolver o problema, por meio da intervenção dos responsáveis pela distribuição, fixando-se manualmente algumas atribuições de disciplinas. Por outro lado, a resolução do DPD por PLI foi obtida em menos de um minuto, sem nenhuma intervenção humana.

Capítulo 6

Contribuições e Conclusões

A partir dos estudos levantados e dos resultados analisados, observa-se que a utilização da programação linear inteira é uma abordagem viável para a resolução do problema, principalmente pela flexibilidade na adição e remoção de restrições. Destaca-se que na FACOM a distribuição é realizada apenas uma vez ao ano, e que portanto é viável dedicar maior poder computacional e tempo de processamento à tarefa, já que a obtenção de resultados ótimos, mesmo que para um problema NP-difícil, é de interesse geral.

Não foi possível desenvolver, em tempo hábil para elaboração deste texto, uma análise sobre outros dados reais da FACOM. Estes resultados seriam úteis para permitir comparação com distribuições anteriores.

6.1 Trabalhos futuros

Apesar de certo polimento neste trabalho, existe ainda uma série de melhorias que podem ser feitas em cima do que foi desenvolvido no campo de funcionalidades ou de levantamento de possíveis novas restrições, além do desenvolvimento de uma medida de satisfação que melhor expresse o sentimento real dos envolvidos. Um exemplo de novas restrições ainda não implementadas seria a adição de um atributo que defina o turno de uma turma e a possibilidade de restringir os professores a atuarem apenas em determinados turnos. Além disso, o trabalho visou o desenvolvimento de um sistema capaz de resolver o problema para a FACOM, mas poderia ser estendido para outras unidades dentro da universidade.

Além disso, poderia ser adicionada uma funcionalidade que permita identificar quais professores foram prejudicados em distribuições anteriores. Com isto seria possível aplicar uma vantagem a eles ao multiplicar, durante a nova distribuição, os pesos de suas preferências por um valor maior que 1.

No campo da otimização se destaca também o desenvolvimento de heurísticas, que podem acelerar extremamente a resolução do problema, o que se torna necessário de acordo com o tamanho do corpo docente e da quantidade de turmas e preferências analisadas, além de possíveis novas restrições.

Um programa exato resolvido por PLI para o problema do horário já existe na Facom e é usado todo semestre. Uma possível extensão seria a junção dos dois programas, uma vez que ambos compartilham os mesmos dados.

Ademais, seria interessante desenvolver um software para lidar com os dados de entrada do DPD. O algoritmo de emparelhamento utilizado atualmente possui uma interface amigável. Uma possível extensão seria a integração desta interface ao trabalho desenvolvido. Isto também seria útil para os docentes, podendo oferecer funções que facilitem a definição de preferências, como por exemplo se basear nas escolhas de anos anteriores.

Referências Bibliográficas

Al-Yakoob, S. e Sherali, H. (2013). A column generation mathematical programming approach for a class-faculty assignment problem with preferences. *Computational Management Science*, 12. Citado na página 6.

Al-Yakoob, S. M. e Sherali, H. D. (2006). Mathematical programming models and algorithms for a class–faculty assignment problem. *European Journal of Operational Research*, 173(2):488–507. ISSN 0377-2217. URL https://www.sciencedirect.com/science/article/pii/S0377221705002067. Citado nas páginas 6 e 7.

Avella, P. e Vasilyev, I. (2005). A computational study of a cutting plane algorithm for university course timetabling. *J. Scheduling*, 8:497–514. Citado na página 6.

Carter, M. W. e Laporte, G. (1998). Recent developments in practical course timetabling. In Burke, E. e Carter, M., editors, *Practice and Theory of Automated Timetabling II*, p. 3–19, Berlin, Heidelberg. Springer Berlin Heidelberg. ISBN 978-3-540-49803-2. Citado na página 6.

Domenech, B. e Lusa, A. (2016). A milp model for the teacher assignment problem considering teachers' preferences. *European Journal of Operational Research*, 249(3):1153–1160. ISSN 0377-2217. URL https://www.sciencedirect.com/science/article/pii/S0377221715008139. Citado na página 6.

Gamrath, G., Anderson, D., Bestuzheva, K., Chen, W.-K., Eifler, L., Gasse, M., Gemander, P., Gleixner, A., Gottwald, L., Halbig, K., Hendel, G., Hojny, C., Koch, T., Le Bodic, P., Maher, S. J., Matter, F., Miltenberger, M., Mühmer, E., Müller, B., Pfetsch, M. E., Schlösser, F., Serrano, F., Shinano, Y., Tawfik, C., Vigerske, S., Wegscheider, F., Weninger, D., e Witzig, J. (2020). The SCIP Optimization Suite 7.0. Technical report, Optimization

Online. URL http://www.optimization-online.org/DB_HTML/2020/03/7705.html. Citado na página 17.

Gunawan, A., Ng, K., e Ong, H. (2008). A genetic algorithm for the teacher assignment problem for a university in indonesia. *International Journal of Information and Management Sciences*, 19. Citado na página 7.

Hosny, M. (2012). A heuristic algorithm for the faculty assignment problem. Citado na página 6.

Szwarc, E., Bach-Dąbrowska, I., e Bocewicz, G. (2018). *Competence Management in Teacher Assignment Planning: 24th International Conference, ICIST 2018, Vilnius, Lithuania, October 4–6, 2018, Proceedings*, p. 449–460. ISBN 978-3-319-99971-5. Citado na página 6.

Szwarc, E., Bocewicz, G., Banaszak, Z., e Wikarek, J. (2019). Competence allocation planning robust to unexpected staff absenteeism. *Eksploatacja i Niezawodnosc - Maintenance and Reliability*, 21:440–450. Citado na página 7.

Vanderbei, R. J. (2014). *Linear Programming: Foundations and Extensions*. Springer. Citado na página 9.

Wolsey, L. A. (2020). *Integer Programming*. John Wiley Sons, Ltd. Citado na página 9.