USO DE TÉCNICAS DE LINEARIZAÇÃO NA OBTENÇÃO DE MODELOS DE PROGRAMAÇÃO LINEAR INTEIRA PARA O PROBLEMA DA DIVERSIDADE MÁXIMA DE GRUPOS

Pablo L. B. Soares, Tatiane F. Figueiredo, Otávio A. G. Nogueira
Laboratório de Pesquisa & Desenvolvimento do NEMO, Universidade Federal do Ceará
Rua Felipe Santiago – Nº 411, Campus de Russas, CEP 62900-000. Russas, Ceará, Brasil
{pablo.soares, tatianefernandes}@ufc.br,
otavio.g.nogueira@gmail.com

RESUMO

O Problema da Diversidade Máxima de Grupos (PDMG) consiste em agrupar um determinado conjunto de n elementos em m grupos disjuntos, de modo que a diversidade, obtida pela soma das distâncias individuais entre cada par de elementos, em cada grupo seja a maior possível. O problema é NP-Difícil e por isso diversas pesquisas têm focado em encontrar soluções através de heurísticas e meta-heurísticas. Na direção oposta das soluções aproximadas, outras pesquisas têm focado na construção de modelos matemáticos para o PDMG. Neste trabalho, dois novos modelos matemáticos de programação linear inteira, para o PDMG, são obtidos através da aplicação de técnicas de linearização de funções quadráticas. Os modelos gerados são comparados, por meio de experimento computacional, em 240 instâncias da literatura, com outro modelo também linearizado e inteiro da literatura. Os resultados mostram que um dos novos modelos obtido alcançou os melhores valores em termos de tempo computacional e GAP.

PALAVRAS CHAVE. Problema da diversidade máxima de grupos, Linearização, Programação linear inteira.

Programação Matemática (PM), Teoria e Algoritmos em Grafos (TAG)

ABSTRACT

The Maximally Diversity Grouping Problem (MDGP) consists of grouping a given set of n elements into m disjoint groups, so that the diversity, obtained by the sum of the individual distances between each pair of elements, in each group is the largest possible. The problem is NP-Hard and therefore several researches focused on solving MDGP using heuristics and meta-heuristics. In the opposite direction of approximate solutions, other research has been focused on the construction of mathematical models for the MDGP. In this work, two new mathematical models of integer linear programming, for MDGP, are obtained through the application of linearization techniques of quadratic functions. The generated models are compared, through a computational experiment, in 240 instances from the literature, with another model also linearized and integer from the literature. The results show that one of the new models obtained reached the best values in terms of computational time and GAP.

KEYWORDS. Maximally diverse grouping problem. Linearization. Integer linear programming.

Mathematical Programming (MP), Graphs algorithms and theory (GAT)

1. Introdução

De acordo com Araujo e Figueiredo [2018], o Problema da Diversidade Máxima de Grupos (PDMG) consiste em agrupar um determinado conjunto de n elementos em m grupos disjuntos, de modo que a diversidade em cada grupo seja a maior possível. O conceito de diversidade é definido pela soma das distâncias entre cada par de elementos pertencentes a um mesmo grupo. O objetivo do PDMG é maximizar a diversidade geral, definida como o somatório do resultado da diversidade dos m grupos.

Por ser tratar de um problema NP-Difícil [Feo e Khellaf, 1990], diversas pesquisas tem focado em encontrar soluções para o problema através de heurísticas e meta-heurísticas. Busca locais e de vizinhança são apresentadas por Lai e Hao [2016]; Brimberg et al. [2015]. Enquanto meta-heurísticas baseadas em população, como algoritmos genéticos [Singh e Sundar, 2019; Fan et al., 2011], colônia de formigas [Rodriguez et al., 2013] e meta-heurísticas baseadas em memória adaptativa como a Tabu Search [Gallego et al., 2013] também são utilizadas para resolução do PDMG. Versões híbridas que mesclam estas e outras abordagens, também podem ser encontradas em Ramos-Figueroa et al. [2020]; Yang et al. [2022]. Na contramão das abordagens aproximativas, heurísticas e meta-heurísticas, destacam-se as pesquisa de Fan et al. [2011], Araujo e Figueiredo [2018] e Schulz [2021] que são focadas na construção de modelos matemáticos para o PDMG. A seguir é apresentada a definição formal do PDMG, assim como o modelo matemático proposto por Fan et al. [2011] para sua resolução.

Dado um grafo não direcionado ponderado $G = (V, E, c)$, com $|V| = n$ vértices e $|E| = q$ arestas ponderadas, onde $c : ij \to \mathbb{R} : (i, j) \in E, i < j$ é uma função que representa o peso de cada aresta $(i, j) \in E$, tal que $c_{ij} = 0$ para todo $ij \notin E$. Considere V o conjunto de elementos e E o conjunto de arestas que representa a diversidade de cada par de elementos do conjunto, onde m é uma constante inteira positiva que indica a quantidade de grupos que devem ser formados, sendo a_g e b_g, respectivamente, os valores mínimos e máximos de elementos permitidos em cada grupo g, tal que $a_g \leq b_g$ para todo $g = \{1, \dots, m\}$. Assim, o PDMG consiste no problema de particionar os vértices de V em m grupos de forma a maximizar a soma das distâncias entre pares vértices pertencentes a cada um dos m grupos gerados.

Seja x_{ig} o conjunto de variáveis binárias que indicam se cada vértice $i \in V$ foi alocado ao grupo g ($x_{ig} = 1$), ou não ($x_{ig} = 0$), para cada grupo $i = \{1, \dots, m\}$, o Modelo Quadrático Binário (MQB) proposto por Fan et al. [2011] é definido como:

$$
\begin{aligned}
\max \quad & \sum_{g=1}^{m}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c_{ij}x_{ig}x_{jg} & \\
s.a \quad & \sum_{g=1}^{m} x_{ig} = 1, & \forall i = \{1, \dots, n\}, \\
& \sum_{i=1}^{n} x_{ig} \leq b_g, & \forall g = \{1, \dots, m\}, \\
& \sum_{i=1}^{n} x_{ig} \geq a_g, & \forall g = \{1, \dots, m\}, \\
& x_{ig} \in \{0, 1\}, & \forall i = \{1, \dots, n\}, \forall g = \{1, \dots, m\}.
\end{aligned}
\tag{1}
$$

As restrições $\sum_{g=1}^{m} x_{ig} = 1$ garantem que cada vértice i deve estar em exatamente um grupo g, enquanto as demais restrições definem os limites de tamanho máximo e mínimo de cada grupo, além das restrições que fixam o domínio correto das variáveis. Por fim, a função objetivo define quais arestas devem ser contabilizadas, ou seja, quais pares de elementos participam de cada grupo. Note que a função objetivo deste modelo é quadrática, que por sua vez, pode ser reescrita utilizando técnicas de linearização, gerando assim modelos de Programação Linear Inteira (PLI).

O uso de técnicas de linearização consiste em substituir os termos quadráticos da função objetivo ou das restrições do modelo por termos lineares a custo do acréscimo de variáveis e/ou novas restrições ao modelo [Fortet, 1960; Beasley, 1998; Soares, 2018]. De acordo com Furini e Traversi [2019] a principal vantagem de descrever um MQB como um modelo PLI é o fato de poder utilizar diferentes técnicas para sua resolução e diferentes ferramentas computacionais, tais como solver matemáticos.

Desta forma, este trabalho compara o modelo de PLI desenvolvido por Araujo e Figueiredo [2018], sem o acréscimo de desigualdades válidas, que foi obtido a partir da aplicação da linearização de Glover e Woolsey [1974] utilizando o modelo proposto por Fan et al. [2011] com dois novos modelos de PLI gerados pela aplicação das linearizações de Glover [1975] e Rodrigues [2010] também no modelo de Fan et al. [2011]. Para nosso maior conhecimento, este é o primeiro trabalho que apresenta o uso das linearizações mencionadas para resolução do PDMG. Para avaliar o desempenho e a competitividade dos três modelos em termos de qualidade e eficiência da solução, realizamos experimentos computacionais em 240 instâncias da literatura. O restante desse trabalho está organizado da seguinte forma. As Seções 2, 3 e 4 descrevem como foram obtidos os três modelos de PLI para resolução do PDMG através da aplicação das três respectivas técnicas de linearização propostas por Glover e Woolsey [1974], Glover [1975], Rodrigues [2010], respectivamente. Ao final de cada subseção é apresentado o número de variáveis e restrições que foram adicionados em cada modelo para que o mesmo fosse linearizado. Por fim, a Seção 5 é dedicada aos experimentos e resultados computacionais, e a Seção 6 são apresentadas as considerações e direções futuras serão apresentadas.

2. Linearização de Glover e Woolsey

A linearização introduzida e descrita por Glover e Woolsey [1974], também conhecida como linearização clássica, consiste em duas etapas: a primeira, substituir cada produto $x_{ig}x_{jg}$, da função objetivo, por uma variável não negativa y_{ijg}; e a segunda em adicionar as seguintes restrições: $y_{ijg} \leq x_{ig}$, $y_{ijg} \leq x_{jg}$ e $y_{ijg} \geq x_{ig} + x_{jg} - 1$ para todo $g = \{1, \ldots, m\}, 1 \leq i < j \leq n$ que garantem que $y_{ijg} = x_{ig}x_{jg}$. Dessa forma, podemos reescrever (1) como:

$$
\begin{array}{lll}
\max & \displaystyle\sum_{g=1}^{m}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} c_{ij}y_{ijg} & \\
s.a & \displaystyle\sum_{g=1}^{m} x_{ig} = 1, & \forall i = \{1, \ldots, n\}, \\
& \displaystyle\sum_{i=1}^{n} x_{ig} \leq b_g, & \forall g = \{1, \ldots, m\}, \\
& \displaystyle\sum_{i=1}^{n} x_{ig} \geq a_g, & \forall g = \{1, \ldots, m\}, \\
& y_{ijg} \leq x_{ig}, & \forall g = \{1, \ldots, m\}, 1 \leq i < j \leq n, \\
& y_{ijg} \leq x_{jg}, & \forall g = \{1, \ldots, m\}, 1 \leq i < j \leq n, \\
& y_{ijg} \geq x_{ig} + x_{jg} - 1, & \forall g = \{1, \ldots, m\}, 1 \leq i < j \leq n, \\
& y_{ijg} \geq 0, & \forall g = \{1, \ldots, m\}, 1 \leq i < j \leq n, \\
& x_{ig} \in \{0, 1\}, & \forall i = \{1, \ldots, n\}, \forall g = \{1, \ldots, m\}.
\end{array}
\tag{2}
$$

Pode-se observar que, em relação ao modelo (1), o modelo (2) acrescenta $m(n^2 - n)/2$ novas variáveis e $m(4n^2 - 4n)/2$ restrições.

3. Linearização Glover

A ideia principal da linearização apresentada por Glover [1975], consiste na introdução de um conjunto de variáveis contínuas $w_{ig} \in \mathbb{R}$ tais que $w_{ig} = \left(x_{ig} \sum_{j=i+1}^{n} c_{ij} x_{jg}\right)$ para todo $g = \{1, \ldots, m\}$ e todo $i = \{1, \ldots, n-1\}$; e também na adição de restrições chamadas de *Big-M*. Para isso, sejam $\overline{C}_{ig}$ e $\underline{C}_{ig}$ os limites superiores e inferiores, respectivamente, para $\sum_{j=i+1}^{n} c_{ij} x_{jg}$, tais que $\overline{C}_{ig} = \sum_{j=i+1}^{n} \max(0, c_{ij})$ e $\underline{C}_{ig} = \sum_{j=i+1}^{n} \min(0, c_{ij})$, para todo $g = \{1, \ldots, m\}$ e todo $i = \{1, \ldots, n-1\}$. Dessa forma, podemos reescrever (1) como:

$$
\begin{aligned}
\max \quad & \sum_{g=1}^{m} \sum_{i=1}^{n-1} w_{ig} & & \\
s.a \quad & \sum_{g=1}^{m} x_{ig} = 1, & & \forall i = \{1, \ldots, n\}, \\
& \sum_{i=1}^{n} x_{ig} \leq b_g, & & \forall g = \{1, \ldots, m\}, \\
& \sum_{i=1}^{n} x_{ig} \geq a_g, & & \forall g = \{1, \ldots, m\}, \\
& w_{ig} \leq \overline{C}_{ig} x_{ig}, & & \forall g = \{1, \ldots, m\}, \forall i = \{1, \ldots, n-1\}, \\
& w_{ig} \leq \sum_{j=i+1}^{n} c_{ij} x_{jg} - \underline{C}_{ig}(1 - x_{ig}), & & \forall g = \{1, \ldots, m\}, \forall i = \{1, \ldots, n-1\}, \\
& x_{ig} \in \{0, 1\} & & \forall i = \{1, \ldots, n\}, \forall g = \{1, \ldots, m\}.
\end{aligned}
\tag{3}
$$

Note que quando $x_{ig} = 1$, a restrição $w_{ig} \leq \overline{C}_{ig} x_{ig}$ torna-se redundante enquanto a restrição com $\underline{C}_{ig}$ leva a $w_{ig} \leq \sum_{j=i+1}^{n} c_{ij} x_{jg}$. Por outro lado, quando $x_{ig} = 0$, $w_{ig} \leq \overline{C}_{ig} x_{ig}$ estabelece que $w_{ig} \leq 0$, levando a restrição com $\underline{C}_{ig}$ a ser redundante. O modelo (3) acrescenta $m(n-1)$ variáveis e $m(2n-2)$ restrições possuindo menos variáveis que (2), entretanto sua relaxação pode ser mais fraca.

4. t-Linearização

Rodrigues [2010] propôs um *framework*, chamado de t-linearização, para linearizar o Problema Quadrático da Mochila, e que vem sendo aplicado ultimamente em outros problemas que possuem modelos quadráticos binários, tais como, Problema do Corte Máximo [Soares et al., 2017], Problema da Diversidade Máxima [Soares e Campêlo, 2021] e Problema da Clique Máxima Ponderada em Arestas [Soares e Campêlo, 2018]. Essa abordagem consiste basicamente de duas etapas, descritas a seguir: a primeira etapa, o termo quadrático $\sum_{g=1}^{m} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_{ij} x_{ig} x_{jg}$ da função objetivo do modelo (1) é substituído por uma única variável real t e logo após é adicionado como uma restrição que limita superiormente o valor de t; a segunda etapa consiste em substituir a restrição quadrática, recém adicionada, por um conjunto exponencial de restrições lineares $t \leq \sum_{g=1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{i-1} c_{\pi(j)\pi(i)} x_{\pi(i)\pi(g)}, \forall \pi \in \mathrm{S}_n$, onde S_n é o conjunto de todas as permutações de $\{1, \ldots, n\}$. Dessa forma, podemos reescrever (1) como:

$$
\begin{aligned}
\max \quad & t \\
s.a \quad & \sum_{g=1}^{m} x_{ig} = 1, && \forall i = \{1,\dots,n\}, \\
& \sum_{i=1}^{n} x_{ig} \leq b_g, && \forall g = \{1,\dots,m\}, \\
& \sum_{i=1}^{n} x_{ig} \geq a_g, && \forall g = \{1,\dots,m\}, \\
& t \leq \sum_{g=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{i-1} c_{\pi(j)\pi(i)} x_{\pi(i)\pi(g)} && \forall \pi \in \mathrm{S}_n, \\
& t \in \mathbb{R}, \\
& x_{ig} \in \{0,1\} && \forall i = \{1,\dots,n\}, \forall g = \{1,\dots,m\}.
\end{aligned} \tag{4}
$$

O modelo (4) acrescenta, em relação ao modelo (1), uma única variável, mas também acrescenta $m(2^n)$ restrições lineares que definem os mesmos pontos inteiros. Na verdade, tais restrições lineares que definem a envoltória convexa dos pontos inteiros podem ser separadas em tempo polinomial quando os coeficientes do termo quadrático são não-negativos [Rodrigues, 2010].

5. Experimentos e análise dos resultados computacionais

Na subseção 5.1 são descritas os três conjuntos de instâncias utilizadas nos experimentos computacionais. Já na subseção 5.2 sé mostrado os parâmetros utilizados na execução das linearizações, assim como a configuração do ambiente utilizado. A subseção 5.3 é dedicada aos experimentos realizados e na análise dos resultados obtidos.

5.1. Instâncias

O conjunto de dados utilizado contém 240 instâncias que estão divididos nos seguintes subconjuntos, descritos a seguir:

- ***RanReal:*** esse conjunto consiste de 80 instâncias, onde cada instância é representada por uma matriz quadrática. Cada matriz é preenchida por valores reais c_{ij} que representam as distâncias, geradas por uma distribuição uniforme $(0, 100)$, entre cada elementos i e j de cada matriz;

- ***RanInt:*** esse conjunto consiste de 80 instâncias, onde seus valores foram obtidos através de uma distribuição uniforme $(0, 100)$, tal qual as instâncias *RanReal*, porém os valores gerados são inteiros;

- ***Geo:*** esse conjunto consiste de 80 instâncias e segue a mesma estrutura e tamanho dos dois conjuntos anteriores, porém, os valores das matrizes são calculados como distâncias euclidianas entre pares de pontos que possuem coordenadas geradas aleatoriamente entre $[0, 10]$. O número de coordenadas para cada ponto é gerado aleatoriamente no intervalo de $[2, 21]$.

Estas e outras instâncias fazem parte de um conjunto de *benchmark*, denominado de MDGPLIB, e estão disponíveis em `https://grafo.etsii.urjc.es/optsicom/mdgp/`. Cada uma das instância representa um grafo ponderado e possui os seguintes parâmetros: n, que indica a quantidade de vértices no grafo; m, que indica a quantidade de grupos que serão formados; m pares (a_g, b_g) que indicam a quantidade mínima e máxima de elementos de vértices em cada grupo; e uma *flag* que indica em qual das duas categoria a instância faz parte. Em uma categoria, o tamanho dos grupos são iguais $a_g = b_g = \lfloor n/m \rfloor$ e na outra categoria, os limites mínimos e máximos de cada grupo (a_g, b_g) para cada instância são gerados de forma aleatório de acordo com o intervalo predefinido. Os valores de a_g são gerados no intervalo $[a_g^{min}, a_g^{max}]$ e os valores de b_g são gerados no intervalo $[b_g^{min}, b_g^{max}]$. A Tabela 1 apresenta os parâmetros, onde cada uma das linhas, de 3 a 6, da tabela, representa 10 instâncias formadas pela combinação dos parâmetros em cada categoria.

Tabela 1: Sumário dos parâmetros das 80 instâncias de cada subconjunto

$(a_g = b_g)$				$(a_g \leq b_g)$				
n	m	$\lfloor n/m \rfloor$	#	a_g^{min}	a_g^{max}	b_g^{min}	b_g^{max}	#
10	2	5	10	3	5	5	7	10
12	4	3	10	2	3	3	5	10
30	5	6	10	5	6	6	10	10
60	6	10	10	7	10	10	14	10
Total			40	Total				40

5.2. Configuração do ambiente e parâmetro de execução

Os três modelos de PLI gerados a partir da aplicação das lineareizações foram implementados utilizando a linguagem C++ através do ambiente de desenvolvimento *Code::Blocks*. Além disso, o software IBM/ILOG CPLEX 12.7 foi utilizado como *solver* por possuir uma biblioteca, denominada *Concert Technology*, que permite o desenvolvimento de modelos matemáticos utilizando a linguagem de programação C++. Os experimentos foram realizados em um processador Intel Core^{TM} i$5-4570$ com 3.20 GHz, 16 GB de memória RAM e sistema operacional Ubuntu 16.04 LTS. Importante mencionar que como o modelo (4) possui um número exponencial de restrições, foi implementado, em forma de *CallBack*, o algoritmo de separação descrito em Rodrigues [2010]. O procedimento é chamado pelo *Solver* sempre que uma solução inteira é encontrada, ou seja, as restrições da t-linearização são usadas como *lazy constraints*. A seguir, os parâmetros do solver *Cplex* que foram utilizados: *Threads*, que informa a quantidade de *threads* utilizada na execução do *solver*, setado em 1 em todas as execuções; *TiLim*, que informa a quantidade máxima, em segundos, de tempo de execução, setado em 1800 segundos em todas as execuções.

5.3. Experimento computacionais e análise dos resultados obtidos

Nas Tabelas 2 e 3 apresentadas a seguir, estão presentes as colunas relacionadas ao número de vértices (n), à identificação de cada *Instância*, o valor ótimo obtido, o *Tempo* e execução (segundos), assim como o número de *Nós* gerados por cada modelo (2), (3) e (4), respectivamente. O número de *Nós* é um parâmetro retornado pelo próprio solver *Cplex*, que refere ao número de nós do *branch-and-bound* que foram explorados durante a resolução da instância. As Tabelas 4 e 5 apresentam os casos onde nenhum dos modelos propostos obteve solução ótima no tempo máximo fixado. Por esse motivo, a coluna *Melhor Valor*, apresenta o limite inferior obtido por cada modelo, seguido da coluna *Gap*, parâmetro também retornado pelo solver *Cplex*, e por fim novamente a coluna *Nós*.

Ao analisar os resultados da Tabela 2 e 3 é importante notar que há uma notável diferença entre o tempo de execução do modelo (3) em comparação com os outros dois modelos propostos. O tempo de execução do modelo (3) é consideravelmente inferior aos demais modelos em quase todas instâncias apresentadas nestas tabelas, com exceção das instâncias: Int_01, Real_09 com $a_g \leq b_g$ e Int_09 e Real_09 com $a_g = b_g$, todas com 10 vértices. Porém, note que as 4 instâncias descritas podem ser consideradas de fácil resolução, pois todos os modelos utilizaram apenas alguns décimos de segundos para sua resolução, portanto as mesmas podem ser desconsideradas da análise, sem prejuízo algum. A qualidade do modelo (3) pode ser evidenciado também visualizando os grupos de instâncias de difícil resolução, como é o caso das instâncias do tipo *Real* com 10 e 12 vértices e visualizando a média do tempo de execução de todas as instâncias apresentadas no final das tabelas. Em relação ao número de nós gerados, os resultados seguem de forma similar ao tempo de execução gasto pelos modelos.

Por fim, ao analisar os resultados da Tabela 4 e 5, novamente o modelo (3) se sobressai em comparação com os outros dois modelos propostos. Com exceção das instâncias Int_02, Int_03 com $a_g \leq b_g$ e 30 vértices, o modelo (3) apresentou os melhor limites inferiores para todas as demais instâncias apresentadas nas tabelas. Em relação aos *Gaps* gerados, os resultados seguem de forma similar ao limite superior obtido pelos modelos. Note que até mesmo para instâncias Int_02, Int_03 mencionadas, o modelo (3) obtém os melhores *Gaps*, isso se deve ao fato do solver *Cplex* considerar tanto o limite inferior quanto o limite superior para o cálculo deste parâmetro, ficando evidente pela coluna *Nós* que o modelo (3) também é o que conseguem construir uma maior ramificação na sua árvore de soluções no tempo limite estabelecido. É importante ressaltar que os resultados obtidos pelo modelo (1) foram suprimidos das tabelas por um questão espacial, não afetando a análise de resultados descrita, ou seja, o modelo (3) também apresenta um resultado superior quando comparado ao modelo (1).

6. Conclusão e Direções Futuras

Neste trabalho foram gerados dois novos modelos de PLI para o Problema da Diversidade Máxima de Grupos, obtidos através do uso das técnicas de linearização já difundidas. O desempenho dos modelos foi analisado e comparado com o único modelo linearizado apresentado na literatura para o problema. Para realização dos experimentos computacionais, foi utilizada uma amostra de 240 instâncias pertencentes a um *benchmark* com tamanhos variando entre 10 – 60 vértices, com tamanhos de grupos iguais e diferentes. Embora a execução dos modelos implementados tenha sido limitada pelo tempo, é possível observar que aumentar o tempo limite de execução não resultaria, de forma significativa, em melhores soluções para instâncias de tamanhos 30 e 60 vértices.

Na comparação dos resultados, constatamos que o modelo (3) se sobressai em comparação com os outros dois modelos propostos com exceção de 6 instâncias de fácil resolução, que podem ser desconsideradas, sem prejuízo algum da análise realizada. Em trabalho futuros espera-se melhorar ainda mais a qualidade do modelo (3), através do uso de desigualdades válidas, com o objetivo de obter soluções ótima para as instâncias com tamanhos acima de 30 vértices ou investigar o uso de outras linearizações. Além disso, há também a possibilidade de criação de uma novas heurísticas que possam gerar bons limites inferiores e auxiliando na diminuição dos ramos gerados pelo *brand-and-bound*.

Referências

Araujo, C. e Figueiredo, T. (2018). O problema da diversidade máxima de grupos: uma abordagem de programação linear inteira. *Anais do XL SBPO*.

Beasley, J. E. (1998). Heuristic algorithms for the unconstrained binary quadratic programming problem. Technical report, Working Paper, The Management School, Imperial College, London, England.

Brimberg, J., Mladenović, N., e Urošević, D. (2015). Solving the maximally diverse grouping problem by skewed general variable neighborhood search. *Information Sciences*, 295:650–675.

Fan, Z.-P., Chen, Y., Ma, J., e Zeng, S. (2011). Erratum: A hybrid genetic algorithmic approach to the maximally diverse grouping problem. *Journal of the Operational Research Society*, 62(7): 1423–1430.

Feo, T. A. e Khellaf, M. (1990). A class of bounded approximation algorithms for graph partitioning. *Networks*, 20(2):181–195.

Fortet, R. (1960). Applications de l'algebre de boole en recherche opérationelle. *Revue Française de Recherche Opérationelle*, 4(14):17–26.

Furini, F. e Traversi, E. (2019). Theoretical and computational study of several linearisation techniques for binary quadratic problems. *Annals of Operations Research*, 279(1):387–411.

Gallego, M., Laguna, M., Martí, R., e Duarte, A. (2013). Tabu search with strategic oscillation for the maximally diverse grouping problem. *Journal of the Operational Research Society*, 64(5): 724–734.

Glover, F. (1975). Improved linear integer programming formulations of nonlinear integer problems. *Management science*, 22(4):455–460.

Glover, F. e Woolsey, E. (1974). Converting the 0-1 polynomial programming problem to a 0-1 linear program. *Operations research*, 22(1):180–182.

Lai, X. e Hao, J.-K. (2016). Iterated maxima search for the maximally diverse grouping problem. *European Journal of Operational Research*, 254(3):780–800.

Ramos-Figueroa, O., Quiroz-Castellanos, M., Mezura-Montes, E., e Schütze, O. (2020). Metaheuristics to solve grouping problems: A review and a case study. *Swarm and Evolutionary Computation*, 53:100643.

Rodrigues, C. D. (2010). *Abordagens híbridas na solução de problemas de programação inteira da teoria e prática*. PhD thesis, Universidade Federal do Ceará, Fortaleza, CE.

Rodriguez, F. J., Lozano, M., García-Martínez, C., e González-Barrera, J. D. (2013). An artificial bee colony algorithm for the maximally diverse grouping problem. *Information Sciences*, 230: 183–196.

Schulz, A. (2021). The balanced maximally diverse grouping problem with block constraints. *European Journal of Operational Research*, 294(1):42–53.

Singh, K. e Sundar, S. (2019). A new hybrid genetic algorithm for the maximally diverse grouping problem. *International Journal of Machine Learning and Cybernetics*, 10(10):2921–2940.

Soares, P. e Campêlo (2018). t-linearizalização aplicada ao problema da clique máxima ponderada em arestas. *Anais do L SBPO*.

Soares, P. e Campêlo, M. (2021). t-linearization for the maximum diversity problem. *Optimization Letters*, 15(8):2879–2895.

Soares, P., Campêlo, M., Rodrigues, C. D., e Michelon, P. (2017). t-linearizalização de funções quadráticas de variáveis binárias. *Anais do XLIX SBPO*, p. 2569–2580.

Soares, P. L. B. (2018). *Problemas quadráticos binários: abordagem teórica e computacional*. PhD thesis, Universidade Federal do Ceará, Fortaleza, CE.

Yang, X., Cai, Z., Jin, T., Tang, Z., e Gao, S. (2022). A three-phase search approach with dynamic population size for solving the maximally diverse grouping problem. *European Journal of Operational Research*.

Tabela 2: Resultado das instâncias com tamanhos 10 e 12 e grupos diferentes ($a_g \leq b_g$)

n	Instância	Valor Ótimo	Tempo (s)			Nós		
			Modelo (2)	Modelo (3)	Modelo (4)	Modelo (2)	Modelo (3)	Modelo (4)
10	Geo_01	3864.69	0.35	0.04	0.21	414	412	797
	Geo_02	1404.8	0.18	0.05	0.13	303	372	470
	Geo_03	1662.12	0.34	0.05	0.14	223	290	466
	Geo_04	3678.98	0.23	0.06	0.23	395	434	799
	Geo_05	2452.55	0.24	0.02	0.15	151	208	743
	Geo_06	1161.77	0.23	0.03	0.15	219	379	479
	Geo_07	3206.54	0.2	0.05	0.23	357	385	772
	Geo_08	2243.15	0.26	0.03	0.17	283	348	485
	Geo_09	3828.65	0.24	0.03	0.17	412	417	493
	Geo_10	4559.18	0.27	0.03	0.12	167	198	602
12	Geo_01	807.677	134.76	23.83	203.48	193350	100294	644912
	Geo_02	921.548	116.01	8.1	182.77	168663	42411	614677
	Geo_03	2368.84	133.87	26.01	137.56	143211	135871	617824
	Geo_04	2041.37	131.56	26.91	376.12	171610	119660	999986
	Geo_05	2684.17	129.67	33.62	348.31	179694	140973	939747
	Geo_06	1612.75	170.5	57.49	273.91	258747	239612	924436
	Geo_07	1821.88	281.13	84.51	814.35	516939	423407	1714912
	Geo_08	2481.76	166.25	30.61	359.7	244811	149980	1089140
	Geo_09	1940.36	259.56	91.43	374.63	374227	386249	1077232
	Geo_10	984.986	79.48	9.69	168.89	105446	42396	526623
10	Int_01	1325	0.21	0.01	0.1	71	80	475
	Int_02	1204	0.26	0.02	0.09	72	139	363
	Int_03	1374	0.18	0.02	0.08	45	95	363
	Int_04	1308	0.17	0.01	0.06	39	88	284
	Int_05	1349	0.19	0.01	0.07	65	104	467
	Int_06	1228	0.31	0.01	0.07	59	83	298
	Int_07	1221	0.29	0.01	0.08	54	111	465
	Int_08	1208	0.21	0.01	0.06	57	66	257
	Int_09	1278	0.25	0.02	0.12	110	144	420
	Int_10	1268	0.26	0.02	0.07	50	99	415
12	Int_01	1.059	58.64	2.88	79.14	51315	7429	269255
	Int_02	978	28.93	2.73	42.17	22197	7268	159879
	Int_03	993	43.98	3.95	36.53	24300	10482	156095
	Int_04	1139	47.4	8.09	104.79	43053	23319	352569
	Int_05	1199	26.03	4.66	70.24	23275	12450	299209
	Int_06	1112	23.05	1.78	39.93	13746	3697	154434
	Int_07	1248	26.97	4.86	80.48	25542	15966	288559
	Int_08	1010	24.03	1.6	68.85	13025	4475	284804
	Int_09	1080	47.21	3.9	63.94	38130	17007	211195
	Int_10	1173.0	14.35	2.64	33.14	8053	5394	125050
10	Real_01	1437.81	0.24	0.01	0.10	97	92	471
	Real_02	1202.09	0.22	0.01	0.06	21	33	277
	Real_03	1292.79	0.37	0.01	0.07	37	49	293
	Real_04	1285.91	0.2	0.02	0.11	85	130	492
	Real_05	1574.6	0.17	0.01	0.08	59	56	409
	Real_06	1153.93	0.23	0.01	0.10	94	126	396
	Real_07	1402.68	0.34	0.01	0.07	42	42	393
	Real_08	1267.58	0.33	0.02	0.09	87	122	330
	Real_09	1420.13	0.19	0.01	0.1	84	139	407
	Real_10	1381.34	0.16	0.01	0.07	45	48	367
12	Real_01	1050.35	46.22	2.06	10.141	47033	6013	368019
	Real_02	1093.06	32.56	5.22	64.37	30684	18093	226188
	Real_03	1011.77	37.28	1.92	33.91	25598	5876	141963
	Real_04	1248.46	40.13	4.18	61.27	46853	10944	220965
	Real_05	1225.48	24.59	2.98	74.76	24161	7472	300084
	Real_06	1036.79	46.54	7.96	57.29	49794	29251	192277
	Real_07	1207.33	39.58	7.97	112.94	50732	30671	381933
	Real_08	921.411	35.3	3.7	76.36	30211	11787	290739
	Real_09	980.793	28.47	4.49	49.81	23383	19541	175296
	Real_10	1235.08	32.31	6.76	72.80	37534	21315	269675
	Média		53.3	22.7	89.4	50371.8	38519.1	234261.6

Tabela 3: Resultado das instâncias com tamanhos 10 e 12 e grupos iguais ($a_g = b_g$)

n	Instância	Valor Ótimo	Tempo (s)			Nós		
			Modelo (2)	Modelo (3)	Modelo (4)	Modelo (2)	Modelo (3)	Modelo (4)
10	Geo_01	3660.67	0.36	0.05	0.15	426	337	494
	Geo_02	1404.8	0.19	0.05	0.15	303	372	477
	Geo_03	1662.12	0.33	0.05	0.14	223	290	484
	Geo_04	3504.88	0.26	0.03	0.16	399	418	492
	Geo_05	1994.6	0.27	0.04	0.13	321	385	468
	Geo_06	1161.77	0.23	0.03	0.13	219	379	476
	Geo_07	3031.27	0.28	0.04	0.15	365	360	482
	Geo_08	2243.15	0.24	0.05	0.16	281	326	477
	Geo_09	3828.65	0.25	0.03	0.16	412	417	486
	Geo_10	3752.03	0.27	0.03	0.14	353	362	477
12	Geo_01	716.46	26.91	1.49	110	26352	7137	479061
	Geo_02	840.18	29.63	1.29	93.63	27871	5789	429107
	Geo_03	2368.84	39.59	10.65	129.31	49317	59846	582373
	Geo_04	1657.54	37.94	33.83	137.89	43673	184820	591391
	Geo_05	2120.14	41.91	23.1	137.82	50010	145718	622577
	Geo_06	1482.28	38.84	5.35	126.60	44686	28801	577005
	Geo_07	1445.27	37.07	8.37	120.03	40193	50236	522736
	Geo_08	2261.4	42.51	9.1	134.85	53617	69307	610283
	Geo_09	1734.33	36.91	63.39	138	42735	298120	597524
	Geo_10	783.734	28.64	0.9	93.51	28774	3192	412296
10	Int_01	1292	0.14	0.01	0.07	41	58	319
	Int_02	1204	0.24	0.01	0.09	72	139	356
	Int_03	1374	0.17	0.01	0.08	45	95	421
	Int_04	1201	0.25	0.02	0.07	51	93	326
	Int_05	1208	0.24	0.02	0.07	61	102	334
	Int_06	1228	0.31	0.01	0.07	59	83	330
	Int_07	1140	0.13	0.01	0.09	86	123	378
	Int_08	1208	0.28	0.01	0.06	39	67	299
	Int_09	1278	0.25	0.02	0.1	110	144	428
	Int_10	1112	0.09	0.02	0.08	74	147	376
12	Int_01	985	2.74	0.88	33.27	977	2452	148172
	Int_02	856	2.29	0.63	38.30	773	2121	165718
	Int_03	993	2.9	0.93	37.72	829	1968	165242
	Int_04	937	2.83	0.94	58.31	1102	2748	253409
	Int_05	1003	2.74	0.47	54.49	475	2577	235034
	Int_06	1000	2.83	0.4	22.59	431	2187	93627
	Int_07	993	2.03	0.6	37.29	459	3322	160429
	Int_08	863	2.58	0.31	43.63	544	1716	192222
	Int_09	989	2.57	0.82	34.14	804	1463	136630
	Int_10	923	2.39	0.56	24.56	465	2881	104640
10	Real_01	1427.85	0.19	0.01	0.08	74	51	328
	Real_02	1202.09	0.21	0.01	0.05	21	33	243
	Real_03	1292.79	0.35	0.01	0.08	37	49	307
	Real_04	1263.77	0.18	0.01	0.08	55	74	355
	Real_05	1385.65	0.16	0.01	0.08	104	73	333
	Real_06	1153.93	0.23	0.02	0.10	94	126	394
	Real_07	1236.48	0.2	0.02	0.10	102	167	392
	Real_08	1267.58	0.19	0.01	0.07	96	133	334
	Real_09	1420.13	0.19	0.01	0.1	84	139	386
	Real_10	1195.92	0.21	0.01	0.08	50	68	318
12	Real_01	956.43	3.56	0.8	52.23	1216	2451	216854
	Real_02	972.55	2.74	0.78	47.62	1241	2845	187996
	Real_03	1011.77	3.06	0.42	36.44	1204	2551	149732
	Real_04	986.692	2.95	0.99	46.94	1904	3474	196696
	Real_05	993.349	2.81	0.47	32.42	1168	2530	135134
	Real_06	937.348	2.94	0.88	36.07	1409	2406	142457
	Real_07	977.853	2.32	0.75	34.57	1163	2337	136705
	Real_08	869.72	2.57	0.7	32.57	1021	2390	133918
	Real_09	881.72	2.48	0.4	28.47	542	2618	121566
	Real_10	1031.01	3.26	0.86	38.23	1553	2092	161861
Média			7.02	2.86	33.24	7186.1	15128.4	144569.4

Tabela 4: Resultado das instâncias com tamanhos 30 e 60 e grupos diferentes ($a_b \leq b_g$)

n	Instância	Melhor Valor			Gap (%)			Nós		
		(2)	(3)	(4)	(2)	(3)	(4)	(2)	(3)	(4)
30	Geo_01	14288.3	14398.1	14139.9	76	50	78	67573	950618	71122
	Geo_02	13689	13901.6	13514.4	75	49	79	117680	921618	72444
	Geo_03	11301.7	11367.7	11129.3	74	50	78	131868	883018	75134
	Geo_04	6019.48	6077.91	5921.22	72	51	77	103297	637666	74054
	Geo_05	15026.8	15176.2	14902.2	74	48	78	86331	951369	72000
	Geo_06	14787.3	14942.4	14567.6	75	49	79	90022	1002400	74179
	Geo_07	7880.57	7979.02	7839.64	75	50	78	73333	786541	72810
	Geo_08	13233.3	13308	13151.6	76	46	79	86831	726945	71216
	Geo_09	11023.7	11188.3	10957.9	75	50	79	83523	877558	67708
	Geo_10	10823.5	10904.2	10628.5	74	51	79	85677	817068	68910
60	Geo_01	46691	48231.9	46807.0	83	63	83	3132	123918	50629
	Geo_02	44722.2	46355	45208.7	83	57	82	2985	163243	43829
	Geo_03	54193.4	56027.9	54745.3	82	63	82	4298	122326	44700
	Geo_04	47748.9	49837.4	48631.6	83	58	82	3150	150677	48241
	Geo_05	53952.5	55645.8	54550.6	82	67	82	4178	229414	44613
	Geo_06	26343.7	26966.6	26373.6	82	63	82	4525	103076	45300
	Geo_07	38773.8	40823	39111.5	83	68	83	4679	147423	44060
	Geo_08	41752.9	43291.8	42070.7	83	64	83	4588	190922	43815
	Geo_09	37875.5	39142.3	38305.8	82	62	82	4510	100835	52023
	Geo_10	53587.3	55429.3	54025.3	83	61	83	4838	132087	49002
30	Int_01	5413	5555	5113	67	45	74	75237	307818	76643
	Int_02	5781	5595	5278	66	46	74	74629	324418	72219
	Int_03	5245	5221	5198	67	48	72	88992	367228	75969
	Int_04	5400	5602	4803	66	45	75	94435	315660	77011
	Int_05	5830	5840	5506	65	45	72	73231	378118	73505
	Int_06	5303	5335	4579	66	44	75	69392	314818	68507
	Int_07	5216	5220	4848	67	47	74	98233	323159	71297
	Int_08	5061	5241	4659	68	43	75	75582	492751	73439
	Int_09	5546	5580	5008	66	46	74	104355	339118	70124
	Int_10	5471	5543	5104	67	46	74	93894	336047	79907
60	Int_01	16048	17802	16760	81	64	80	4225	41213	55415
	Int_02	16487	18027	15994	80	62	81	4540	52126	37596
	Int_03	16608	17729	15967	80	64	81	4374	43730	48817
	Int_04	15980	17749	16562	80	63	80	4872	45039	41010
	Int_05	16539	18747	17347	80	62	80	4938	47763	46604
	Int_06	16141	18080	16845	81	64	80	4401	45648	40400
	Int_07	16546	18298	16864	80	63	80	5139	45951	49973
	Int_08	16394	18458	17011	81	60	80	4854	70925	51409
	Int_09	15965	17783	16638	81	64	80	4740	51089	41471
	Int_10	16122	18059	16392	81	63	80	4746	44688	48025
30	Real_01	5248.9	5682.13	5429.46	68	43	71	88692	424659	80640
	Real_02	5508.02	5685.05	5220.71	68	44	73	87016	438357	75182
	Real_03	5315.45	5375.76	4775.31	67	47	75	76981	450170	73900
	Real_04	5651.65	5934.69	5426.97	67	46	73	104369	384232	81439
	Real_05	5634.38	5632.27	5122.9	66	47	73	70931	437227	73899
	Real_06	5394.81	5592.01	4954.77	67	44	74	96681	410785	74201
	Real_07	5317.21	5427.88	4917.17	67	45	74	99046	424779	78706
	Real_08	5190.99	5450.88	4886.91	69	40	75	80331	583359	74485
	Real_09	5015.51	5451.24	4966.94	69	44	73	59155	399720	73673
	Real_10	5357.01	5498.69	5142.87	65	43	72	75656	451940	78115
60	Real_01	16342.4	19045.6	15981.3	80	61	81	5235	45954	49020
	Real_02	16158.1	19242.3	17217.3	81	60	80	4496	59167	47589
	Real_03	16182.3	19568.5	16278.6	81	60	81	4604	50293	47021
	Real_04	16853.3	19525.7	16339.3	80	61	81	4754	44417	43616
	Real_05	18086.5	19697.1	17123.8	79	60	80	4180	45299	43801
	Real_06	16245.4	18909.6	16398.2	80	61	80	4820	43769	51600
	Real_07	15743.4	19381.2	16145.5	81	60	81	4478	55891	45115
	Real_08	16361.8	19286.7	16272	80	59	81	4185	68763	43000
	Real_09	16079.4	19115.5	16587.2	81	61	80	4792	50833	44469
	Real_10	15746.8	18944	16818.6	81	63	80	4767	60511	50606
	Média				75.3	54.2	78.1	45783.3	315602.6	60253.5

Tabela 5: Resultado das instâncias com tamanhos 30 e 60 e grupos iguais ($a_g = b_g$)

n	Instância	Melhor Valor			Gap (%)			Nós		
		(2)	(3)	(4)	(2)	(3)	(4)	(2)	(3)	(4)
30	Geo_01	13615.9	13784.1	13498.4	74	48	79	97220	1077011	69584
	Geo_02	13089.6	13206.1	13059.7	74	48	80	100220	685317	63491
	Geo_03	10711.8	10767.9	10576.3	74	46	79	103620	786197	68981
	Geo_04	5410.4	5434.14	5342.07	73	45	78	101241	972124	61662
	Geo_05	13743.6	13826.1	13670.4	73	47	80	105360	1162859	67096
	Geo_06	14158.3	14330.6	14091	74	48	80	106976	658176	66800
	Geo_07	7560.2	7587.4	7474.72	73	43	79	103920	740364	64828
	Geo_08	13252.7	13340	13134.6	74	47	79	104706	988765	69157
	Geo_09	10633	10652.7	10493.6	74	48	79	105749	1119654	65065
	Geo_10	10334.2	10346.8	10234.7	74	47	79	100720	1062190	66035
60	Geo_01	44539.7	45405.3	44830.2	83	62	83	3682	159351	44249
	Geo_02	42327.4	43135	42618.8	83	62	83	4082	216023	46190
	Geo_03	50637.4	51356.5	50632.4	83	62	83	3384	203807	50412
	Geo_04	45024.8	46037.2	45422	83	61	83	3280	191719	42402
	Geo_05	48903.9	50076.7	49379.7	83	63	83	3113	182773	46602
	Geo_06	24588.3	25107.1	24793.7	83	61	83	3737	173270	45802
	Geo_07	37275.9	38003.1	37551.7	83	61	83	4893	232457	52802
	Geo_08	39406.9	40383	39788.1	83	62	83	4938	217291	53328
	Geo_09	35454	36204	35888.8	83	62	83	3098	191154	49405
	Geo_10	50743	51924.7	51234.3	83	62	83	2192	180466	43078
30	Int_01	5149	5288	4710	66	53	76	96041	991363	71256
	Int_02	5435	5585	5138	65	36	74	85320	426694	71978
	Int_03	5176	5206	4816	64	39	74	98520	435403	69600
	Int_04	5206	5288	4660	62	35	74	88851	422608	67899
	Int_05	5310	5385	4884	64	36	74	93520	478565	69947
	Int_06	4866	4973	4578	64	38	74	94300	461868	72180
	Int_07	5180	5127	4684	63	39	75	108741	422881	66049
	Int_08	5153	5303	4698	65	37	75	90441	481595	69829
	Int_09	5377	5320	4821	64	47	75	79771	574134	76163
	Int_10	5186	5444	4939	66	36	74	95488	470400	68034
60	Int_01	15818	17397	16452	81	61	81	5382	55921	44440
	Int_02	15015	17073	15771	81	62	81	3590	53280	46627
	Int_03	15263	17650	15659	81	60	81	4874	51253	44809
	Int_04	15582	17037	16424	81	62	80	3678	57068	52486
	Int_05	15246	17228	16062	81	62	81	4572	53172	49635
	Int_06	15078	17408	16121	81	61	81	3086	55043	41602
	Int_07	15703	17330	15501	81	60	82	5293	63101	42416
	Int_08	16027	17583	15319	80	61	82	3327	57773	48699
	Int_09	16023	17605	15848	80	60	81	3924	58421	46094
	Int_10	14968	16932	15812	81	62	81	5351	62485	45655
30	Real_01	5247.69	5464.94	4895.28	65	36	74	89120	639254	69857
	Real_02	5265.98	5547.19	4788.75	66	34	75	87120	641664	72500
	Real_03	5073.98	5204.59	4789.58	64	38	74	85820	753854	70466
	Real_04	5263.69	5561.75	4958	66	35	75	90115	628664	72300
	Real_05	5233.66	5447.95	4916.02	65	36	74	90720	680524	70040
	Real_06	5071.51	5432.79	4749.14	65	34	75	87220	613375	72253
	Real_07	5037.83	5186.26	4750.68	66	37	75	90920	506854	73573
	Real_08	5064.18	5355.53	4863.14	66	37	75	88820	731963	70987
	Real_09	5075.96	5277.19	4629.09	65	34	75	89920	628253	71378
	Real_10	5138.32	5269.71	4638.71	64	35	75	92441	688554	76036
60	Real_01	15003.0	17893.7	15752.4	82	60	81	3932	69021	45402
	Real_02	16009.0	18758.4	16413.4	80	58	81	4730	70310	55214
	Real_03	15660.6	18227.2	15989.7	81	60	81	5310	70158	38031
	Real_04	15331.8	18361.4	15955.8	81	59	81	3512	62959	52020
	Real_05	16226.3	18297.9	16495.3	81	61	81	3171	91017	48032
	Real_06	14589.3	18107.7	15609	82	58	81	4068	77126	46000
	Real_07	15301.4	18162.5	15791.4	81	59	81	5438	72050	49431
	Real_08	15529.1	18257.8	15916.2	81	59	81	3278	78199	50654
	Real_09	15560.8	18561	16465.1	81	57	81	3778	70225	45047
	Real_10	15605.8	18208	16154.8	81	60	81	3080	60940	46844
	Média				74.7	50.7	78.9	49545.2	402816	58307.2