XXVIII Congresso {virtual} de Iniciação Científica da Unicamp 2020

IMECC - DMA - UNICAMP

Resumo de pesquisa de Iniciação Científica - Bolsa PIBIC/CNPq - 2019/2020

Um estudo sobre o problema de alocação

Orientanda: Beatriz Aria Valladão
Orientadora: Dra. Kelly Cristina Poldi

1 Descrição do problema

O Problema de Alocação (Burkard *et al.* [2]) constitui-se em alocar uma determinada tarefa a um certo agente obedecendo condições, analisando a melhor combinação dessas designações. Por exemplo, um problema com m tarefas e n agentes pode ser representado pelo modelo matemático

$$\begin{array}{lll} \text{minimizar} & f(x) = \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} & \\ \text{sujeito a:} & \displaystyle\sum_{j=1}^{n} x_{ij} = 1 & i = 1, \dots, m, \\ & \displaystyle\sum_{i=1}^{m} x_{ij} \leq 1 & j = 1, \dots, n, \\ & x_{ij} \in \{0, 1\} & i = 1, \dots, m \quad j = 1, \dots, n. \end{array} \tag{1}$$

em que c_{ij} é o custo da tarefa i ser realizada pelo agente j e x_{ij} indica se o agente j foi designado para a tarefa i ($x_{ij} = 1$) ou não ($x_{ij} = 0$). Nesse exemplo, assumimos que toda tarefa deve ser alocada por um único agente (primeiro conjunto de restrições) e que um agente pode ser designado para no máximo uma tarefa (segundo conjunto de restrições).

O projeto visa estudar a alocação de turmas nos principais prédios de salas de aula da UNICAMP, o Ciclo Básico I (CB) e o Ciclo Básico II (PB). O cronograma de referência foi o da semana letiva dos dias **19 a 23 de agosto de 2019**, e foram retirados do site do Ciclo Básico da UNICAMP [1].

Definimos como entrada do modelo os dados das turmas (tarefas): código da disciplina, dias da semana que são oferecidas as aulas, respectivos horários, duração das aulas e a quantidade de matriculados na turma; e os dados das salas de aula (agentes): número de identificação da sala e sua capacidade. Então, vamos estudar um modelo que visa buscar a sala mais adequada (de acordo com as condições desejadas) para cada uma das turmas, e compararemos essas alocações com as utilizadas originalmente.

2 Modelos desenvolvidos

Em primeira análise, consideramos alocar as turmas dos prédios CB, PB e também os dias da semana, separadamente. Ademais, separamos os **modelos 1**, **2** e **3** em duas modelagens, uma considerando que cada turma ocupa um espaço fixo de 2 horas/aula na grade e outra já a *segunda versão* trabalha com os horários e durações reais de cada turma.

Em seguida, foram criados os **modelos 4** e **5** onde consideramos alocar as salas do CB e do PB conjuntamente, utilizando os horários efetivos das aulas. Como foco, para disciplinas com mais de uma aula durante a semana, queremos alocá-las, se não na mesma sala, em salas próximas.

2.1 Modelos 1,2 e 3

Os modelos da **primeira versão** ($modelo.1$) não serão exibidos neste resumo devido à sua simplicidade.

Na **segunda versão** ($modelo.2$), temos m turmas para alocar, n salas disponíveis e p horários de aula. Sejam c_j a capacidade da sala j, t_i a quantidade de alunos na turma i, h_i o horário da aula da turma i e d_i a duração da aula da turma i, queremos determinar

$$x_{ijk} = \begin{cases} 1 \text{ se a turma } i \text{ foi alocada na sala } j \text{ no horário } k \\ 0 \text{ caso contrário} \end{cases}$$

com $i = 1, \dots, m$, $j = 1, \dots, n$ e $k = 1, \dots, p$.

Definimos a expressão $\text{uso}_{ij} = \displaystyle\sum_{k=h_i}^{h_i+d_i-1} x_{ijk}$, que representa a quantidade de horários da sala j que estão ocupados durante o período de aula da turma i.

Com isso, construímos um conjunto de restrições comum aos modelos da segunda versão, que

- não permitem que uma turma i seja alocada em uma sala j com capacidade menor que a quantidade t_i de alunos na turma:
$$x_{ijk} \cdot t_i \leq c_j, \qquad \text{para todo } j = 1, \dots, n \text{ e } k = 1, \dots, p;$$
- garantem que se a turma i está alocada na sala j, todos os horários necessários $(h_i, \dots, h_i + d_i - 1)$ devem ser alocados também:
$$\text{uso}_{ij} = x_{ijh_i} \cdot d_i, \qquad \text{para todo } i = 1, \dots, m \text{ e } j = 1, \dots, n;$$
- permite alocar no máximo uma turma i em cada sala j e em cada horário k:
$$\sum_{i=1}^{m} x_{ijk} \leq 1, \qquad \text{para todo } j = 1, \dots, n \text{ e } k = 1, \dots, p;$$
- garantem que para cada turma i, todos os horários necessários serão alocados para alguma sala j:
$$\sum_{j=1}^{n} \text{uso}_{ij} = d_i, \qquad \text{para todo } i = 1, \dots, m;$$
- as variáveis de decisão x_{ijk} serão binárias:
$$x_{ijk} \in \{0, 1\}, \qquad \text{para todo } i = 1, \dots, m,\ j = 1, \dots, n \text{ e } k = 1, \dots, p.$$

Modelo 1.2

Nesse modelo, buscamos encontrar uma solução que satisfaça somente as restrições básicas de alocação de sala, assim sendo, consideramos a função objetivo como constante, z_{12} (2) para o modelo da segunda versão.

$$\begin{array}{ll} \text{minimizar} & z_{12} \\ \text{sujeito a:} & [\text{conjunto básico de restrições da segunda versão}] \end{array} \tag{2}$$

Modelo 2.2

Já nesse modelo, a função objetivo minimiza a quantidade de salas utilizadas durante o período analisado, ou seja, $\sum_{j=1}^{n} u_j$, onde $u_j = \begin{cases} 1 & \text{se } \sum_{i=1}^{n} x_{ijh_i} > 0, \\ 0 & \text{caso contrário,} \end{cases}$ indica se a sala foi usada ou não. Com z_{22} sendo a função objetivo para o modelo da segunda versão (3).

$$\begin{array}{lll} \text{minimizar} & z_{22} = \sum_{j=1}^{n} u_j & \\ \text{sujeito a:} & [\text{conjunto básico de restrições da segunda versão}] & \\ & \sum_{i=1}^{m} \sum_{k=1}^{p} x_{ijk} \leq p \cdot u_j, & j = 1, \dots, n, \\ & u_j \in \{0, 1\}, & j = 1, \dots, n. \end{array} \tag{3}$$

Modelo 3.2

Utilizando ainda as duas versões formuladas, outra modelagem proposta tenta alocar turmas em salas cujo tamanho seja o mais consistente possível com o tamanho da turma.

Seja $f_{ij} = \begin{cases} c_j - t_i & \text{se } t_i \leq c_j \\ M & \text{caso contrário} \end{cases}$ a quantidade de lugares livres (se houver) na sala j se a turma i for alocada nela, onde M é um número grande.

$$\begin{array}{ll} \text{minimizar} & z_{32} = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} \cdot x_{ijh_i} \\ \text{sujeito a:} & [\text{conjunto básico de restrições da segunda versão}] \end{array} \tag{4}$$

2.2 Modelos 4 e 5

Definimos como **"disciplinas satisfeitas"** aquelas as quais todas as turmas são designadas para a mesma sala (caso haja mais de uma aula durante a semana). Dessa forma, buscamos satisfazer a maior quantidade de disciplinas possível, simulando uma alocação mais verossímia com a implementada.

Temos m turmas para alocar, n salas disponíveis, p horários diferentes para as aulas e q dias da semana. Além disso, sejam c_j a capacidade da sala j, t_i a quantidade de alunos na turma i, h_{il} o horário da aula da turma i no dia l, e d_{il} a duração da aula da turma i no dia l.

Além disso, na modelagem, as 18 salas de aula do Ciclo Básico I (CB) são as salas $1, \dots, 18$, e as 18 salas do Ciclo Básico II (PB) são as salas $19, \dots, 36$.

Para fins da modelagem, fizemos um pré-processamento dos dados e criamos um vetor indicador ind_k. Se a turma i **não** tem aula no dia l, atribuímos $h_{il} = p+1$ (indicando que é um horário inválido) e $d_{il} = 1$ (para que as restrições sejam condizentes). Com isso, $\text{ind}_k = \begin{cases} 0 & \text{se } k = p+1 \\ 1 & \text{caso contrário} \end{cases}$.

Definimos também uma expressão $\text{uso}_{ijl} = \sum_{k=h_{il}}^{h_{il}+d_{il}-1} x_{ijkl}$, que representa a quantidade de horários da sala j que estão ocupados durante o período de aula da turma i.

Queremos então determinar $x_{ijkl} = \begin{cases} 1 & \text{se a turma } i \text{ foi alocada na sala } j \text{ no horário } k \text{ do dia } l \\ 0 & \text{caso contrário} \end{cases}$, com $i = 1, \dots, m$, $j = 1, \dots, n$, $k = 1, \dots, p+1$ e $l = 1, \dots, q$.

Além disso, há restrições comuns aos dois modelos, que

- não permitem que uma turma i seja alocada em uma sala j com capacidade menor que a quantidade t_i de alunos na turma:

$$x_{i,j,k,l} \cdot t_i \leq c_j, \qquad \text{para todo } i = 1, \dots, m,\ j = 1, \dots, n,\ k = 1, \dots, p \text{ e } l = 1, \dots, q;$$

- permite alocar no máximo uma turma i para cada sala, horário e dia j, k e l, respectivamente:

$$\sum_{i=1}^{m} x_{i,j,k,l} \leq 1, \qquad \text{para todo } j = 1, \ldots, n,\ k = 1, \ldots, p \text{ e } l = 1, \ldots, q;$$

- garantem que se a turma i está alocada na sala j, todos os horários necessários $(h_i, \ldots, h_i + d_i - 1)$ devem ser alocados também:

$$\text{uso}_{ijl} = x_{i,j,h_{il},l} \cdot d_{il}, \qquad \text{para todo } i = 1, \ldots, m,\ j = 1, \ldots, n \text{ e } l = 1, \ldots, q;$$

- garantem que cada turma i no dia l será alocada (se exigido) em alguma sala j:

$$\sum_{j=1}^{n} x_{i,j,h_{il},l} = \text{ind}_{h_{il}}, \qquad \text{para todo } i = 1, \ldots, m \text{ e } l = 1, \ldots, q;$$

- as variáveis de decisão x_{ijkl} serão binárias:

$$x_{ijkl} \in \{0, 1\}, \qquad \text{para todo } i = 1, \ldots, m,\ j = 1, \ldots, n,\ k = 1, \ldots, p \text{ e } l = 1, \ldots, q.$$

$$\text{Seja dist}_{j_1 j_2} = \begin{cases} -1 \ \text{ se } \ j_1 = j_2 \\ 0 \ \text{ se } \ j_1 \leq 18 \ \textbf{ e } \ j_2 \leq 18 \ \textbf{ e } \ j_1\%2 = 0 \ \textbf{ e } \ j_2\%2 = 0 \ \textbf{ e } \ |j_1 - j_2| \leq 6 \\ 0 \ \text{ se } \ j_1 \leq 18 \ \textbf{ e } \ j_2 \leq 18 \ \textbf{ e } \ j_1\%2 \neq 0 \ \textbf{ e } \ j_2\%2 \neq 0 \ \textbf{ e } \ |j_1 - j_2| \leq 6 \\ 0 \ \text{ se } \ j_1 \geq 18 \ \textbf{ e } \ j_2 \geq 18 \ \textbf{ e } \ |j_1 - j_2| \leq 4 \\ +1 \ \text{ caso contrário} \end{cases}$$

uma medida de distância (peso) entre duas salas quaisquer j_1 e j_2. Se forem iguais, a distância é -1; se forem no CB, do mesmo lado do prédio (ambas no lado par ou no lado ímpar) e a distância real entre as salas for menor ou igual a 3 salas, a distância dada é 0; caso forem no PB, e a distância real for menor ou igual a 4, a distância dada será 0. Caso contrário, considera-se que as salas não são suficientemente próximas, e portanto é dado um peso de distância maior, igual a $+1$.

Modelo 4

$$\begin{aligned} \min \quad & z_4 = \sum_{i=1}^{m} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \sum_{l_1=1}^{q-1} \sum_{l_2=l_1+1}^{q} \text{dist}_{j_1 j_2} \cdot x_{i,j_1,h_{il_1},l_1} \cdot x_{i,j_2,h_{il_2},l_2} \\ \text{s.a.:} \quad & [\text{conjunto básico de restrições}] \end{aligned} \tag{5}$$

Modelo 5

Nessa formulação, ponderamos também a quantidade de salas utilizadas durante a semana. Seja $u_{jl} = \begin{cases} 1 \ \text{ se } \sum_{i=1}^{n} x_{i,j,h_{il},l} > 0 \\ 0 \ \text{ caso contrário} \end{cases}$ a variável que indica se a sala j é utilizada no dia l ou não, e ρ um peso (importância) para esse objetivo.

$$\begin{aligned} \min \quad & z_5 = \sum_{i=1}^{m} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \sum_{l_1=1}^{q-1} \sum_{l_2=l_1+1}^{q} \text{dist}_{j_1 j_2} \cdot x_{i,j_1,h_{il_1},l_1} \cdot x_{i,j_2,h_{il_2},l_2} + \rho \cdot \sum_{j=1}^{n} \sum_{l=1}^{q} u_{jl} \\ \text{s.a.:} \quad & [\text{conjunto básico de restrições}] \\ & \sum_{i=1}^{n} x_{i,j,h_{il},l} \leq p \cdot u_{jl}, \qquad j = 1, \ldots, n \text{ e } l = 1, \ldots, q, \\ & u_{jl} \in \{0, 1\}, \qquad j = 1, \ldots, n \text{ e } l = 1, \ldots, q. \end{aligned} \tag{6}$$

3 Abordagens de solução

3.1 Heurística

Utilizando outro método de solução, redefinimos a representação do problema. A grade de horários foi unificada, tornando a identificação dos horários contínua ao longo da semana (horários $1, \ldots, 15$ representam a segunda-feira, $16, \ldots, 30$ a terça-feira etc). Temos também 180 salas disponíveis na semana (36 salas para cada um dos 5) e 160 disciplinas com mais de uma aula na semana (de um total de 330 disciplinas). Além disso, as turmas de disciplinas são representadas separadamente (por exemplo, turma MA311-A da segunda-feira e turma MA311-A da quarta-feira, são a mesma disciplina porém tratadas como duas turmas diferentes).

Gulosa (G)

Uma heurística gulosa constrói uma solução escolhendo a cada passo a opção localmente ótima, buscando chegar a um ótimo global ao final do processo.

O primeiro passo é ordenar as turmas e as salas de aula por ordem decrescente de capacidade. Então, de maneira mecânica, alocamos à cada turma a maior sala disponível, até todas as turmas estarem designadas. Quando uma turma de identificação "I" é alocada em uma sala j, tenta-se alocar toda disciplina de mesma identificação (e consequentemente mesmo número de matriculados) na mesma sala.

Essa ordenação garante que as turmas serão alocadas em salas que comportem seu tamanho. Caso isso não seja possível, a solução é dada como infactível. Além disso, ao tentar designar todas as aulas de mesma identificação na mesma sala, asseguramos que grande parte das turmas serão satisfeitas.

Aleatória ajustável (AA)

A partir do mecanismo da heurística gulosa, possibilitamos a inserção de aleatoriedade no método para provocar uma variação da solução, podendo esta não atingir a otimalidade. Para isso, foram criadas funções que avaliam uma solução com os mesmos critérios das funções objetivos dos modelos **3** (adequação do tamanho sala designada em relação à quantidade de alunos) e **4** (distância entre as salas alocadas para cada disciplina).

Escolhemos a melhor solução rodando a heurística T vezes, onde em cada iteração: ordenamos as turmas e as salas de maneira decrescente, aplicamos um embaralhamento controlado, e então buscamos uma solução para o problema com o algoritmo guloso descrito acima. Para um dado vetor de tamanho v, aplicamos s permutações aleatórias nos primeiros f elementos do vetor. Definindo a porcentagem $fr \in [0, 1]$ do vetor que poderá ser alterada e a taxa $sr \in [0, 1]$ de embaralhamento, temos que $f = \lceil v \cdot fr \rceil$ e $s = \lceil v \cdot sr \rceil$.

4 Resultados e análise

Comparando os dados originais com o **modelo 2**, há uma redução na quantidade de salas usadas (para os dois prédios). Ainda, o **modelo 3** traz uma visão alternativa do problema. Os modelos 1, 2 e 3 ajudam a observar as diferenças entre as propostas e como a quantidade de salas usadas variam.

Aplicando o **modelo 4**, satisfazemos todas as disciplinas usando 171 salas durante a semana. Se atribuirmos $\rho = 1$ ao **modelo 5**, o número de salas usadas cai para 133, e mantemos a satisfação completa das disciplinas. Explorando o parâmetro ρ, percebemos que priorizar a minimização da quantidade de salas usadas (em uma certa gama de valores) não interfere na satisfação das disciplinas. Ao aplicarmos o **heurística G**, garantimos 92 disciplinas satisfeitas ocupando 135 salas na semana. Sua construção nos garante menor variedade de salas em relação ao modelo 4, apesar da pouca quantia de disciplinas satisfeitas.

Variando os parâmetros da **heurística AA**, observa-se que algumas configurações são mais vantajosas em relação à complexidade computacional, mesmo quando a solução do problema não é ótima (porém, muitas vezes é suficientemente boa).

Referências

[1] Calendário de aulas dos prédios do ciclo básico. https://cronograma.basico.unicamp.br/calendario/predio/1/190819. Data base: 19/08/2019 a 23/08/2019.

[2] R. E. Burkard, M. Dell'Amico, and S. Martello. *Assignment problems.* Springer, 2009.