

Programação de caminhões de minas de céu aberto minimizando o custo de manutenção

Victor São Paulo Ruela
Programa de Pós-Graduação em Engenharia Elétrica
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
Email: victorspruela@gmail.com

Resumo—Este relatório apresenta os resultados do trabalho final da disciplina de Otimização em Redes 2020/1, o qual consiste na implementação de um modelo de programação linear inteiro proposto na literatura [1] para a otimização da programação de equipamentos de mina de céu aberto. Utilizando a linguagem de programação Python e o solver CPLEX, a implementação foi realizada com sucesso e avaliada para diferentes instâncias do problema.

I. INTRODUÇÃO

O processo de mineração possui 5 etapas: prospecção, pesquisa, desenvolvimento, exploração, reclamação [5]. De acordo com [6], otimização é principalmente utilizada nas etapas de desenvolvimento e exploração. Na etapa de desenvolvimento, o objetivo é o planejamento de como a mina deverá ser explorada ao longo dos anos, além da tomada de decisão de quais investimentos serão realizados. As principais aplicações são o planejamento da produção, e a seleção e alocação de equipamentos. Na etapa de exploração, o objetivo é otimizar os recursos disponíveis de forma a se atingir a produção planejada. Sistemas de despacho de caminhões são umas das suas principais aplicações.

A mineração é um negócio de capital muito intenso, o qual requer um investimento da ordem de centenas de milhões de dólares somente para manter uma frota de equipamentos em operação. A manutenção da frota de caminhões é um dos maiores custos envolvidos, sendo responsável por 30-50% dos custos de transporte de materiais em uma mina de céu aberto que utiliza frotas de caminhões e escavadeiras [1]. Um investimento deste tamanho necessita que os equipamentos sejam utilizados da melhor forma ao longo do tempo, garantindo que os custos de operação sejam minimizados e a sua utilização maximizada, uma vez que pequenos ganhos podem resultar em uma economia na ordem de milhões. Logo, o uso de modelos de otimização possui grande potencial de aplicação neste contexto.

Em [6] é feita uma revisão das aplicações de pesquisa operacional no contexto de mineração, onde é constatado um grande interesse em problemas de programação inteira para otimização de diversos aspectos relacionados à frota da mina, visto o impacto que eles podem trazer para o negócio. Portanto, o objetivo deste trabalho será implementar o modelo de otimização da programação de caminhões para minas de céu aberto proposto por [1].

Neste artigo, o autor propõe um modelo de programação linear inteiro visando minimizar os custos de manutenção e respeitando os objetivos de produção anuais. Ele considera a variação do custo de manutenção com a idade do caminhões por meio de faixas de idade, bem como a sua disponibilidade variável para movimentação de material em cada ano. A formulação possui alguns aspectos únicos, como o sequenciamento das faixas de idade utilizando variáveis binárias. Além disso, o modelo é capaz de considerar vários tipos de caminhões com diferentes idades iniciais, bem como otimizar globalmente uma frota completa para todo o ciclo de operação de uma mina.

Algumas variações desta formulação escolhida existem na literatura, as quais propõe principalmente a adição de novas restrições e diferentes objetivos, tornando cenários mais complexos de programação da frota possíveis. Em [2], os autores acrescentaram a opção de compra de caminhões novos ao longo do tempo. Já em [7], uma formulação um pouco diferente é proposta onde é considerada a depreciação dos equipamentos de acordo com a sua idade. Já [3] desenvolveu um modelo mais completo que leva em consideração não só os custos de manutenção, mas também o consumo de combustível e a depreciação dos equipamentos. Portanto, nota-se que esse é um problema relevante na literatura, e com grandes oportunidades de continuidade.

II. METODOLOGIA

A. Formulação do problema

O problema foi formulado com base na descrição disponível em [1], a qual foi adaptada de acordo com a implementação realizada. Inicialmente, são feitas as seguintes definições sobre os índices, parâmetros e variáveis de decisão do problema, conforme visto na Tabela II-A, onde $t \in T = [1, \dots, t_{max}]$, $b \in B = [1, \dots, b_{max}]$ e $y \in Y = [1, \dots, y_{max}]$. A partir destas definições, o problema é formulado na sequência.

Índices

t	Identificador do caminhão
b	Faixas de idade
y	Período de tempo (anos)
c	Faixa de operação crítica

Parâmetros

$C_{t,b,y}$	Custo de manutenção (\$/hora) para um caminhão t na faixa b no y -ésimo período de tempo
FE_t	Custo de reparo do motor do caminhão t
$A_{t,y}$	Horas disponíveis do caminhão t no período de tempo y
M	Tamanho da faixa de idade (horas)
R_y	Total de horas de operação para um período de tempo y
I_t	Idade inicial dos caminhão t

Variáveis de decisão

$X_{t,b,y}$	Número de horas alocadas para o caminhão t , faixa de idade b no y -ésimo período de tempo
$Y_{t,b,y}$	1, se o caminhão t na faixa b utilizou todas as horas disponíveis no período de tempo y

Minimizar

$$\sum_{y \in Y} \sum_{t \in T} \sum_{b \in B} X_{t,b,y} C_{t,b,y} + \sum_{y \in Y} \sum_{t \in T} Y_{t,c,y} FE_t \quad (1)$$

Sujeito a

$$\sum_{b \in B} X_{t,b,y} \leq A_{t,y}, \quad \forall t \in T, \quad \forall y \in Y \quad (2)$$

$$\sum_{y \in Y} X_{t,b,y} \leq M, \quad \forall t \in T, \quad \forall b \in B \quad (3)$$

$$\sum_{k=1}^y X_{t,b,y} \geq MY_{t,b,y}, \quad \forall t \in T, \quad \forall b \in B, \quad \forall y \in Y \quad (4)$$

$$X_{t,(b+1),y} \geq M \sum_{k=1}^y Y_{t,b,k}, \quad \forall t \in T, \quad \forall b \in B, \quad \forall y \in Y \quad (5)$$

$$\sum_{t \in T} \sum_{b \in B} X_{t,b,y} = R_y, \quad \forall y \in Y \quad (6)$$

$$\sum_{y \in Y} \sum_{b \in B} X_{t,b,y} \leq Mb_{max} - I_t, \quad \forall t \in T \quad (7)$$

$$X_{t,b,y} \in \mathbb{Z}^+, \quad \forall t \in T, \quad \forall b \in B, \quad \forall y \in Y \quad (8)$$

$$Y_{t,b,y} \in \{0, 1\}, \quad \forall t \in T, \quad \forall b \in B, \quad \forall y \in Y \quad (9)$$

A restrição 2 garante o respeito à disponibilidade de horas de cada caminhão, enquanto que 3 não permite que horas alocadas ultrapassem o tamanho da faixa. As restrições 4 e 5 garantem que as faixas de idades sejam alteradas de forma sequencial. Já 6 e 7 garantem que o total de horas planejadas anualmente seja atingida e que os caminhões não sejam utilizados mais do que o valor máximo da faixa de idades, respectivamente. As restrições 8 e 9 definem o domínio das variáveis de decisão.

Isso resulta em um problema de programação linear inteiro, contendo somente variáveis de decisão inteiras e binárias. Levando em conta que o número de restrições e variáveis é bem alto, espera-se que este problema seja bem difícil de resolver em cenários de maior escala.

B. Geração das instâncias

Conforme visto na seção anterior, existem diversos parâmetros que precisam ser definidos para a execução deste modelo. Analisando as informações contidas em [1], constatou-se que os dados de disponibilidade e custo de manutenção dos caminhões estão incompletos, de forma que a instância executada pelo autor não possa ser replicada.

Por esse motivo, foi proposto a verificação da disponibilidade destes dados de minas de céu aberto reais que utilizam sistemas da *Hexagon Mining*, os quais infelizmente também não estavam disponíveis. Então optou-se por gerar estes valores de forma artificial, seguindo as recomendações da literatura. Por outro lado, dados como a quantidade e disponibilidade de caminhões estão parcialmente disponíveis, podendo ser usados também como um guia para esta tarefa. A forma como estes dados são gerados é descrita a seguir.

1) *Disponibilidade dos caminhões*: De acordo com [1], esta disponibilidade é calculada com base no seu fator de utilização efetivo. Em um cenário hipotético onde o caminhão seria operado sem parar por um ano inteiro, teríamos um valor máximo de 8760 (465 * 24) horas disponíveis, resultando num fator com valor de 1. Como a manutenção do equipamento é necessária na prática, e assumindo diariamente são gastas 2 horas para a sua manutenção, teríamos 8395 horas disponíveis anualmente, o que resulta em um fator de utilização de aproximadamente 91.6%. Como podem ocorrer flutuações nesse valor, os valores de disponibilidade serão amostrados de uma distribuição uniforme para um fator de disponibilidade entre 0.9 e 0.95 do total de horas anuais, conforme a Equação 10.

$$A_{t,y} \sim \mathcal{U}_{[0.9, 0.95]} \cdot 8760 \quad (10)$$

Em [3], os autores sugerem que a disponibilidade dos caminhões varia em função da sua idade, decrescendo até a sua idade crítica, e em seguida retornando aos patamares iniciais após o seu reparo. Entretanto, como isso necessitaria da modificação do modelo descrito anteriormente, a mesma não será implementada neste trabalho.

2) *Produção anual*: Seguindo a mesma ideia da seção anterior, a produção anual máxima possível consiste em multiplicar a quantidade de horas no ano pelo número de caminhões disponíveis para operação. Como isso é um cenário irreal por diversos motivos, devido a limitações da própria disponibilidade de material para ser transportado na mina, por exemplo, estes valores serão amostrados de uma distribuição uniforme no intervalo de 70 a 80% desta disponibilidade, conforme a Equação 11.

$$R_y \sim \mathcal{U}_{[0.7, 0.8]} \cdot 8760 \cdot t_{max} \quad (11)$$

Além disso, [3] sugere que a produção da mina deve ser reduzida a partir do penúltimo ano de operação, o que também é observado na instância descrita em [1]. De acordo com as faixas de valores informadas pelos autores, o penúltimo e último ano serão multiplicados por um fator de escala 0.5 e 0.1, respectivamente, quando o número de anos da instância for maior do que 3. Um exemplo pode ser visto na Figura 1. Vale ressaltar que a escolha dessas faixas deve ser feita de forma a garantir que o problema seja factível. Ou seja, não podemos ter uma produção anual que não possa ser atingida com os caminhões disponíveis.

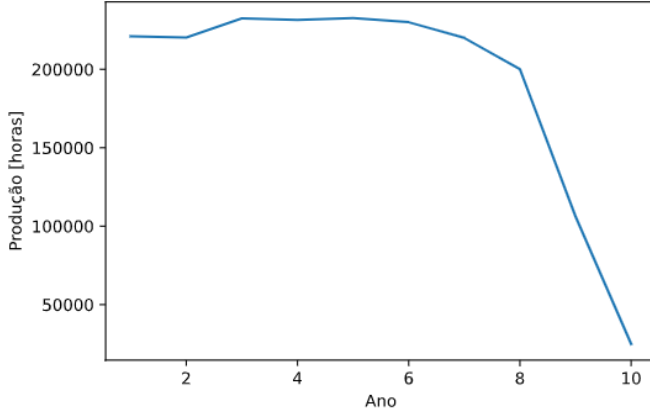


Fig. 1. Exemplo de produção anual

3) *Idade dos caminhões*: Embora fornecidos pelo autor de [1], por padrão os valores serão amostrados de uma distribuição uniforme no intervalo de 0 a 20000 horas.

4) *Custo de manutenção*: De acordo com [3], os custos de manutenção variam de forma não-linear de acordo com a idade do caminhão. Além disso, elas tendem a aumentar até um ponto onde um grande reparo é necessário, o que é representado pela faixa de idade crítica, após a qual o custo tende aos mesmos patamares de quando o caminhão era novo. Observando os valores fornecidos em [1], [2], serão geradas amostras na faixa de 15 a 235 (\$/hora).

Entranto, como os caminhões podem possuir idades iniciais distintas, os custos de manutenção precisarão ser ajustados, uma vez que a formulação do problema não considera a idade inicial para definir o preço atual do caminhão. Portanto, a matriz de custos gerada será ajustada de forma que caminhões com maior idade inicial tenham um custo maior, pois ele já inicia numa faixa de idade diferente.

Como a alocação dos caminhões é definida pelo otimizador, uma curva distinta será gerada para cada ano seguindo este formato, de forma a levar em consideração variações anuais no preço. Como as faixas de idade são incrementadas sequencialmente, é fácil ver que isto não viola o fato de que, embora um caminhão possa ter uma idade inicial diferente a cada ano, o custo sempre será referente à faixa de idade em que ele está atualmente.

Para cada caminhão t e ano y , o custo para uma determinada faixa de idade b é descrito pela Equação 12. Os valores

médios por faixa μ_b são gerados de forma igualmente espaçada dentro do intervalo $[5, 15]$ e em seguida elevados ao quadrado, simulando a não-linearidade. Além disso, um valor de ajuste é amostrado de uma distribuição normal com média 0 e desvio padrão 10 é acrescido. Um exemplo de curva pode ser visto na Figura 2.

$$C_{t,b,y} = (\mu_b)^2 + \mathcal{N}(0, 10) \quad (12)$$

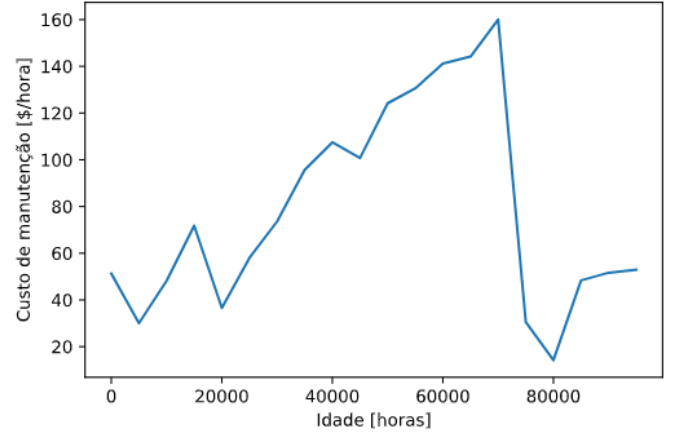


Fig. 2. Exemplo de curva de custos de manutenção em função da idade

C. Método de solução

Como foi formulado um problema de programação linear inteiro de larga escala e com muitas restrições, ele será resolvido utilizando o algoritmo *Branch and Cut* através do CPLEX versão 12.10.0.0. O trabalho será implementado na linguagem de programação Python versão 3.7.7, sendo necessários os pacotes *pandas*, *numpy*, *matplotlib*, *seaborn* e *docplex*. Instruções para a instalação e configuração da biblioteca *docplex* estão disponíveis em [4]. Os testes serão executados em um notebook com Intel Core i7 Quad Core 2.4GHz, 8GB de memória RAM.

D. Desenho dos experimentos

O algoritmo será avaliado sobre 4 instâncias distintas do problema, cada uma possuindo diferente tamanho, conforme descrito na Tabela I. Dessa forma, é possível validar a implementação e observar o quão difícil se torna esse problema aumentando a quantidade de variáveis e restrições.

Na instância *Artigo*, serão considerados os valores disponíveis em [1] para objetivos de produção, número de caminhões e suas respectivas idades iniciais. Os demais valores serão gerados conforme descrito na seção anterior. O número de faixas de idade será igual a 20 para todas as instâncias.

Os valores de cada instância serão persistidos em arquivos de texto para garantir sua reprodutibilidade, uma vez que os dados são gerados de forma aleatória. Como é utilizado um algoritmo exato, somente uma execução é necessária para cada instância. Serão utilizados como critério de parada um *gap* relativo mínimo ou tempo máximo de execução, quando for aplicável.

TABELA I
VALORES DE CADA INSTÂNCIA GERADA

	Pequena	Média	Grande	Artigo
Caminhões	5	20	45	34
Anos	3	5	10	10
M	2000	4000	5000	5000
Idade crítica	16000	40000	75000	75000

III. RESULTADOS

A. Instância pequena

Os valores de custo utilizados podem ser visto na Figura 3, onde é importante notar que eles foram gerados conforme descrito nas seções anteriores. Um total de 480 variáveis de decisão e 567 restrições foram criadas para esta instância e o esforço computacional foi bem baixo para a sua solução, uma vez que o CPLEX encontrou o ótimo global em menos de 0.5 segundos.

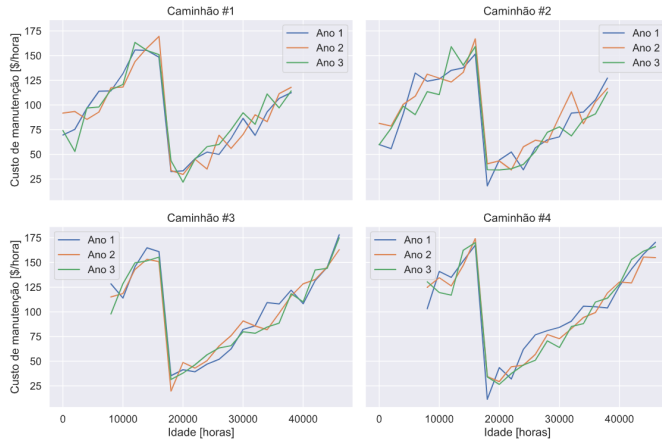


Fig. 3. Custos de manutenção por caminhão para a instância pequena

As idades iniciais foram definidas como $[0, 0, 8000 \text{ e } 8000]$, bem como uma produção de 25000 horas por ano. Portanto, isso consegue simular o cenário onde pelos menos 1 caminhão atingirá a idade crítica e comparar ao cenário onde possa ser vantajoso utilizar os caminhões mais velhos pois o custo é menor após o reparo do motor, verificando se a implementação está correta.

Na Figura 4, que contém a alocação de horas escolhidas pelo otimizador em cada ano, é possível notar que pelo menos dois caminhões atingiriam a idade crítica. Como o custo após o reparo do motor é muito menor, os caminhões mais antigos forma utilizados ao máximo no primeiro ano, de forma que nos dois anos seguintes eles possam ser utilizados com um custo menor.

Para os caminhões novos, observe que eles possuem praticamente a mesma alocação no primeiro ano, porém partir do segundo ano ela foi feita de forma intercalada. Isso aconteceu pois existem pequenas variações de custo entre os anos, como pode ser visto na Figura 3. A partir da Figura 5, é possível ver que a restrição de sequência na escolha das faixas está sendo respeitada corretamente, e também que caminhões com

idade iniciais e custos similares tiveram uma alocação de horas acumuladas bem próxima. Tendo em vista os resultados apresentados, é possível constatar que a implementação do modelo foi realizada com sucesso.

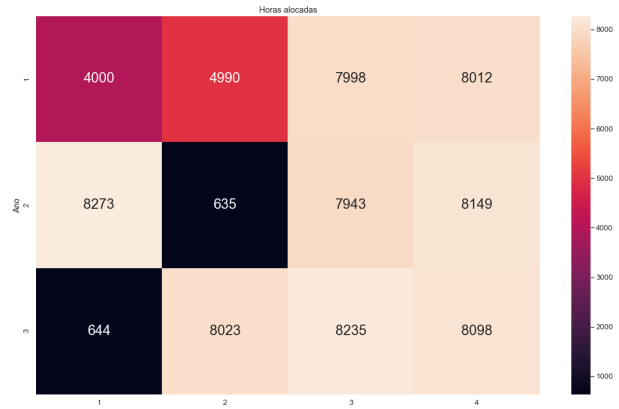


Fig. 4. Alocação de horas por caminhão em cada ano para a instância pequena

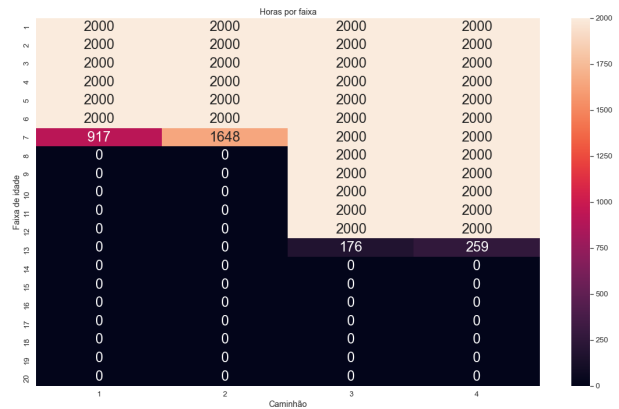


Fig. 5. Alocação de horas por caminhão em cada faixa de idade para a instância pequena

B. Instância média

Um total de 4000 variáveis de decisão e 4425 restrições foram criadas para esta instância e o esforço computacional também foi bem baixo para a sua solução, uma vez que o CPLEX encontrou o ótimo global em cerca de 3 segundos. Para o problema deste tamanho, já fica mais difícil definir quais seriam os resultados esperados, então uma análise mais detalhada será omitida. É interessante observar que o otimizador optou por alocar um dos caminhões além da sua idade crítica, pois era vantajoso do ponto de vista de custo. Isso pode ser visto na Figura 7, enquanto que a alocação escolhida pelo otimizador está na Figura 6.

C. Instância grande

Um total de 18000 variáveis de decisão e 18955 restrições foram criadas para esta instância e o esforço computacional

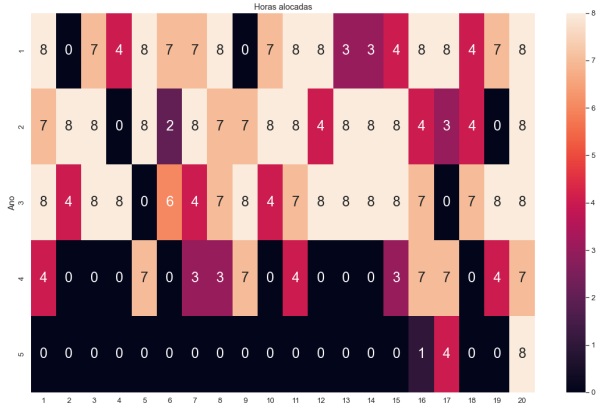


Fig. 6. Alocação de horas por caminhão em cada ano para a instância média. Os valores estão em múltiplos de 1000 horas.

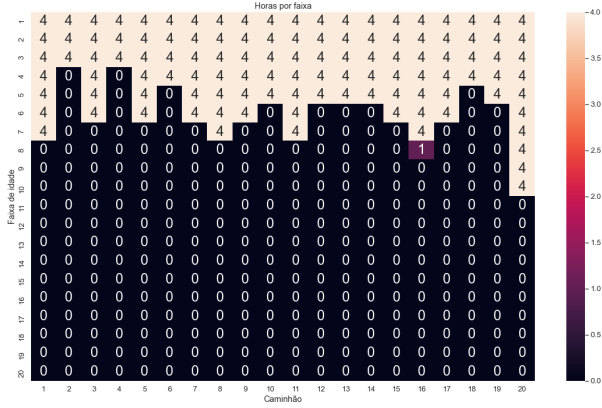


Fig. 7. Alocação de horas por caminhão em cada faixa de idade para a instância média

foi relativamente grande para a sua solução, uma vez que o CPLEX chegou ao gap relativo de 1% em cerca de 3.5 minutos. Da mesma forma que para a instância média, uma análise detalhada da solução é muito difícil para um problema deste tamanho, a qual será omitida. É importante notar que alguns caminhões são alocados de forma bem próxima da idade máxima de 100000 horas, e que vários foram utilizados além da idade crítica como opção para reduzir os custos. Isso pode ser visto na Figura 8, enquanto que a alocação escolhida pelo otimizador está na Figura 9.

Um segundo experimento foi realizado para verificar o quanto este resultado poderia ser melhorado, o que foi feito reduzindo o gap mínimo para 0.1% e deixando o algoritmo executar por 30 minutos. A evolução do gap pode ser vista na Figura 10, onde é possível constatar que após o gap de 1%, o algoritmo conseguiu melhorar muito pouco a solução, demorando mais de 20 minutos para chegar a um gap final de 0.68%, bem longe do limite especificado. Isso mostra que a etapa de polimento das soluções é muito difícil para este problema.

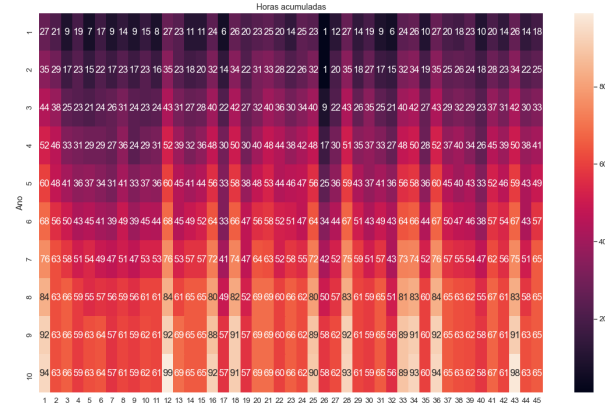


Fig. 8. Idade dos caminhões ao longo dos anos. Os valores estão em múltiplos de 1000 horas.

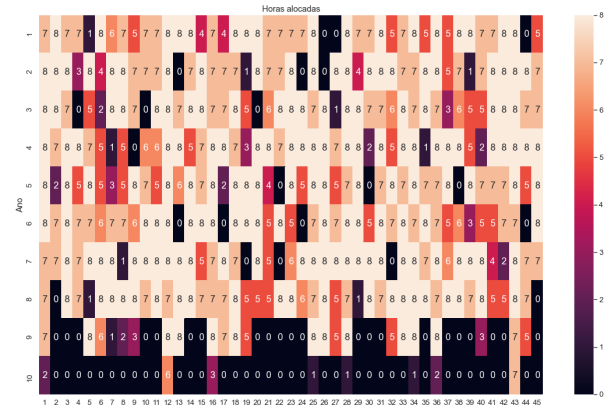


Fig. 9. Alocação de horas por caminhão em cada ano para a instância grande. Os valores estão em múltiplos de 1000 horas.

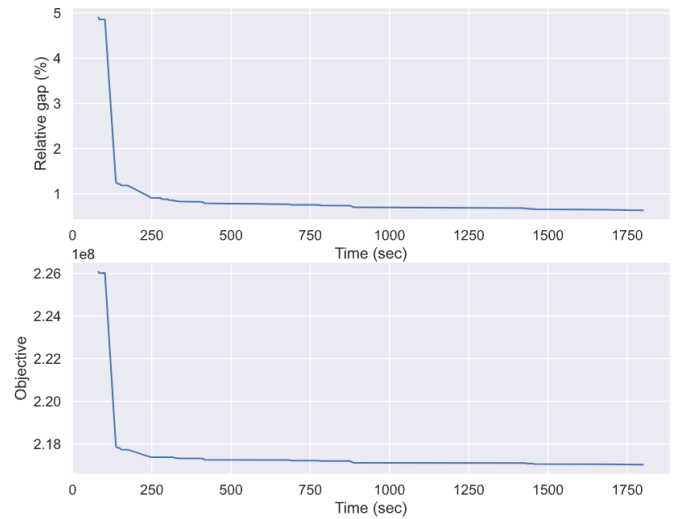


Fig. 10. Progresso do algoritmo para a instância grande

D. Instância artigo

Um total de 13600 variáveis de decisão e 14324 restrições foram criadas para esta instância e o esforço computacional

foi muito grande para a sua solução, uma vez que o CPLEX executou durante 30 minutos para chegar a um gap relativo de 6.40%. Observe que embora possua um tamanho menor que a instância grande, sua solução foi muito mais difícil. Isso se deve ao fato de que mais da metade dos caminhões possuem mais de 40000 horas, ou seja, rapidamente eles irão chegar próximo da idade crítica, de forma que as restrições sejam mais fortes devido à escassez de recursos. Conforme a Figura 11, após 400 segundos de execução o algoritmo consegue evoluir muito pouco a qualidade da solução até o limite de tempo estipulado. É interessante notar que o algoritmo demorou menos de 1 minuto para chegar num gap de cerca de 14%.

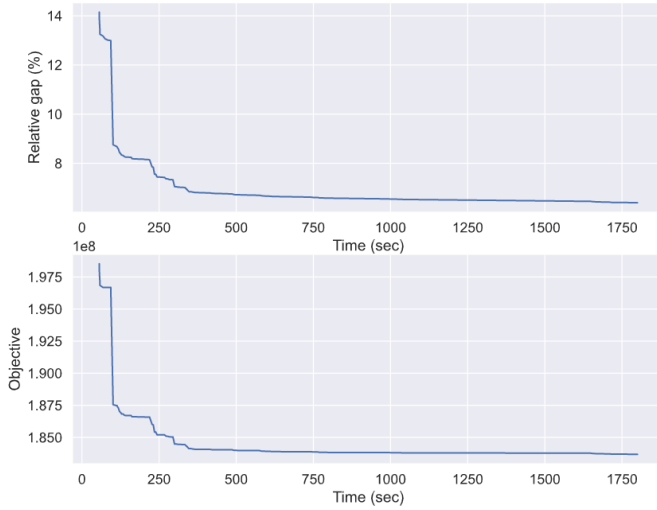


Fig. 11. Exemplo de curva de custos de manutenção em função da idade

Como boa parte dos caminhões estão com idade avançada, espera-se que grande parte seja utilizada até a seu limite de horas de operação, o que pode ser visto na Figura 12, que exibe o total de horas acumuladas ao longo dos anos. A alocação nas faixas de idade podem ser vistas na Figura 13. Contrariando um pouco a lógica, o algoritmo encontrou ser mais barato realizar o reparo do motor de 20 caminhões ao invés de utilizar ao máximo os mais novos. Isso pode ter sido causado pela característica de crescimento quadrático dos custos nas faixas de idade próximas da crítica. Portanto, pequenas alterações na solução pode levar a um aumento grande do objetivo, o que pode explicar a lentidão da evolução do algoritmo. Isso leva a crer que o uso de alguma meta-heurística neste ponto seja interessante.

Alguns parâmetros do CPLEX foram ajustados como forma de tentar melhorar o desempenho do algoritmo. Como houve pouco progresso após determinado tempo, a estratégia de *probing* foi ajustada para o nível mais agressivo, o parâmetro de ênfase alterado para priorizar a otimalidade e a heurística de busca local foi ativada, conforme recomendação da IBM e as observações anteriores. Executando o algoritmo por 30 minutos, foi possível chegar a um gap de 5.76%, o que mostra o ajuste de parâmetros tem potencial para melhorar o

desempenho computacional obtido, embora tenha se chegado a uma solução não muito melhor neste teste. Como o autor de [1] reportou um esforço computacional 30 minutos para atingir um gap de 5%, os resultados estão coerentes com a literatura, embora os dados de custo e disponibilidade foram distintos.

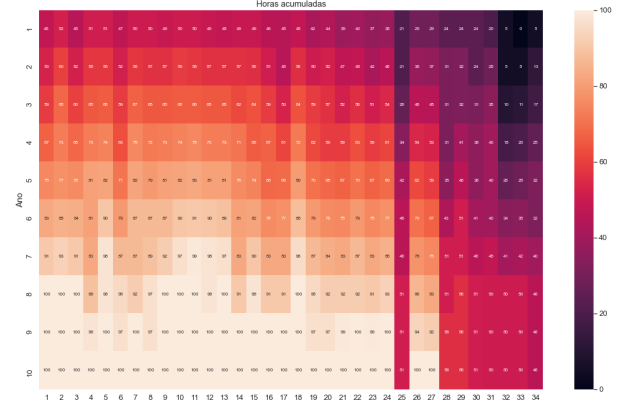


Fig. 12. Idade dos caminhões ao longo dos anos. Os valores estão em múltiplos de 1000 horas.

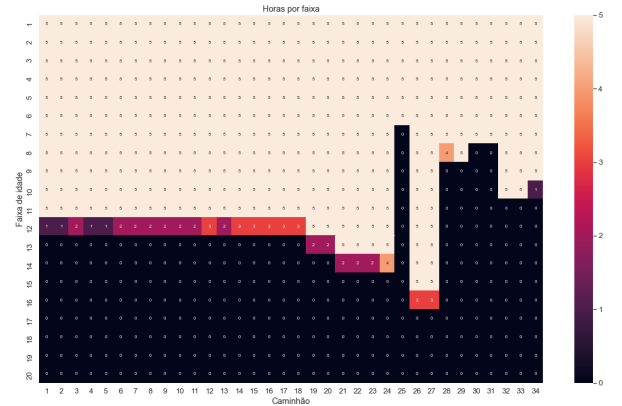


Fig. 13. Alocação de horas por caminhão em cada faixa de idade para a instância grande

IV. CONCLUSÃO

Neste trabalho foi implementado um modelo de programação linear inteiro para otimização da programação de equipamentos de mina de céu aberto proposto na literatura [1]. Utilizando o solver CPLEX e dados gerados artificialmente ou extraídos das referências utilizadas, foi possível analisar as soluções encontradas pelo algoritmo para 4 instâncias diferentes do problema. Por meio delas, foi possível constatar o quanto o problema fica mais difícil de acordo com o tamanho da instância, e também em relação aos dados considerados de custo, idades iniciais e disponibilidade. Para as 3 instâncias artificiais, o CPLEX não teve grandes dificuldades para a sua solução. Entretanto, para dados mais

próximos da realidade, foi notável o quão mais difícil o problema se tornou.

A maior dificuldade do trabalho consistiu na geração dos dados artificiais e no entendimento das equações da formulação disponível na literatura. Até a obtenção de uma implementação que respondia de acordo com o esperado, diversos problemas ocorreram, dentre eles falta de respeito à ordem das faixas de idade e dados de custo inconsistentes com a idade inicial dos caminhões, por exemplo. Além disso, visualizar um problema onde cada variável de decisão possuía 3 índices não foi muito trivial, mas após o correto entendimento de como isso poderia ser feito, avaliar os resultados do algoritmo se tornou bem mais eficiente.

Todos os códigos utilizados no desenvolvimento deste trabalho também estão disponíveis online no repositório <https://github.com/vicrsp/otredes-pgge>.

REFERÊNCIAS

- [1] Topal, E. and Ramazan, S., 2010. A new MIP model for mine equipment scheduling by minimizing maintenance cost. *European Journal of Operational Research*, 207(2), pp.1065-1071.
- [2] Fu, Z., Topal, E. and Erten, O., 2014. Optimisation of a mixed truck fleet schedule through a mathematical model considering a new truck-purchase option. *Mining Technology*, 123(1), pp.30-35.
- [3] Nakousi, C., Pascual, R., Anani, A., Kristjanpoller, F. and Lillo, P., 2018. An asset-management oriented methodology for mine haul-fleet usage scheduling. *Reliability Engineering & System Safety*, 180, pp.336-344.
- [4] IBM® Decision Optimization CPLEX® Modeling for Python (DOcplex) V2.15 documentation: <http://ibmdecisionoptimization.github.io/docplex-doc/>. Acessado em 27/09/2020
- [5] The 5 Stages of the Mining Life Cycle — Supply Chain & Operations — Mining Global. <https://www.miningglobal.com/operations/gifs-5-stages-mining-life-cycle>. Acessado em 06/10/2020
- [6] Newman, Alexandra M., Enrique Rubio, Rodrigo Caro, Andrés Weintraub, and Kelly Eurek. "A review of operations research in mine planning." *Interfaces* 40, no. 3 (2010): 222-245.
- [7] Burt, C., Caccetta, L., Fouché, L. and Welgama, P., 2016. An MILP approach to multi-location, multi-period equipment selection for surface mining with case studies. *Journal of Industrial & Management Optimization*, 12(2), p.403.