



**UNIVERSIDADE  
FEDERAL RURAL  
DE PERNAMBUCO**



Estéfane Paula Bezerra de Oliveira

**Acompanhamento preventivo e agendamento de  
pacientes com câncer utilizando programação inteira  
e matroides**

**Recife**

Maio de 2023

Estéfane Paula Bezerra de Oliveira

## **Acompanhamento preventivo e agendamento de pacientes com câncer utilizando programação inteira e matroides**

Artigo apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE  
Departamento de Estatística e Informática  
Curso de Bacharelado em Sistemas de Informação

**Orientadora: Silvana Bocanegra**

Recife  
Maio de 2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal Rural de Pernambuco  
Sistema Integrado de Bibliotecas  
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

---

B574a OLIVEIRA, ESTEFANE PAULA BEZERRA DE  
Acompanhamento preventivo e agendamento de pacientes com câncer utilizando programação inteira e matroides /  
ESTEFANE PAULA BEZERRA DE OLIVEIRA. - 2023.  
25 f. : il.

Orientadora: Silvana Bocanegra.  
Inclui referências.

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco, Bacharelado em  
Sistemas da Informação, Recife, 2023.

1. programacao-inteira. 2. matroides. 3. otimizacao. 4. ampl. 5. cancer. I. Bocanegra, Silvana, orient. II. Título

CDD 004

---

ESTEFANE PAULA BEZERRA DE OLIVEIRA

ACOMPANHAMENTO PREVENTIVO E AGENDAMENTO DE  
PACIENTES COM CÂNCER UTILIZANDO PROGRAMAÇÃO  
INTEIRA E MATROIDES

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Aprovada em: 26 de Abril de 2023.

BANCA EXAMINADORA

Silvana Bocanegra (Orientadora)  
Departamento de Estatística e Informática  
Universidade Federal Rural de Pernambuco

Marcelo Gama  
Departamento de Estatística e Informática  
Universidade Federal Rural de Pernambuco

# Acompanhamento preventivo e agendamento de pacientes com câncer utilizando programação inteira e matroides

Estéfane Paula Bezerra de Oliveira <sup>1</sup>, Silvana Bocanegra <sup>2</sup>

<sup>1</sup>Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco  
Rua Dom Manuel de Medeiros, s/n, - CEP: 52171-900 – Recife – PE – Brasil

estefane.oliveira@ufrpe.br, silvana.bocanegra@ufrpe.br

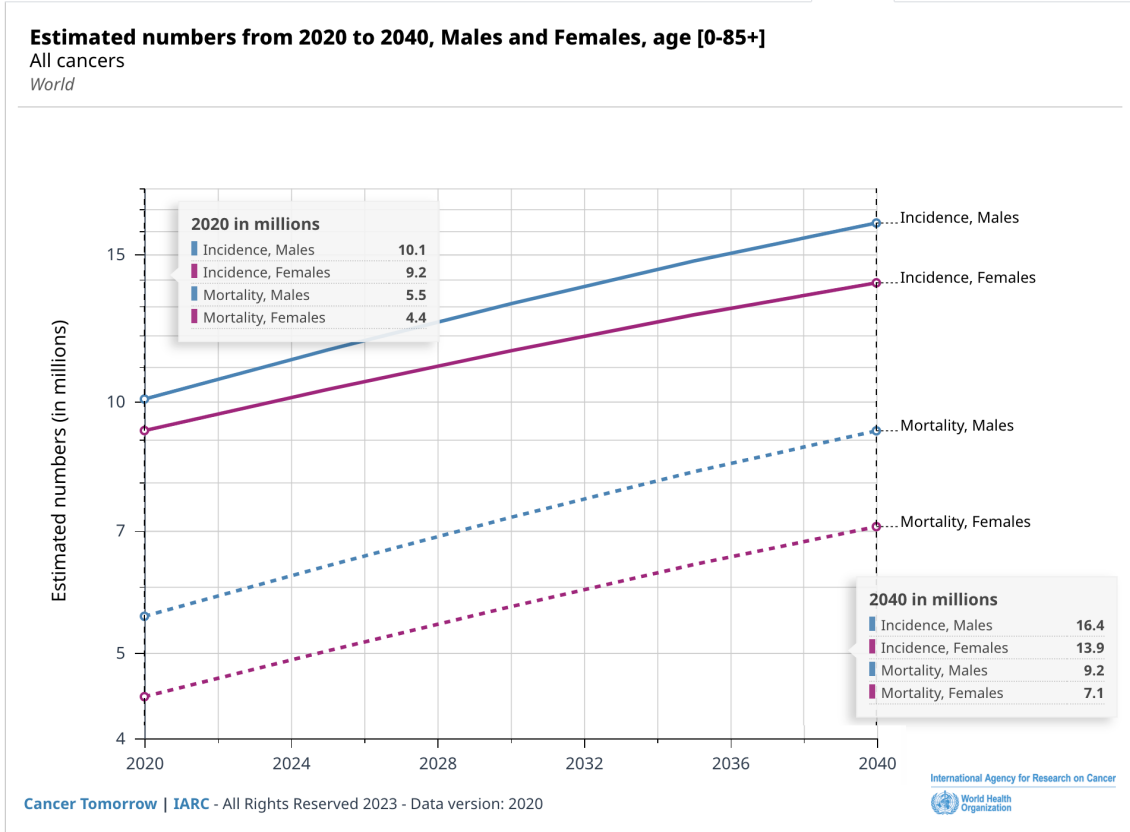
**Resumo.** *Dentro do contexto de crescimento do número de casos de câncer globalmente, modelos de otimização podem ajudar médicos e profissionais de saúde a garantir a melhor utilização dos recursos disponíveis. Este trabalho apresenta dois modelos de otimização restrita, um para a priorização de pacientes para o acompanhamento preventivo de câncer cervical e outro para o escalonamento de pacientes oncológicos em tratamento, ambos utilizando programação inteira e AMPL. Além disso, é apresentada uma resolução do modelo de escalonamento de pacientes utilizando matroides. Os modelos foram testados em instâncias geradas aleatoriamente, obtendo como resultado uma lista de pacientes hipotéticos prioritários para execução de exames, bem como uma lista de pacientes agendados em seus respectivos turnos.*

**Abstract.** *Within the context of the growing number of cancer cases globally, optimization models can help physicians and health professionals to ensure the best use of available resources. This paper presents two constrained optimization models, one for the prioritization of patients for preventive follow-up of cervical cancer and the other for the scheduling of oncology patients in treatment, both using integer programming and AMPL. In addition, a performance analysis of a patient scheduling algorithm using matroids is presented. The models were tested on randomly generated instances, resulting in a list of hypothetical priority patients for performing tests, as well as a list of patients scheduled in their respective shifts.*

## 1. Introdução

O câncer é uma das doenças mais prevalentes e letais ao redor do mundo. Segundo a Agência Internacional de Pesquisa sobre o Câncer, estima-se que em 2020, 19.3 milhões de pessoas ao redor do mundo tiveram câncer, e 9.9 milhões morreram da doença [1]. Em 2040, a previsão é que esses números aumentem consideravelmente, posicionando o câncer como o principal problema de saúde pública no mundo e entre as quatro principais causas de morte prematura (antes dos 70 anos de idade) na maioria dos países, como pode ser visto na Figura 1.

No Brasil, de acordo com o relatório divulgado em 2019 pelo INCA (Instituto Nacional de Câncer) [2], são previstos 625 mil novos casos de câncer para cada ano do triênio 2020/2022, dentre os quais 7.4 por cento são referentes ao câncer cervical, tornando-o o terceiro em número de casos em mulheres em território nacional. Mundialmente, esse número representa 528.000 novos casos e 266.000 mortes por ano [3].



**Figura 1. Estimativa de incidência e mortalidade de câncer entre 2020 a 2040 por gênero, entre idades de 0 a 85 anos. Fonte: Agência Internacional de Pesquisa sobre o Câncer, 2020. [1]**

Diante desse cenário, se faz essencial o investimento em tecnologias que possam atuar em prevenção, diagnóstico e tratamento de câncer. O acompanhamento preventivo, por exemplo, pode ajudar a reduzir o risco de desenvolver a doença e/ou aumentar as chances de sucesso caso a detecção ocorra no início da doença. Nem todos os tipos de câncer podem ser prevenidos, já que alguns são causados por fatores genéticos ou hereditários, porém para tipos como o câncer cervical, parece existir uma correlação entre os fatores de risco e a chance de desenvolver a doença, permitindo assim que seja possível estimar o grau de risco que uma determinada pessoa se encontra com base nesses fatores.

Além da prevenção, o escalonamento de pacientes já com a doença demanda tempo. O agendamento depende das necessidades de cada paciente e do tipo de tratamento que estão recebendo. Alguns pacientes podem precisar de exames frequentes e consultas com seus médicos, enquanto outros podem precisar de tratamentos mais intensivos, como cirurgia ou quimioterapia. A otimização pode auxiliar na carga de trabalho dos médicos e enfermeiros, no uso do equipamento e aumentar as chances de sucesso no tratamento de pacientes, dando espaço para apoiar a alta na demanda de casos prevista para os próximos anos.

Este trabalho tem como principal objetivo desenvolver modelos matemáticos-computacionais baseados em otimização para auxiliar na prevenção de câncer cervical

e no agendamento de pacientes oncológicos. Como objetivos específicos é possível elencar:

1. Propor um modelo para auxiliar médicos na definição da prioridade no acompanhamento preventivo de câncer cervical estimado a partir dos fatores de risco;
2. Propor modelo que auxilie profissionais da saúde no processo de escalonamento de pacientes para tratamentos oncológicos;
3. Analisar o desempenho de um algoritmo guloso com matróides para a resolução do problema de priorização de pacientes.

Em termos de organização, este trabalho está dividido nas seguintes seções: Trabalhos relacionados, que apresenta uma revisão de trabalhos na literatura que buscam resolver problemas similares de otimização e agendamento de pacientes; Fundamentação teórica, onde são apresentados os conceitos a serem tratados nesse trabalho, como fatores de risco do câncer cervical, otimização restrita, AMPL e matróides; Materiais e métodos, que abrange os modelos propostos neste trabalho e a metodologia utilizada; Resultados (Estudo de Caso), com a apresentação dos resultados obtidos a partir da execução do modelo com os parâmetros selecionados; e por fim a seção de Discussões e Trabalhos futuros, onde são trazidas algumas discussões e questões a serem abordadas em trabalhos futuros.

## **2. Trabalhos Relacionados**

As aplicações de otimização restrita na área de saúde são inúmeras, desde a gestão da capacidade de atendimento, localização de unidades de saúde, até a entrega eficiente de suprimentos [4]. Dentro dessas aplicações, o citado problema de escalonamento de pacientes é abordado em alguns artigos como [4], [5] [6], e [7]. Cada pesquisa propõe seu próprio modelo de como otimizar o escalonamento de pacientes no tratamento com radioterapia, definindo o conjunto de restrições e criando variáveis de decisões que fazem sentido para cada contexto.

Problemas reais relacionados a serviços de saúde, em geral, são dinâmicos, estocásticos, não lineares e contínuos. Devido a complexidade de criar modelos que levem em consideração todos esses critérios, são propostos modelos mais simplificados, que cheguem perto o suficiente da solução ótima. Alguns exemplos são os modelos criados em [5] [6], e [7] que utilizam programação linear inteira para resolver o problema de escalonamento de pacientes. Cada um dos modelos seguem diferentes tipos de restrições, como já citado, porém todas consideram variáveis de decisão semelhantes, que podem assumir os valores de 1 se o paciente for afeto/agendado àquele turno ou 0 se não.

Em [5], o modelo de otimização é aplicado em um Hospital da Luz em Lisboa. O objetivo principal é maximizar o número de tratamentos por dia e minimizar o tempo de inatividade do serviço. Para alcançar esse objetivo, o estudo começa com a compreensão do funcionamento do serviço de radioterapia e a coleta de dados sobre os tempos de tratamento de cada paciente. Um escalonamento inicial é elaborado, considerando 9 tipos de tratamento de câncer diferentes, e revisado semanalmente para planejar a semana seguinte. O problema é resolvido com o Solver do Excel e a solução é registrada em uma planilha Excel usando um programa em Visual Basic for Applications (VBA).

De forma análoga, [6] propõe dois modelos para otimização do escalonamento de pacientes no Hospital da Luz utilizando programação Linear Inteira, que são testados

numa instância gerada aleatoriamente com dimensão próxima da real. É realizada uma simplificação dos tipos de câncer, considerando apenas 2 tipos. Os resultados obtidos foram analisados comparativamente procurando-se igualmente estudar uma possibilidade de combinação de ambos modelos. Como no exemplo acima, o modelo é executado utilizando VBA e o OpenSolver [6].

Em [7], é discutido o desenvolvimento e a validação de um escalonador para aceleradores lineares usados em tratamento radioterápico de um hospital no Reino Unido. O artigo compara o agendamento de tratamentos de radioterapia com outros problemas de agendamento de consultas e mostra que embora exista um grande número de modelos matemáticos e de simulação sendo usados para abordar o problema, estes não consideram todas as restrições do escalonamento do mundo real. O artigo apresenta um novo modelo de programação linear inteira para agendamento de tratamentos de radioterapia e avalia a eficácia do modelo. Neste trabalho, todos os experimentos são executados usando o ILOG CPLEX 12.1 em um ambiente Linux [7].

Em termos de estimar riscos de doenças, [8] desenvolve um modelo matemático computacional baseado na teoria de conjuntos de lógica fuzzy para estimar o risco de uma pessoa desenvolver a doença e auxiliar os especialistas na prevenção. Foram estudados os principais fatores de risco e criadas variáveis linguísticas para subsidiar a construção do sistema de regras fuzzy. Com a seleção de variáveis como idade, idade da primeira relação sexual, histórico de câncer dos pais, presença de ISTs, tabagismo, quantidade de parceiros homens no último ano e vacinação contra HPV, o risco do paciente desenvolver câncer de colo uterino pode ser classificado em cinco categorias: muito baixo, baixo, moderado, alto e muito alto.

Em [9], é abordado um problema de otimização que consiste em priorizar um conjunto de requisitos para serem desenvolvidos, considerando as restrições de tempo e esforço da equipe de desenvolvimento, além de equilibrar a satisfação dos clientes. O modelo apresentado neste trabalho foi utilizado como base para a criação do modelo de priorização de pacientes com base nos fatores de risco.

### **3. Fundamentação Teórica**

Nesta seção, serão abordados conceitos essenciais relacionados ao câncer e seus fatores de risco, bem como será apresentada a otimização restrita como uma abordagem matemática para lidar com problemas complexos na área da saúde. Além disso, serão discutidos o uso do software AMPL (A Mathematical Programming Language) como uma ferramenta poderosa para modelagem e resolução de problemas de otimização, e a aplicação de matroides como uma estrutura matemática possível para representar e solucionar problemas combinatórios. O entendimento desses conceitos é fundamental para a compreensão dos modelos de otimização de problemas relacionados ao câncer apresentados na Seção 4.

#### **3.1. Câncer e fatores de risco**

O câncer do colo do útero, também conhecido como câncer cervical, surge devido à infecção crônica por determinados tipos do papilomavírus humano (HPV) que são considerados oncogênicos. Os tipos de HPV 16 e 18, em particular, estão associados a um maior risco de progressão para lesões precursoras que, se não forem detectadas e tratadas, podem evoluir para o câncer ao longo do tempo [2].



A incidência de câncer cervical varia ao redor do mundo, com taxas mais elevadas em países em desenvolvimento devido à falta de acesso ao rastreamento e aos cuidados de saúde adequados. Segundo a Organização Mundial da Saúde (em inglês WHO), o câncer do colo do útero é o quarto tipo de câncer mais comum em mulheres em todo o mundo, com cerca de 604 mil casos novos, representando 6,5% de todos os tipos de câncer em mulheres [3]. Estes são alguns fatores de risco associados ao desenvolvimento do câncer cervical de acordo com o INCA [2]:

1. Infecção pelo vírus do papiloma humano (HPV);
2. Tabagismo;
3. Uso prolongado de anticoncepcionais hormonais;
4. Sistema imunológico comprometido;
5. Histórico familiar de câncer de colo de útero;
6. Vida sexual precoce e/ou múltiplos parceiros sexuais;
7. Idade;
8. Idade da Primeira relação sexual.

A infecção por HPV é uma causa necessária para o desenvolvimento do câncer do colo do útero [2]. O HPV é transmitido por meio do contato sexual e pode infectar tanto homens quanto mulheres. Fatores como imunodeficiência, idade precoce de início da atividade sexual, tabagismo e a multiplicidade de parceiros sexuais aumentam a chance de infecção pelo vírus, logo são considerados fatores de risco para o desenvolvimento deste tipo de câncer.

O uso prolongado de contraceptivos hormonais também pode aumentar o risco da doença, especialmente em mulheres que os utilizam por mais de cinco anos. É importante destacar que o câncer cervical é tratável se for detectado precocemente através de exames de rotina, como o exame Papanicolaou (PAP) e testes de HPV.

### 3.2. Otimização restrita

Uma das abordagens utilizadas para resolver problemas complexos com muitas restrições é chamada otimização restrita. Segundo [10], constrained optimization - em português otimização restrita - pode ser definida como um conjunto de métodos para identificar de forma eficiente e sistemática a solução ótima para um problema que possui diversas soluções possíveis e restrições identificadas. O processo para modelagem de um problema de otimização restrita é descrito como:

1. Identificar as variáveis de decisão;
2. Modelar o problema estudado como uma função, chamada de função objetivo;
3. Mapear as restrições do problema e;
4. Computar os valores de entrada que podem maximizar ou minimizar a função, dependendo do objetivo do estudo.

Ainda segundo os autores, a notação padrão para qualquer problema de otimização com o objetivo de encontrar o máximo é a seguinte :

Maximizar:

$$z = f(x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_k) \quad (1)$$

sujeito a:

$$c_j(x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_k) \leq C_j \quad (2)$$

para  $j = 1, 2, \dots, m$ .

onde  $x_1, x_2, \dots, x_n$  são as variáveis de decisão,  $f(x_1, x_2, \dots, x_n)$  é a função objetivo e  $c_j(x_1, x_2, \dots, x_n, p_1, p_2, \dots, p_k) \leq C_j$  são as restrições. Note que as restrições podem incluir desigualdade e igualdade e que a função objetivo e as restrições também incluem parâmetros  $p_1, p_2, \dots, p_k$ , que não são variadas no problema de otimização.

No que tange à complexidade, problemas de otimização restrita podem ser modelados de diversas maneiras: estático, quando não inclui um componente de tempo, logo decisões tomadas em um momento não afetam as seguintes; ou dinâmico, onde o problema de otimização e parâmetros mudam com o passar do tempo; determinístico, onde as informações são assumidas como certas; ou estocástico, onde a incerteza das informações é adereçada; linear, onde os custos e os benefícios são tidos como constantes; ou não linear, onde a função objetivo e as restrições têm uma relação não linear com os parâmetros do modelo; e por fim inteira (ou discreta), quando as variáveis de decisão podem assumir apenas valores discretos; ou contínua, quando estas podem assumir valores fracionários [10].

A resolução de problemas de programação inteira, como é o caso desse trabalho, pode ser difícil e demorada, principalmente quando o número de variáveis de decisão é grande pois é sabido que estes problemas são NP-completos. No entanto, existem vários métodos de resolução eficientes disponíveis, como o método Branch and Bound e o método Branch and Cut e heurísticas [9]. Além disso, existem softwares especializados, como o Gurobi, CPLEX e SCIP, que são capazes de resolver problemas de programação inteira de forma eficiente [11].

### 3.3. AMPL

A Mathematical Programming Language (AMPL) é uma linguagem abrangente e poderosa para modelagem matemática. Foi desenvolvida por Robert Fourer, David Gay e Brian W. Kernighan nos Laboratórios Bell com o objetivo de fornecer uma notação intuitiva e semelhante à álgebra para a formulação de modelos [12]. O AMPL pode ser utilizado para resolver problemas de programação linear, não-linear e inteira. Na Figura 2, é apresentado um modelo básico criado em AMPL.

```
set P;  
param a {j in P};  
param b;  
param c {j in P};  
param u {j in P};  
var X {j in P};  
maximize Total_Profit: sum {j in P} c[j] * X[j];  
subject to Time: sum {j in P} (1/a[j]) * X[j] <= b;  
subject to Limit {j in P}: 0 <= X[j] <= u[j];
```

**Figura 2. Modelo básico em AMPL (arquivo "file prod.mod"). Extraído de [11].**

Além da criação do modelo, se faz necessário criar um arquivo com os dados para serem executados, como pode ser conferido na Figura 3.

```

set P := bands coils;

param:      a      c      u      :=
  bands    200    25    6000
  coils    140    30    4000 ;

param b := 40;

```

**Figura 3. Exemplo de dados utilizados no modelo da Figura 2 (arquivo "file prod.dat"). Extraído de [11].**

O AMPL não resolve os problemas diretamente, apenas os traduz para uma forma que possa ser compreendida e resolvida por algoritmos de otimização como o MINOS ou o CPLEX. De acordo com [11], o CPLEX, ou IBM ILOG CPLEX Optimizer, é capaz de solucionar problemas complexos de programação, incluindo problemas de programação inteira, programação linear com grande escala, programação quadrática convexa e não convexa, além de problemas convexos quadraticamente restritos. No entanto, para problemas mais difíceis, como problemas de otimização global ou problemas de programação inteira mista não lineares, são necessários solvers diferentes.

### 3.4. Matroides

Os problemas de otimização restrita em que as restrições formam uma estrutura de matroide podem ser resolvidos usando a técnica de matroides. As matróides são generalizações abstratas de matrizes, onde tem uma definição de "estrutura independente", e essa estrutura se prende a matróides finitos ou definidos sobre conjuntos finitos. Combinados com um método guloso, é possível encontrar soluções ótimas de maneira eficiente, sem a necessidade de testar todas as combinações possíveis de soluções.

A teoria das matroides surgiu com o objetivo de unir e melhorar os resultados de um determinado problema, onde Hassler Whitney conseguiu relacionar as ideias de independência em teoria dos grafos e independência linear no estudo de espaços vetoriais. O grande interesse pela teoria das matroides é pelo fato do algoritmo se concentrar em problemas combinatórios, fazendo com que mostre a melhor possibilidade possível.

Segundo Oxley [13], seja uma matróide  $M = (E, F)$ , onde a entrada  $E$  é um conjunto finito de elementos e  $F$  é uma família de subconjuntos de  $E$  tal que obedecem as seguintes proposições:

- I.  $\emptyset \in F$  e todos os subconjuntos próprios de um conjunto  $I \in F$  estão em  $F$ , ou seja, se  $I_1 \subset I_2$  e  $I_2 \in F$ , então  $I_1 \in F$ ;
- II. Se  $I_p$  e  $I_{p+1}$  são subconjuntos de  $F$  contendo  $p$  e  $p + 1$  elementos, respectivamente, então existe  $e \in I_{p+1} - I_p$  tal que  $I_p + e \in F$ .

O método guloso consiste em construir uma solução passo a passo, selecionando a cada etapa a solução que fornece a maior melhoria possível para a função objetivo, sujeita às restrições do matroide. Esses algoritmos são chamados de ganancioso/guloso porque percorrem o caminho mais ambicioso, sem olhar consequências. Eles são mais fáceis de usar e compreender, possuindo diversas aplicações na literatura. Segundo Cormen [14],

o método guloso é bastante poderoso e funciona bem para uma ampla faixa de problemas, porém nem sempre produz soluções ótimas.

No entanto, nem todos os problemas de otimização restrita podem ser resolvidos usando técnicas de matroides. É necessário verificar se a estrutura de restrições do problema forma um matroide antes de aplicar esse método.

#### 4. Materiais e Métodos

Em geral, as etapas realizadas nesse trabalho foram as seguintes:

1. Formulação do problema: Entendimento do problema, definição da função objetivo, bem como as restrições que limitam as possíveis soluções. Escrita das restrições em forma matemática, utilizando variáveis e parâmetros de decisão. O modelo proposto no trabalho de [5] foi utilizado como base para a criação do modelo de escalonamento e a modelo descrito em [9] foi a referência utilizada na criação do modelo para estimar o risco do paciente vir a desenvolver câncer cervical.
2. Seleção da técnica de otimização mais adequada para resolução do problema. Nesse caso, foram escolhidas programação inteira para os modelos na primeira parte desse trabalho.
3. Resolução do problema: nessa etapa o modelo matemático foi traduzido para linguagem matemática, o software de resolução foi escolhido e os parâmetros foram definidos. Com isso, modelo foi executado para encontrar a solução ótima do problema.
4. Avaliação da solução: por fim, foi verificado se a solução encontrada atende às restrições e objetivos do problema. Nessa etapa, alguns parâmetros e variáveis foram ajustados e o processo de resolução foi repetido.

##### 4.1. Modelo para acompanhamento preventivo de pacientes baseado em fatores de risco

Como citado nas seções anteriores, o câncer cervical é um tipo de câncer que tem altas chances de cura se detectado previamente através de exames como o exame do Papanicolaou e o exame HPV. Esses exames são limitados, então se faz necessário priorizar os casos mais urgentes para a execução deste exame. O objetivo deste modelo é estimar o grau de risco de uma mulher vir a desenvolver câncer cervical baseado nos fatores de risco e propor uma ordem de prioridade baseado nessa informação.

Seja  $M$  o conjunto de mulheres de uma localidade e  $H$  o conjunto de hospitais que realizam esse exame nessa área. Seja  $R$  o conjunto de todos os fatores de risco associados ao câncer cervical. Cada mulher  $i$ , terá um conjunto de fatores de risco associados a ela, denotado por  $R_i$ ,  $R_i \subseteq R$ . Cada  $r \in R$  possui um peso associado,  $p_r \in \mathbb{Z}^+$ , que é o peso daquele fator de risco no desenvolvimento do câncer cervical.

Assumindo que existem  $m$  mulheres para fazer o exame,  $M = \{1, 2, 3, \dots, m\}$  e  $n$  hospitais,  $H = \{1, 2, 3, \dots, n\}$ , onde cada mulher pode ser atendida pela sede e por outros

hospitais de preferência da sua região, o desafio é encontrar um subconjunto de mulheres  $S \subseteq M$ , que devem ser escolhidas primeiro para executar o teste em cada hospital.

Este problema pode ser formulado como um problema de otimização, no qual temos:

**Variáveis de decisão:**

$$x_{ij} = \begin{cases} 1 & \text{se a mulher } i \text{ é selecionada para realizar o exame no hospital } j; \\ 0 & \text{caso contrário. } \forall i \in M, \forall j \in H \end{cases}$$

**Função objetivo:**

$$\max \sum_{i=1}^m \sum_{j=1}^n r_i x_{ij} \quad (3)$$

onde  $r_i$  representa o risco da mulher  $i$ .

**Restrições:**

- *Limite de testes disponíveis nos hospitais:* o número de mulheres atendidas não ultrapassa o limite de testes de cada hospital. Desta forma, temos:

$$\sum_{i=1}^m D_{ij} x_{ij} \leq C_j, \quad j = 1, 2, \dots, n. \quad (4)$$

onde  $C_j$  representa a capacidade do hospital  $j$  e  $D$  é a matriz que relaciona mulheres a hospitais.

- *Limite de testes por paciente:* cada mulher só pode fazer o teste uma vez, ou seja, só pode ser atendida em um hospital. Desta forma, temos:

$$\sum_{j=1}^n x_{ij} \leq 1, \quad i = 1, 2, \dots, m. \quad (5)$$

**Cálculo do risco:**

Para o cálculo do risco de cada paciente, foram considerados os fatores selecionados em [8], que são eles: idade, idade da primeira relação sexual, histórico familiar, presença de infecções sexualmente transmissíveis (ISTs), tabagismo, quantidade de parceiros sexuais no último ano e se a paciente tomou a vacina contra o HPV. No estudo citado, é explicada a correlação entre os fatores de risco e o desenvolvimento do câncer. Para fins desse trabalho, foi considerada a lógica apresentada na Tabela 1 para definição do peso dos fatores de risco:

Assim é possível determinar o peso de cada fator de risco, conforme apresentado nas Tabelas 2-8.

**Tabela 1. Definição dos pesos.**

| Contribuição do fator para o risco | Peso |
|------------------------------------|------|
| Pouco impacto                      | 10   |
| Médio impacto                      | 50   |
| Alto impacto                       | 100  |

**Tabela 2. Fator de risco: Idade.**

| Idade              | Classificação | Risco |
|--------------------|---------------|-------|
| Entre 0 e 11 anos  | Criança       | 10    |
| Entre 12 e 18 anos | Adolescente   | 10    |
| Entre 19 e 29 anos | Adulto jovem  | 50    |
| Entre 30 e 59 anos | Adulto        | 100   |
| Maior que 60 anos  | Idoso         | 100   |

**Tabela 3. Fator de risco: Idade da primeira relação sexual.**

| Idade da 1ª relação sexual | Classificação | Peso |
|----------------------------|---------------|------|
| Entre 0 e 11 anos          | Criança       | 100  |
| Entre 12 e 18 anos         | Adolescente   | 100  |
| Maior que 19 anos          | Adulto        | 50   |

**Tabela 4. Fator de risco: Histórico de câncer na família.**

| Histórico de câncer na família | Classificação | Peso |
|--------------------------------|---------------|------|
| Entre 0 a 1 cânceres           | Fraco         | 10   |
| Entre 2 a 4 cânceres           | Médio         | 50   |
| Maior que 5 cânceres           | Forte         | 100  |

**Tabela 5. Fator de risco: Número de ISTs.**

| Número de ISTs   | Classificação | Peso |
|------------------|---------------|------|
| Entre 0 e 1 ISTs | Fraco         | 10   |
| 2 ISTs           | Médio         | 50   |
| Maior que 3 ISTs | Forte         | 100  |

**Tabela 6. Fator de risco: Tabagismo.**

| Fumante | Peso |
|---------|------|
| Sim     | 50   |
| Não     | 10   |

**Tabela 7. Fator de risco: Quantidade de parceiros sexuais no último ano.**

| Quantidade de parceiros sexuais | Classificação | Peso |
|---------------------------------|---------------|------|
| Entre 0 e 5 parceiros           | Poucos        | 10   |
| Entre 6 e 15 parceiros          | Moderado      | 50   |
| Maior que 16 parceiros          | Muitos        | 100  |

**Tabela 8. Fator de risco: Vacinação HPV.**

| Vacinação HPV | Peso |
|---------------|------|
| Sim           | 10   |
| Não           | 50   |

O risco associado a um paciente  $i$ ,  $r_i$ , é calculado como a soma dos fatores de risco que essa pessoa apresenta, baseados em [8]:

1. Paciente 1: (idade é adulto) e (idade da primeira relação é adulto) e (presença de ISTs é forte) e (tabagismo é não) e (histórico familiar é forte) e (parceiros é moderado) e (vacina contra hpv é não)
2. Paciente 2: (idade é adulto jovem) e (idade da primeira relação é criança/adolesc) e (presença de ISTs é medio) e (tabagismo é não) e (histórico familiar é fraco) e (parceiros é pouco) e (vacina contra hpv é não)

O cálculo de  $r_1$  e  $r_2$  pode ser visto nas Tabelas 9 e 10. Para facilitar a leitura, usaremos a seguinte legenda: I: idade; IRS: idade da primeira relação sexual, H: histórico familiar; ISTs: presença de infecções sexualmente transmissíveis, T: tabagismo;QP: quantidade de parceiros sexuais no último ano; V: vacina contra HPV.

**Tabela 9. Exemplo do cálculo do risco para o paciente 1.**

| Fator de risco | Classificação | Peso |
|----------------|---------------|------|
| I              | adulto        | 100  |
| IRS            | adulto        | 50   |
| IST            | forte         | 100  |
| T              | não fumante   | 10   |
| H              | forte         | 100  |
| QP             | moderado      | 50   |
| V              | não vacinado  | 100  |
| Total          |               | 510  |

**Tabela 10. Exemplo do cálculo do risco para o paciente 2.**

| Fator de risco | Classificação | Peso |
|----------------|---------------|------|
| I              | adulto jovem  | 50   |
| IRS            | adulto        | 50   |
| IST            | fraco         | 10   |
| T              | fumante       | 100  |
| H              | médio         | 50   |
| QP             | muitos        | 100  |
| V              | vacinado      | 10   |
| Total          |               | 370  |

Neste exemplo, o paciente 1 tem um risco mais alto do que o paciente 2, logo deve ter a prioridade para a realização do exame.

## 4.2. Modelo para agendamento de pacientes em tratamento oncológico

O objetivo do segundo modelo de otimização escolhido é melhorar a eficiência do agendamento dos pacientes a fim de minimizar os tempos de inatividade da escala, dando assim resposta a um maior número de pacientes [5]. Para fins de aplicação deste modelo, são consideradas as informações a seguir:

1. O atendimento dos pacientes funciona 8h por dia apenas em dias úteis e tem duração de 15 minutos.
2. Cada paciente faz um único atendimento por dia, durante aproximadamente quinze dias úteis em média. Esse período pode variar dependendo da patologia.
3. O tratamento alvo deste estudo é o da radiologia externa, onde assume-se que é possível atender apenas um paciente por vez em cada turno devido à limitação de máquinas disponíveis (acelerador linear).
4. Uma vez agendado um paciente para um turno, este não é modificado até o término do tratamento.
5. É sabido a lista de novos pacientes e pacientes em tratamento ao iniciar o escalonamento, bem como os turnos ocupados e disponíveis.
6. Pacientes com tratamentos semelhantes são priorizados para estarem em turnos próximos com o objetivo de reduzir o tempo de preparação do equipamento.
7. O tempo de duração do tratamento (15min) inclui o tempo de preparação das salas entre um paciente e outro e não inclui a consulta inicial com o médico Radio-Oncologista, que deve ser feita antes do início do tratamento.

Para a formulação do problema em [5], foram definidos (retirados integralmente do estudo original):

### Conjuntos ou entradas:

$K$  = tratamentos possíveis;

$P$  = pacientes a iniciar tratamento;

$T_d$  = turnos com vaga disponível no dia  $d$ .

### Índices:

$k$  – tipo de tratamento ( $k \in K$ );

$p$  – paciente a iniciar tratamento ( $p \in P$ );

$d$  – dia da semana ( $d = 1, \dots, 5$ );

$T_d$  – turno do dia  $d$  ( $t_d \in T_d$ ).

### Parâmetros:

$k_p$  – tipo de tratamento do paciente  $p$  ( $k_p \in K$ );

$d_p$  – dia que o paciente  $p \in P$  tem que iniciar o tratamento ( $d_p = 1, \dots, 5$ );

$c_p^{t_d}$  – penalização (ou custo) de inserir o paciente  $p$ , com tratamento  $k$ , no turno  $t_d$  do dia  $d$ . Esta penalização vai depender do tratamento que antecede e sucede uma dada vaga disponível e o tratamento do novo paciente a ser atribuído no dia a que este inicia o tratamento,  $d_p$ . Contudo, mantém-se constante para os dias seguintes ao de início de tratamento, ou seja,  $c_p^{t_d} = c_p^{t_d d_p}$ ,  $d > d_p$ .



**Variáveis de decisão:**

Para cada paciente,  $p \in P$ , a iniciar tratamento  $k_p$  no dia  $d_p$ :

$$x_p^{t_d d} = \begin{cases} 1 & \text{se o paciente } p \text{ é atribuído ao turno } t_d \text{ do dia } d, \forall d \geq d_p; \\ 0 & \text{caso contrário.} \end{cases}$$

**Função objetivo proposta:**

$$\min \sum_{p \in P} \sum_{d=d_p}^5 \sum_{t_d \in T_d} c_p^{t_d d} x_p^{t_d d} \quad (6)$$

O objetivo é minimizar o custo total de alocação de pacientes, encontrando a melhor alocação de pacientes para cada turno, considerando a disponibilidade dos pacientes e a capacidade de cada turno.

**Restrições:**

$$\text{s.a} = \begin{cases} \sum_{p \in P} x_p^{t_d d} \leq 1 & t_d \in T_d; d = 1, \dots, 5; \\ \sum_{t_d \in T_{d_p}} x_p^{t_d d_p} = 1 & p \in P; \\ x_p^{t_d d} \leq x_p^{t_d d+1} & p \in P; t_d \in T_d; d = d_p, \dots, 4; \\ x_p^{t_d d} \in \{1, 0\} & p \in P; t_d \in T_d; d_p \leq d \leq 5. \end{cases}$$

Explicando as restrições, na primeira equação das restrições, garante-se que a cada turno de cada dia é atribuído um doente no máximo. Na segunda, garante-se que cada doente,  $p$ , é atribuído a um turno no dia pretendido,  $d_p$ . Nas restrições da terceira linha, impõe-se que cada paciente,  $p$ , é atribuído sempre ao mesmo turno,  $t_d$  a que for atribuído no dia em que inicia o tratamento,  $d_p$ . Na linha final, define-se o domínio das variáveis.

**Penalização ou custo:**

A penalização ou custo propostos no modelo do estudo [5] são definidos na Tabela 11:

**Tabela 11. Definição de penalizações. Retirado de [5].**

| Penalização | Sequência entre tratamentos  |
|-------------|--|
| C           | Se o novo paciente é atribuído após um paciente com o mesmo tipo de tratamento no escalonamento inicial, $k_p$ |
| 3C          | Se o novo paciente é atribuído após uma vaga, mas próximo de um paciente com o mesmo tipo de tratamento, $k_p$ |
| 4C          | Se o novo paciente é atribuído após um paciente com o tipo de tratamento diferente de $k_p$ .                  |

## 5. Resultados (Estudo de Caso)

O objetivo inicial da pesquisa era utilizar dados reais do Hospital das Clínicas localizado em Recife - PE, porém, devido à necessidade de aprovação do comitê de ética e demais trâmites para liberação de dados, não houve tempo hábil para sua realização. É comum que estudos que envolvam dados de pacientes passem por uma série de processos burocráticos a fim de garantir a segurança e privacidade dos mesmos. Apesar de não ter sido possível utilizar esses dados, a pesquisa seguiu com outras fontes e métodos para alcançar seus objetivos.

Nesta seção é ilustrada a aplicação dos conceitos teóricos discutidos na Seção 4. Os modelos matemáticos são traduzidos para a linguagem matemática AMPL e são finalmente testados com a utilização de um conjunto de dados fictícios criados a partir das regras definidas na Seção 4.

Todos os modelos foram testados em um Macbook Pro 2017, com o processador 3,1 GHz Dual-Core Intel Core i5 e sistema operacional macOS Ventura 13.2.1. Todos os arquivos contendo o código do modelo, os dados e os comandos para execução estão disponíveis em repositório online cujo o link pode ser encontrado em [15].

### 5.1. Resultados do acompanhamento preventivo de pacientes baseado em fatores de risco

Para os experimentos com o modelo de priorização de mulheres para execução de exames citopatológico com base nos fatores de risco, foram consideradas as seguintes informações:

1. Devido a limitações da versão gratuita do software AMPL IDE, foi utilizado um conjunto reduzido de dados, limitado a 100 pacientes.
2. Foi escolhida a cidade de Ipojuca na região metropolitana de Recife-PE como inspiração para o problema, que possui cinco unidades de saúde que fazem a coleta de material do colo de útero para exame citopatológico. Devido ao uso de um conjunto de dados reduzido, a capacidade de cada unidade de saúde também foi reduzida. A capacidade hipotética de cada unidade de saúde está descrita na Tabela 12.
3. Cada paciente foi aleatoriamente associado a uma ou múltiplas unidades de saúde de sua região. Na Tabela 13 pode ser visto um exemplo dessa associação, onde o valor da coluna é "1" se o paciente pode ser atendido por aquele hospital e "0" se não pode.
4. O risco dos pacientes foi definido aleatoriamente utilizando as regras de peso apresentadas na Seção 4.1.

Na Figura 4, é possível visualizar código em AMPL utilizado para testar o modelo:

Por fim, os pacientes selecionados para realizar o exame dado o risco associado a eles pode ser visto na Tabela 14:

**Tabela 12. Capacidade das unidades de saúde para coleta de material do colo de útero para exame citopatológico.**

| Unidade de saúde | Capacidade de exames |
|------------------|----------------------|
| 1                | 34                   |
| 2                | 10                   |
| 3                | 13                   |
| 4                | 5                    |
| 5                | 5                    |

**Tabela 13. Atribuição de pacientes a hospitais/unidades de saúde.**

| Pacientes | Unid.1 | Unid.2 | Unid. 3 | Unid. 4 | Unid. 5 |
|-----------|--------|--------|---------|---------|---------|
| 1         | 0      | 0      | 0       | 1       | 1       |
| 2         | 0      | 1      | 0       | 0       | 0       |
| 3         | 1      | 1      | 0       | 1       | 1       |
| 4         | 0      | 1      | 0       | 0       | 0       |
| 5         | 0      | 0      | 0       | 1       | 0       |
| 6         | 0      | 0      | 1       | 0       | 0       |
| 7         | 1      | 0      | 0       | 1       | 0       |
| 8         | 0      | 0      | 0       | 1       | 0       |
| 9         | 0      | 1      | 0       | 1       | 0       |
| 10...     | 1      | 0      | 0       | 0       | 0       |

```

# Sets
set patients := 1..100;
set hospitals := 1..5;

# Parameters
param r{i in patients};                #risk of each patient
param capacity{j in hospitals};        #capacity of each hospital
param base >= 1;
param preferable_hospital {patients, hospitals} >= 0; #hospitals preferred by the patient

# Variables
var x{i in patients, j in hospitals} binary; # indica se o paciente i é alocado no hospital j.

# Objective function
maximize Risk:
    sum{i in patients, j in hospitals} base^preferable_hospital[i,j] * x[i,j] * r[i];

# Constraints
# a patient can be tested in only one hospital
subject to one_exam_per_patient{i in patients}:
    sum{j in hospitals} x[i,j] <= 1;

# maximum no. of tests per hospital
subject to capacity_hospitals{j in hospitals}:
    sum{i in patients} x[i,j] <= capacity[j];

```

**Figura 4. Modelo escrito em AMPL. Arquivo .mod.**

O tempo de execução do código acima foi 0.173982 segundos.

**Tabela 14. Pacientes com prioridade para a realização dos exames nas unidades de saúde.**

| Unidade de saúde | Pacientes selecionados  |
|------------------|---|
| 1                | 3, 10, 12, 18, 19, 20, 21, 25, 28, 29, 34, 35, 41, 42, 47, 50, 59, 60, 62, 67, 70, 72, 78, 81, 83, 85, 87, 89, 91, 92, 95, 97, 98, 99 |
| 2                | 4, 14, 17, 22, 39, 49, 55, 66, 69, 75   |
| 3                | 6, 16, 24, 32, 51, 53, 54, 68, 79, 88   |
| 4                | 7, 9, 26, 52, 86  |
| 5                | 36, 37, 46, 48, 74  |

## 5.2. Resultados do agendamento de pacientes oncológicos

Para os testes com o modelo, foram consideradas as seguintes informações:

1. Cada paciente faz um único atendimento por dia.
2. É possível atender apenas 4 pacientes por vez em cada turno devido à limitação de máquinas disponíveis.
3. O custo da alocação de pacientes com tratamentos semelhantes é menor do que a de tratamentos distintos devido ao tempo de preparação do equipamento. Para fins dessa atividade, os custos foram definidos aleatoriamente.
4. O custo de um turno sem alocação de pacientes é de 500 unidades.
5. Devido a limitações da versão demo do software AMPL IDE, foram utilizados um conjunto reduzido de dados, limitado a 45 pacientes.

Na Figura 5, é apresentado o modelo descrito na linguagem AMPL:

```
#Sets
set patient:= 1..45; #no. of patients to be allocated
set days:= 1..5; #planning horizon
set shift := 1..2; #no. of shift in a day - morning and afternoon
set S := 1..2; #shift with no patients allocated
set D := 1..5; #day corresponding to have no patients allocated
set N:= 1..5;

#Parameters
param availability{i in patient, j in days};
param costpershift{i in patient, j in days, s in shift};
param emptycost{n in N, l in D, m in S};
var emptyshift{n in N, l in D, m in S} integer;

#Decision variable
var patientassignment{i in patient, j in days, s in shift} binary;
# = 1 if patient i is available on jth day working on sth shift, 0 otherwise

#Objective function
minimize Cost:
sum{i in patient, j in days, s in shift} costpershift[i,j,s]*patientassignment[i,j,s]
+ sum{ n in N, l in D, m in S}emptycost[n,l,m]*emptyshift[n,l,m];

#constraints

#maximum no. of shifts per day
subject to maximum_shifts_perday {i in patient, j in days}:
sum{s in shift} patientassignment[i,j,s]*availability[i,j] <= 1;

#patients per shift
subject to schedule{j in days, s in shift, l in D, m in S}:
sum{i in patient} availability[i,j]*patientassignment[i,j,s] + sum{n in N} emptyshift[n,l,m]=4;
```

**Figura 5. Modelo escrito em AMPL. Arquivo .mod**

Ao rodar o modelo com os dados, o resultado é mostrado na Tabela 15 :

**Tabela 15. Pacientes agendados para tratamentos oncológicos com base em sua prioridade.**

|       | Segunda     | Terça       | Quarta   | Quinta      | Sexta       |
|-------|-------------|-------------|--|-------------|-------------|
| Manhã | Paciente 9  | Paciente 18 | Paciente 2<br>Paciente 3<br>Paciente 5<br>Paciente 9 | Paciente 18 | Paciente 23 |
|       | Paciente 31 | Paciente 19 |  | Paciente 19 | Paciente 24 |
|       | Paciente 33 | Paciente 20 |  | Paciente 20 | Paciente 28 |
|       | Paciente 39 | Paciente 41 |  | Paciente 41 | Paciente 29 |
|       | Paciente 40 | Paciente 45 |  | Paciente 45 | Paciente 30 |
| Tarde | Paciente 8  | Paciente 12 | Paciente 8   | Paciente 12 | Paciente 21 |
|       | Paciente 10 | Paciente 13 | Paciente 10  | Paciente 13 | Paciente 22 |
|       | Paciente 32 | Paciente 42 | Paciente 32  | Paciente 42 | Paciente 25 |
|       | Paciente 34 | Paciente 43 | Paciente 34  | Paciente 43 | Paciente 26 |
|       | Paciente 35 | Paciente 44 | Paciente 35  | Paciente 44 | Paciente 27 |

O tempo de execução do código acima foi 0.182588 segundos.

## 6. Parte 2: Matróides

Nesta seção apresenta-se a modelagem do problema utilizando matróides. Como foi citado na Seção 3, nem todos os problemas podem ser resolvidos com essa técnica. Portanto, o primeiro passo será provar que esse problema tem uma estrutura de matróides.

### 6.1. Modelagem do problema

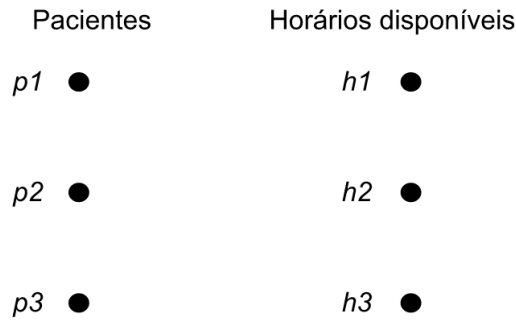
Para fins de modelagem do problema, são consideradas as seguintes informações:

- A duração dos atendimentos é de 15 minutos.
- Cada paciente faz um único atendimento por dia.
- É possível atender apenas um paciente por vez em cada turno devido à limitação de máquinas disponíveis.
- O custo da alocação de pacientes com tratamentos semelhantes é menor do que a de tratamentos distintos devido ao tempo de preparação do equipamento.

Os seguintes conjuntos são entradas do problema:

- Um conjunto  $P$  finito e ordenado de pacientes a serem alocados;
- Um conjunto  $H$  finito de horários disponíveis;
- Um conjunto de  $C$  finito de  $n$  pesos não negativos ou custos de alocar os pacientes nos horários disponíveis.

Considere que existe uma função  $\Phi$  de  $P$  para subconjuntos do conjunto  $H$ . Segundo Gale [16], um subconjunto  $A$  de  $P$  é atribuível se existe um função injetora  $\phi$  de  $A$  para  $H$  tal que  $\phi(a) \in \Phi(a)$  para todo  $a$  em  $A$ . Ele prova em [16] que uma família de conjuntos atribuíveis contém um elemento ótimo e também prova que essa família é uma matróide. Portanto, se modelado o problema de alocação de pacientes como uma família de conjuntos atribuíveis, estes formarão uma matroide e conseguiremos definir um conjunto atribuível ótimo  $\bar{A}$  como solução do problema.



**Figura 6. Exemplo ilustrativo de uma atribuição.**

Se tomarmos por exemplo uma lista de 3 pacientes e 3 horários disponíveis:

Se levarmos em conta a restrição do problema, onde apenas um paciente é alocado em cada horário disponível, garantimos que essa função de atribuição partindo de P para H é injetora, logo a família F abaixo pode ser caracterizada como uma família de conjuntos atribuíveis de P, onde uma tupla é formada pelo paciente e seu respectivo horário alocado:  $F = \{ \{\}, \{(1,1)\}, \{(1,2)\}, \{(1,3)\}, \{(2,1)\}, \{(2,2)\}, \{(2,3)\}, \{(3,1)\}, \{(3,2)\}, \{(3,3)\}, \{(1,1), (2,2)\}, \{(1,1), (3,3)\}, \{(1,1), (2,3)\}, \{(1,1), (3,2)\}, \{(2,2), (3,3)\}, \{(2,2), (3,1)\}, \{(2,2), (1,3)\}, \{(3,3), (2,1)\}, \{(3,3), (1,2)\}, \{(1,2), (2,1)\}, \dots \}$

Como foi provado por Gale em [16], uma família F de conjuntos atribuíveis é uma matróide, logo seguimos para a solução do nosso problema de encontrar um conjunto ótimo.

Podemos implementar essa estrutura de matróide da seguinte forma:

1. Criamos um conjunto de pacientes;
2. Ordenamos o conjunto de pacientes pelo custo em ordem crescente;
3. Inicializamos um conjunto vazio de pacientes agendados;
4. Percorremos os pacientes em ordem crescente de custo:
  - (a) Verificamos se o paciente já foi alocado;
  - (b) Se sim, ignoramos o paciente e partimos para o próximo;
  - (c) Se não, verificamos se o paciente pode ser adicionado ao conjunto de pacientes agendados sem criar conflitos de horário:
    - Se puder, adicionamos o paciente ao conjunto de pacientes agendados;
    - Se não puder, ignoramos o paciente;
5. Retornamos o conjunto de pacientes agendados e o custo total.

O código implementado em Python pode ser visto na Figura 7:

A função *patientScheduling* recebe como parâmetro uma lista, onde cada item é uma tupla com três elementos: o ID do paciente, o horário desejado para a consulta e o custo associado à alocação desse paciente. Os horários disponíveis são representados como uma lista de inteiros que indicam os horários disponíveis.

O algoritmo começa definindo duas funções auxiliares: *isIndependent(S)* e *isAllocated(P)*. A primeira função verifica se um conjunto de pacientes S é independente, ou seja, se não há conflitos de horários entre eles. A segunda função verifica se um paciente P já foi alocado.

```

1 import timeit
2
3 def patientScheduling(patients):
4     # Define a função de independência
5     def isIndependent(S):
6         slotsUsed = set()
7         for patient in S:
8             if patient[1] in slotsUsed:
9                 return False
10            slotsUsed.add(patient[1])
11        return True
12
13    # Checa se o paciente já foi alocado
14    def isAllocated(P):
15        patientsAllocated = set()
16        for patient in P:
17            if patient[0] in patientsAllocated:
18                return False
19            patientsAllocated.add(patient[0])
20        return True
21
22    # Inicializa a lista de pacientes alocados e o custo total
23    assignment = []
24    allocated_patients = []
25    totalCost = 0
26
27    # Executa o algoritmo guloso para matroides
28    for patient in sorted(patients, key=lambda x: x[2]):
29        if isAllocated(allocated_patients + [patient]):
30            if isIndependent(assignment + [patient]):
31                assignment.append(patient)
32                allocated_patients.append(patient)
33                totalCost += patient[2]
34
35    # Imprime a lista de pacientes alocados e o custo total
36    for i in range(len(assignment)):
37        print("O paciente {} sera alocado no horario {}".format(
38            assignment[i][0], assignment[i][1]))
39    print("Custo total: {}".format(totalCost))
40

```

Figura 7. Código em Python.

Em seguida, a função *patientScheduling* inicializa a lista *assignment* de pacientes alocados, a lista *allocated\_patients* de pacientes já alocados e o custo total *totalCost*.

O algoritmo guloso para matroides é então executado em um loop for que itera sobre a lista de pacientes ordenados por custo em ordem crescente. Para cada paciente, é verificado se ele já foi alocado e se é independente dos pacientes já alocados. Se essas condições forem satisfeitas, o paciente é adicionado à lista de pacientes alocados *assignment*, à lista de pacientes já alocados *allocated\_patients* e o custo associado a essa alocação é adicionado ao custo total *totalCost*.

Por fim, a função imprime a lista de pacientes alocados e o custo total. Cada paciente alocado é exibido junto com o horário em que será atendido.

## 6.2. Resultados (Estudo de Caso)

Para os testes com o modelo, foram consideradas as seguintes informações:

1. A lista de pacientes a ser alocados é composta por 100 pacientes;
2. A cada semana, 10 turnos estão disponíveis para alocação, sendo  $t_1$  correspondendo ao turno de segunda-feira pela manhã,  $t_2$  segunda-feira pela tarde e assim sucessivamente.
3. Para cada paciente, foram definidos aleatoriamente os custos de alocação para cada turno.
4. É possível atender apenas 1 paciente por vez em cada turno devido à limitação de máquinas disponíveis.
5. Cada paciente deve ser alocado a apenas 1 turno por semana.

Na Figura 8, é possível visualizar o resultado do modelo executado em Python:

```
/usr/local/bin/python3.9 /Users/estefane/PycharmProjects/matroides/patients_scheduling_matroids.py
0 paciente 50 sera alocado no horario 10
0 paciente 94 sera alocado no horario 8
0 paciente 1 sera alocado no horario 1
0 paciente 5 sera alocado no horario 6
0 paciente 8 sera alocado no horario 5
0 paciente 9 sera alocado no horario 9
0 paciente 10 sera alocado no horario 3
0 paciente 11 sera alocado no horario 2
0 paciente 12 sera alocado no horario 7
0 paciente 22 sera alocado no horario 4
Custo total: 89
Execution time: 0.003265

Process finished with exit code 0
```

Figura 8. Resultado do teste do modelo em Python.

O tempo de execução do código acima foi 0.003265 segundos.

Acerca da eficiência do algoritmo, o tempo de execução deste depende principalmente da complexidade das duas funções auxiliares  $isIndependent(S)$  e  $isAllocated(P)$ .

A função  $isIndependent(S)$  verifica se um conjunto de pacientes é independente, isto é, se cada paciente tem um horário de atendimento diferente. Para cada paciente em  $S$ , a função percorre a lista de pacientes já alocados e verifica se o horário de atendimento do paciente já foi utilizado. Como o número de pacientes já alocados pode ser no máximo  $len(S) - 1$ , o tempo de execução da função é de ordem  $O(n)$ , onde  $n$  é o número de pacientes.

A função  $isAllocated(P)$  verifica se um conjunto de pacientes já foi alocado. Para cada paciente em  $P$ , a função percorre a lista de pacientes já alocados e verifica se o paciente já está lá. Como o número de pacientes já alocados pode ser no máximo  $len(P) - 1$ , o tempo de execução da função também é de ordem  $O(n)$ .

A complexidade total do algoritmo é  $O(n^2 \log n)$ , onde  $n$  é o número de pacientes. Isso ocorre porque a função principal executa um loop sobre os pacientes, que é  $O(n)$ , e, em seguida, chama as funções auxiliares, que são  $O(n)$  cada uma. Além disso, o algoritmo usa a função  $sorted()$  - função interna do Python que cria uma nova lista classificada a



partir de um iterável - para classificar os pacientes com base em seus horários de chegada, o que leva  $O(n \log n)$  tempo. Portanto, a complexidade total é  $O(n^2 \log n)$ .

Em resumo, a análise do tempo de execução do algoritmo usando a notação Big O mostra que ele é relativamente eficiente para um número pequeno de pacientes, mas pode se tornar lento à medida que o número de pacientes aumenta.

## 7. Discussões e Trabalhos futuros

Por meio da aplicação da otimização restrita, foi possível construir dois modelos que podem beneficiar profissionais e instituições de saúde que lidam com o câncer: um para selecionar os pacientes com o maior risco de desenvolver câncer de colo de útero para realização de exames preventivos e outro para otimizar o agendamento de pacientes oncológicos. Para simplificar o primeiro modelo, foi selecionado um subconjunto de 07 fatores para estimar o risco de 100 pacientes. O modelo desenvolvido é um protótipo inicial que pode ser aprimorado no futuro. Do ponto de vista computacional, é possível melhorar os modelo matemático utilizando distâncias reais das unidades de saúde como base para construção da matriz D que relaciona paciente a hospitais. Além disso, seria ideal integrar o modelo com um software para pudesse ser utilizado diretamente por profissionais de saúde e médicos no dia a dia da unidade de saúde para ajudar na priorização.

Em relação ao segundo modelo, foram considerados 45 pacientes apenas devido a limitação da licença demo do software e um limite de 4 pessoas por turno. No futuro, pode-se melhorar esse modelo para considerar um maior número de pacientes e horários para agendamento, além da própria otimização do programa escrito. Infelizmente, ambos os modelos não foram testados com dados reais de pacientes, o que não invalida a pesquisa, mas impossibilitou o teste in loco. No futuro, seria interessante a validação desses modelos em hospitais e unidades de saúde.

A respeito da resolução do modelo utilizando matróides, trabalhos futuros relacionados a esse texto poderiam explorar outras técnicas matemáticas além de matróides para resolver o problema de alocação de pacientes. Além disso, seria interessante investigar como a modelagem do problema pode ser estendida para lidar com outras restrições, como a disponibilidade de recursos médicos específicos para cada paciente ou a preferência do paciente por horários específicos. Outra área de pesquisa relacionada seria a análise da eficácia do algoritmo proposto em casos mais complexos, com um grande número de pacientes e horários disponíveis, por exemplo.

## Referências

- [1] FERLAY, J.; COLOMBET, M.; SOERJOMATARAM, I.; PARKIN, D. M.; PIÑEROS, M.; ZNAOR, A.; BRAY, F. Global cancer observatory: Cancer tomorrow. Technical report, International Agency for Research on Cancer, 2020. Disponível em <https://gco.iarc.fr/tomorrow>, acessado em 10 de Fevereiro de 2023.
- [2] Estimativa 2023: incidência de câncer no brasil. Technical report, Instituto Nacional de Câncer (Brasil) - INCA, 2022. Disponível em <https://www.inca.gov.br/publicacoes/livros/estimativa-2023-incidencia-de-cancer-no-brasil>. Acesso em: 10 de Fevereiro de 2023.
- [3] Who report on cancer: setting priorities, investing wisely and providing care for all. Technical report, World Health Organization, 2020.

- [4] CONFORTI, D.; GUERRIERO, F.; GUIDO, R. Optimization models for radiotherapy patient scheduling. *4OR*, v. 6, p. 263–278, 2008. <https://doi.org/10.1007/s10288-007-0050-8>. Acesso em 10 de Fevereiro de 2023.
- [5] OLIVEIRA, M. M. D. F. R. D. *Escalonamento de pacientes de radioterapia*. 2015. Dissertação (Mestrado em Física) - Instituto Superior de Economia e Gestão, Lisboa, 2015. Disponível em: <http://hdl.handle.net/10400.5/11164>. Acesso em 10 de Fevereiro de 2023.
- [6] AFONSO, T. M. *Escalonamento de pacientes num serviço hospitalar de radioterapia*. 2017. Dissertação (Mestrado em Física) - Instituto Superior de Economia e Gestão, 2017. Disponível em: <http://hdl.handle.net/10400.5/15597>. Acesso em 10 de Fevereiro de 2023.
- [7] BURKE, E.; LEITE-ROCHA, P.; PETROVIC, S. An integer linear programming model for the radiotherapy treatment scheduling problem. 03 2011. Disponível em: <https://doi.org/10.48550/arXiv.1103.3391>. Acesso em 10 de Fevereiro de 2023.
- [8] NETO, M. A. D. F. Um modelo fuzzy para estimar o risco de desenvolver câncer de colo de Útero. Disponível em: [https://www.dropbox.com/s/v27fs1lowy9wpfv/TCC\\_Manoel\\_v06.pdf?dl=0](https://www.dropbox.com/s/v27fs1lowy9wpfv/TCC_Manoel_v06.pdf?dl=0). Acesso em: 01 de Fevereiro de 2023, 2021.
- [9] MOURA, M. A. Algoritmos de otimização para solução do problema do próximo release. Disponível em: <http://200.17.137.109:8081/novobsi/bsi-na-ufrpe-recife/trabalhos-de-conclusao-de-curso/2014.2/Mariana.pdf>. Acesso em: 01 de Fevereiro de 2023, 2015.
- [10] CROWN, W.; BUYUKKARAMIKLI, N.; THOKALA, P.; MORTON, A.; SIR, M. Y.; MARSHALL, D. A.; TOSH, J.; PADULA, W. V.; IJZERMAN, M. J.; WONG, P. K.; PASUPATHY, K. S. Constrained optimization methods in health services research—an introduction: Report 1 of the ispor optimization methods emerging good practices task force. *Value in Health*, v. Volume 20, issue 3, p. P310–319, March 2017. Disponível em: <https://doi.org/10.1016/j.jval.2017.01.013>. Acesso em 10 de Fevereiro de 2023.
- [11] HONGWEI, J. *A tutorial of ampl for linear programming*. May 2014. Disponível em: [https://www.cs.uic.edu/hjin/files/ampl\\_tutorial.pdf](https://www.cs.uic.edu/hjin/files/ampl_tutorial.pdf). Acesso em: 01 de Fevereiro de 2023.
- [12] FOURER, R.; GAY, D. M.; KERNIGHAN, B. *Ampl: A modeling language for mathematical programming*. 2nd. ed. Duxbury Press, Pacific Grove, 2003. Disponível em: <https://vanderbei.princeton.edu/307/textbook/AMPLbook.pdf>. Acesso em 15 de Abril de 2023.
- [13] OXLEY, J. G. *Matroid theory*. Oxford University Press, 2006.
- [14] CORMEN, T. *Algoritmos: teoria e prática*. 3. ed. Campus, 2012.
- [15] OLIVEIRA, E. Repositório com o código criado, 2023. <https://github.com/stefsoliveira/tcc>.

- [16] GALE, D. Optimal assignments in an ordered set: An application of matroid theory. *Journal of Combinatorial Theory*, v. 4, p. 176–180, 1968. [https://doi.org/10.1016/S0021-9800\(68\)80039-0](https://doi.org/10.1016/S0021-9800(68)80039-0). Acesso em 10 de Fevereiro de 2023.