ALGORITMOS NSGA-II E SPEA2 PARA O PROBLEMA MULTIOBJETIVO DE LOCALIZAÇÃO DE MAMÓGRAFOS

Marcos Vinícius Andrade de Campos[1,2], Romário dos Santos Lopes de Assis[1,4], Marcone Jamilson Freitas Souza[1,3], Maria Amélia Lopes Silva[4], Eduardo Camargo de Siqueira[5], Sérgio Ricardo de Souza[1]

[1] Programa de Pós-Graduação em Modelagem Matemática e Computacional, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Av. Amazonas 7675, CEP 30.510-000, Belo Horizonte (MG), Brasil

[2]Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG), CEP 30.575-180, Belo Horizonte (MG), Brasil

[3]Departamento de Computação – Universidade Federal de Ouro Preto (UFOP), CEP 35.400-000, Ouro Preto (MG), Brasil

[4]Universidade Federal de Viçosa (UFV) – campus Florestal, CEP 35.690-000 Florestal (MG), Brasil

[5]Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro (IFTM), CEP 38.600-000, Paracatu (MG), Brasil

marcos.campos@ifmg.edu.br, romario@ufv.br, marcone@ufop.edu.br, mamelia@ufv.br, eduardosiqueira@iftm.edu.br, sergio@cefetmg.br

RESUMO

A cobertura de exames de mamografia no Brasil é afetada pela distribuição inadequada dos equipamentos de mamografia. Em vista disso, muitas mulheres não têm acesso ao exame ou se deslocam por grandes distâncias para obtê-lo. Este artigo aborda o Problema Multiobjetivo de Localização de Mamógrafos, em que, além da maximização da cobertura, tem-se também, como objetivo, a minimização dos deslocamentos. Para resolvê-lo, é utilizado o *MOEA Framework* e implementados dois algoritmos baseados nos algoritmos genéticos NSGA-II e SPEA2. Os resultados foram avaliados em relação às métricas Hipervolume e *Spacing*, e mostraram-se capazes de fornecer opções de qualidade ao tomador de decisão.

PALAVRAS CHAVE. Problema Multiobjetivo de Localização de Mamógrafos, NSGA-II, SPEA2.

Tópicos: PO na Área de Saúde

ABSTRACT

The coverage of mammography screenings in Brazil is affected by the inadequate distribution of mammography equipment. Given this, many women do not have access to the screening or must travel long distances to obtain it. This work addresses the Multi-objective Problem of Mammography Unit Location in which, in addition to maximizing coverage, the objective is also to minimize displacements. The MOEA Framework is used to solve it, and two algorithms, based on the NSGA-II and SPEA2 genetic algorithms, are implemented. The results were evaluated concerning the Hypervolume and Spacing metrics and proved capable of providing quality options to the decision-maker.

KEYWORDS. Multi-objective Mammography Unit Location Problem, NSGA-II, SPEA2.
Paper topics: OR in Health

https://proceedings.science/p/157495?lang=pt-br

1. Introdução

As chances de cura de um câncer de mama podem alcançar 95% caso o tumor seja detectado precocemente [Witten e Parker, 2018]. Para isso, é importante que as mulheres tenham acesso adequado ao exame de mamografia, sendo esse o principal método de detecção precoce de neoplasias malignas ligadas à mama [Xavier et al., 2016].

Trabalhos como os de Amaral et al. [2017] e Miranda e Patrocinio [2018] mostraram o quanto é ineficiente a distribuição dos equipamentos de mamografia no Brasil, prejudicando o acesso do público alvo ao exame. Em resumo, o país possui uma quantidade significativa de equipamentos, a qual possibilitaria uma cobertura mais abrangente do que a observada hoje, mas sua má distribuição limita o acesso de muitas mulheres ao exame. A partir desse diagnóstico, iniciaram-se os estudos do Problema de Localização de Mamógrafos (PLM), cujo objetivo é melhorar a cobertura da demanda por mamografias, estabelecendo, assim, uma distribuição mais adequada dos equipamentos.

O trabalho de Souza et al. [2019] tratou o PLM no Estado de Rondônia. Foram criadas duas formulações de programação matemática. Na primeira, a cobertura de cada cidade deve ser integral, isto é, uma cidade só pode ser atendida com 100% de cobertura. Na segunda formulação é permitido atendimento parcial. Em Souza et al. [2020], os autores trabalharam com um modelo que permite somente atendimento integral. Também foi apresentado um algoritmo heurístico baseado em *Variable Neighborhood Search* (VNS) [Mladenović e Hansen, 1997] para resolver o problema. As instâncias utilizadas foram baseadas em dados do Estado de Minas Gerais. Posteriormente, de Campos et al. [2020] focaram no modelo de atendimento parcial, em que parcelas da demanda de uma cidade j podem ser atendidas por mamógrafos de uma ou mais cidades diferentes. No referido trabalho, os autores utilizaram um algoritmo heurístico baseado em *Simulated Annealing* [Kirkpatrick et al., 1983] como alternativa de solução. Dados de Rondônia e de Minas Gerais foram utilizados para testes. O mais recente trabalho relacionado ao PLM, elaborado por de Assis et al. [2021], também focou no modelo de cobertura parcial e utilizou um algoritmo baseado na metaheurística *Iterated Greedy Search* [Ruiz e Stützle, 2008] como método de solução.

Todos os trabalhos supracitados tiveram como objetivo único a maximização da cobertura da demanda por exames de mamografia por meio de uma alocação eficiente dos mamógrafos. Porém, o deslocamento entre cidades pode ser custoso, cansativo e desestimulador. Como consequência, pode-se chegar a uma situação em que há mamógrafo disponível de acordo com os parâmetros estabelecidos, porém a cobertura real não chega aos níveis esperados durante o planejamento. Por isso, além de maximizar a cobertura, neste trabalho tratamos o Problema Multiobjetivo de Localização de Mamógrafos (PMLM), buscando, assim, também minimizar o deslocamento das mulheres em direção aos locais aonde estão instalados os equipamentos de mamografia.

A minimização do deslocamento eventualmente significará perda de cobertura, ao passo que uma solução ótima em relação à cobertura pode gerar deslocamentos maiores. Em situações como essas, em que temos dois ou mais objetivos conflitantes, temos um problema multiobjetivo, e neste tipo de problema não é esperada uma solução única, mas sim um conjunto de soluções não-dominadas. Esse conjunto de soluções formam então a chamada fronteira de Pareto.

O restante deste trabalho se organiza da seguinte forma. Na seção 2 o problema é descrito. Na Seção 3 é apresentada a metodologia utilizada, enquanto os resultados dos experimentos computacionais são apresentados na Seção 4. Por fim, é feita uma discussão final na Seção 5.

2. Descrição do Problema

O PMLM possui características similares à sua versão mono-objetivo apresentadas nos trabalhos de Souza et al. [2019], Souza et al. [2020], de Campos et al. [2020] e de Assis et al. [2021]. Em resumo, essas características são:

- Existe um conjunto de n localidades candidatas a receber um total de p mamógrafos;
- Cada mamógrafo k possui uma capacidade anual de realização de mamografias igual a cap;
- Cada cidade i possui uma demanda anual por exames de mamografia;
- As mulheres de uma cidade j podem ser atendidas por mamógrafos instalados em uma outra cidade i, porém a distância entre i e j não pode ultrapassar R km;
- Para que uma cidade i seja sede de um equipamento de mamografia, é necessário que ela possua estrutura para tal;
- Uma cidade que possui demanda de exames superior a $demMin$ é considerada como tendo estrutura para sediar um mamógrafo;
- Uma cidade j pode ser atendida por mamógrafos que estejam em uma ou mais cidades i;
- Para que um mamógrafo forneça exames a mulheres de uma cidade diferente daquela onde ele está instalado, a cidade sede precisa estar com sua demanda totalmente coberta;
- Uma cidade pode ser sede de um mamógrafo e, ainda assim, ter parte de sua demanda coberta por mamógrafos de outra cidade.

Nesse contexto, é necessário estabelecer em quais cidades os mamógrafos devem ser instalados e, além disso, designar os atendimentos desses mamógrafos à população. Os objetivos são: *(i)* maximizar a cobertura da demanda de exames; e *(ii)* minimizar o somatório das distâncias percorridas pelas mulheres em busca do exame de mamografia.

3. Métodos de Solução

Nesta seção são apresentados os métodos e ferramentas utilizados no desenvolvimento do trabalho. Nas subseções 3.1, 3.2, 3.3 e 3.4 abordamos o *MOEA Framework*, os algoritmos propostos NSGA-II (*Non-dominated Sorting Genetic Algorithm II*) e SPEA2 (*Strength Pareto Evolutionary Algorithm 2*), a representação de uma solução e a função de designação, respectivamente. Na Subseção 3.5 são apresentadas as formas de construção da população inicial. A avaliação das funções objetivo são mostradas na Subseção 3.6. Os operadores de cruzamento e mutação usados são definidos na Subseção 3.7.

3.1. *MOEA Framework*

Neste trabalho é utilizado o *MOEA Framework* [Hadka, 2016], uma biblioteca Java gratuita e de código aberto para desenvolver e experimentar algoritmos evolucionários multiobjetivos (MOEAs) e outros algoritmos de otimização de propósito geral, tanto mono quanto multiobjetivos. O *framework* suporta vários algoritmos genéticos, dentre eles o NSGA-II e o SPEA2, utilizados neste trabalho. Além disso, o *MOEA Framework* também possui ferramentas para projetar, desenvolver, executar e testar estatisticamente algoritmos de otimização.

3.2. Algoritmos Utilizados

Dois algoritmos evolucionários foram utilizados para a resolução do PMLM, sendo eles o NSGA-II [Deb et al., 2002] e o SPEA2 [Zitzler et al., 2001]. Em ambos, o processo inicia-se com uma população inicial de t indivíduos. Cada indivíduo dessa população recebe uma nota baseada em uma função de aptidão. Os indivíduos com maior aptidão são selecionados para cruzamento e mutação. De um conjunto formados por pais e filhos, os t melhores são selecionados para a próxima geração. Repete-se o processo até que o critério de parada seja atingido.

https://proceedings.science/p/157495?lang=pt-br

No NSGA-II, a classificação dos indivíduos se baseia em dois critérios: *(i)* fronteira à qual o indivíduo pertence (critério de convergência em relação à Fronteira de Pareto) e *(ii)* distância de aglomeração (critério de diversidade). Na classificação por fronteira, os indivíduos não-dominados estão na fronteira 0. Na fronteira 1 estão posicionados indivíduos que são dominados somente por membros da fronteira 0. Indivíduos da fronteira 2 são dominados somente por indivíduos da fronteira 1 e assim por diante. Quanto menor a fronteira que o indivíduo estiver, mais apto ele é. A distância de aglomeração funciona como critério de desempate quando dois indivíduos estão na mesma fronteira. Nesse caso, indivíduos em uma região menos povoada terão uma distância de aglomeração maior e levarão vantagem no confronto com indivíduos de regiões mais densas.

O SPEA2 utiliza um arquivo externo no qual são mantidas as melhores soluções obtidas ao longo do processo. Para classificar os indivíduos, o algoritmo combina, na mesma função de aptidão $F(s)$, um componente de convergência à fronteira de Pareto $R(s)$ e um componente de diversidade $D(s)$. Na parcela de convergência, $R(s)$ é determinado pelo somatório das forças dos indivíduos que dominam s. A força de um indivíduo s representa o número de soluções u que s domina. A parcela de diversidade $D(s)$ é dada pela densidade do indivíduo, calculada por:

$$D(s) = \frac{1}{\sigma_s^v + 2} \tag{1}$$

em que σ_s^v significa a distância do indivíduo s em relação ao v-ésimo indivíduo mais próximo dele, sendo v a raiz quadrada da soma do número de indivíduos da população e do arquivo externo. Finalmente, a função de aptidão $F(s)$ é dada por $F(s) = R(s) + D(s)$, sendo que quanto menor for esse número, mais apto é o indivíduo.

3.3. Representação de um indivíduo/solução

Um indivíduo s no PMLM é representado por um vetor de inteiros y de tamanho $p + 1$, no qual y_k representa a cidade de instalação do mamógrafo k e p é a quantidade de mamógrafos. A posição $p + 1$ desse vetor representa a forma de designação dos mamógrafos. A partir do vetor y, uma matriz de designação x é gerada, de dimensão $p \times n$, em que n representa a quantidade de cidades e x_{kj} representa a quantidade de exames que o mamógrafo k fornece para a cidade j. A matriz x é construída de forma determinística e o valor de y_{p+1} é utilizado como parâmetro de construção dela.

Cidades	1	2	3	4	5
Demandas	6000	3000	1000	7000	4000

Vetor y

Mamógrafo	1	2	3	4	P+1
Cidade Sede	4	4	2	1	3

Matriz x

Mamógrafo/Cidade	1	2	3	4	5
1	0	0	0	5000	0
2	0	0	0	2000	3000
3	1000	3000	0	0	1000
4	5000	0	0	0	0

Figura 1: Representação de uma solução.

A Figura 1 mostra um exemplo dessa representação. São cinco cidades, com demandas variadas, e quatro mamógrafos, sendo que cada um deles possui capacidade de realização de

5000 mamografias por ano. É possível observar que os mamógrafos 1 e 2 estão instalados na cidade 4. A matriz x informa que o mamógrafo 1 disponibiliza toda a sua capacidade (5000 exames) para sua cidade-sede (cidade 4), enquanto o mamógrafo 2 contribui com 2000 exames para cobrir a parte da demanda da cidade 4 que não foi coberta pelo mamógrafo 1. A capacidade ociosa de 3000 exames do mamógrafo 2 é utilizada para cobrir parte da demanda da cidade 5, que não possui nenhum mamógrafo.

3.4. Designação de exames

Para montar a matriz x e, assim, determinar quais cidades serão atendidas por cada mamógrafo, é utilizada uma função de designação. Para cada mamógrafo k, busca-se cobrir totalmente a demanda de sua cidade-sede i até o limite da capacidade de exames do equipamento. Se toda a demanda da cidade i estiver coberta e o mamógrafo k ainda possuir capacidade de atendimento, as cidades vizinhas em um raio máximo de R km tornam-se candidatas a se beneficiar da capacidade ociosa de k. A lista de cidades candidatas é, então, ordenada de acordo com a variável y_{p+1} (ver Subseção 3.4.1) e, seguindo a ordem da lista, as cidades são atendidas enquanto houver capacidade ociosa no mamógrafo.

Dois parâmetros são importantes no processo de aproveitamento da capacidade ociosa dos mamógrafos: max_atend e $max_cidades$. O primeiro limita a quantidade de exames que cada cidade candidata receberá, sendo esse limite definido por um percentual max_atend da capacidade ociosa de k. Essa limitação busca evitar que uma única cidade consuma toda a capacidade ociosa disponível de k, proporcionando soluções que maximizam a cobertura. Se, após passar por todas as cidades da lista, ainda houver demanda residual e capacidade ociosa no mamógrafo, permite-se, então, que esse limite seja ultrapassado.

O segundo parâmetro define a quantidade máxima de cidades que poderão ser atendidas, gerando diversidade de soluções, de forma a privilegiar a minimização do deslocamento. Assim como a forma de ordenação da lista de candidatos, o parâmetro $max_cidades$ é definido pela variável y_{p+1} (ver Subseção 3.4.2).

Algoritmo 1: designaExames($y, x, exames, max_atend$)

```
1  k ← 0
2  enquanto (k ≤ p) faça
3      se (exames(k) > 0) então
4          lst_candidatos ← obtemCidadesDaRegiao(k)
5          ordena_lista(lst_candidatas, y_{p+1})
6          max_cidades ← obtemMaxCidades(y_{p+1})
           limite ← exames[k] × max_atend
7          enquanto (houver cidade c ∈ lst_candidatas) faça
8              realizaAtendimento(c, limite, x)
9              se (atendidas ≥ max_cidades) então
10                 passa_para_proximo_mamografo
11             fim
12         fim
13     fim
14     k ← k + 1
15 fim
```

O Algoritmo 1 mostra o funcionamento da função de designação. Um mamógrafo individual é identificado pelo índice k (linhas 1 e 2). Se houver capacidade de atendimento do k-ésimo

mamógrafo (linha 3), é gerada uma lista de cidades candidatas a se beneficiarem dessa sobra (linha 4), a qual é ordenada conforme Subseção 3.4.1 (linha 5). O número máximo de cidades atendidas pelo mamógrafo, além da cidade-sede, é calculado conforme Subseção 3.4.2 (linha 6). Baseado na quantidade de exames disponíveis no k-ésimo mamógrafo, é definido o limite de exames para cada cidade (linha 6). As cidades da lista são atendidas respeitando-se o limite de exames (linha 8) e o limite de cidades (linhas 9 e 10).

3.4.1. Ordenação da Lista de Cidades Candidatas

A lista de cidades candidatas é ordenada de acordo com os seguintes valores que y_{p+1} assume:

(a) 1 ou 6: não haverá ordenação adicional, mantendo-se a sequência original das cidades;

(b) 2, 7 ou 11: cidades com maior demanda residual primeiro;

(c) 3, 8 ou 12: cidades com menor demanda residual primeiro;

(d) 4, 9 ou 13: cidades com menor distância em relação à cidade-sede primeiro;

(e) 5, 10 ou 14: cidades com menor número de opções de atendimento devem ser atendidas primeiramente. Entende-se por opção de atendimento uma cidade com mamógrafo(s) situado(s) em um raio de até R km de distância.

3.4.2. Quantidade de Cidades a Cobrir

O valor que a variável y_{p+1} assume determina, também, quantas cidades, no máximo, poderão ser atendidas pela capacidade ociosa de um mamógrafo da seguinte forma:

(a) De 1 a 5: Poderão ser atendidas n cidades, ou seja, não se impõe um limite;

(b) De 6 a 10: Limite de 2 cidades;

(c) De 11 a 14: Limite de 3 cidades.

3.5. Geração de População Inicial

A população inicial é gerada de forma que parte dos indivíduos sejam de qualidade com relação ao objetivo de cobertura, parte privilegie a minimização do deslocamento e parte totalmente aleatória, conforme descrito abaixo:

(a) **Soluções com viés de cobertura:** para gerar soluções de qualidade com relação à cobertura, foi aplicada a fase de construção GRASP [Feo e Resende, 1995]. Para tanto, ordenam-se as cidades que possuem infraestrutura para receber um mamógrafo de acordo com suas demandas. Dentre as cinco com maiores demandas, uma delas é sorteada para ser sede do primeiro mamógrafo. Enquanto houver mamógrafo sem alocação, repete-se o processo considerando-se a demanda atualizada na ordenação, ou seja, descontando-se a parte da população já atendida por outros mamógrafos. O valor da variável y_{p+1} é definido por um número aleatório entre 1 e 5, de forma que um mamógrafo possa cobrir o máximo de demanda possível.

https://proceedings.science/p/157495?lang=pt-br

Algoritmo 2: construcaoGRASP(y, dem)

```
1 k ← 0
2 enquanto (k ≤ p) faça
3 |   ordena_por_demanda(lst_estrutura)
4 |   lst ← obtemPrimeiros(lst_estrutura, 5)
5 |   i ← escolhaAleatoria(lst)
6 |   y_k ← i
7 |   k ← k + 1
8 fim
9 y_{p+1} ← escolhaAleatoria(1, 5)
```

No Algoritmo 2, um mamógrafo é identificado pelo índice k (linha 1). Para alocar cada mamógrafo (linha 2), é gerada uma lista de cidades com infraestrutura para sediar o equipamento, ordenadas por demanda residual (linha 3). Dentre as primeiras 5 cidades da lista (linha 4), uma delas é escolhida aleatoriamente (linha 5) para sediar o equipamento (linha 6). Na linha 9 é escolhido aleatoriamente um número entre 1 e 5, de forma que a montagem da matriz de designação privilegie a cobertura.

(b) **Soluções com viés de deslocamento:** nestas soluções, a alocação dos mamógrafos é definida aleatoriamente, isto é, cada mamógrafo pode ser instalado em qualquer uma das cidades que possuam infraestrutura. Porém, o valor da variável y_{p+1} é definido por um número aleatório entre 6 e 14, de forma a restringir o número de cidades que um mamógrafo pode atender.

(c) **Soluções aleatórias:** todas as posições do vetor y são geradas aleatoriamente, sendo que, de y_1 a y_p, as opções são as cidades com infraestrutura, enquanto y_{p+1} pode receber valores entre 1 e 14.

3.6. Avaliações de Funções objetivo

As duas funções objetivo são avaliadas com base na matriz x mencionada na Subseção 3.3 e geradas pelo Algoritmo 1. Para calcular a cobertura, basta somar todos os valores da matriz de designação (matriz x). Exemplificando, a cobertura é igual a 20000 na Figura 1. O deslocamento é calculado com base no somatório da matriz x; porém, cada variável x_{kj} deve ser multiplicada pela distância entre a cidade-sede do mamógrafo k e a cidade j por ele atendida.

3.7. Operadores de cruzamento e mutação

O operador de cruzamento utilizado foi o Cruzamento de Ponto Único (*One-Point Crossover*), amplamente utilizado na literatura. Esse operador traça uma linha que divide os cromossomos dos dois pais. Dois filhos são gerados pela combinação das partes dos pais. A Figura 2 ilustra a aplicação desse operador. Nela, o "FILHO 1" herda do "PAI 1" os três primeiros genes ([4 4 2]) e do "PAI 2", os dois últimos ([3 2]). Por sua vez, o "FILHO 2" herda do "PAI 2" os três primeiros genes ([1 1 5]) e do "PAI 1", os dois últimos ([1 3]).

Assim como na natureza, nem todo par de pais irá gerar filhos. Para isso, o parâmetro pc determina a probabilidade de cruzamento de dois indivíduos.

Com relação à mutação, foi utilizada a Mutação Uniforme (*Uniform Mutation*), também amplamente utilizada na literatura. Cada gene do indivíduo pode sofrer mutação e mudar de valor, porém, isso ocorre de acordo com o valor do parâmetro de probabilidade de mutação pm. A Figura 3 ilustra que o terceiro gene do indivíduo sofreu mutação, tendo seu valor sido modificado de 2 para 1.

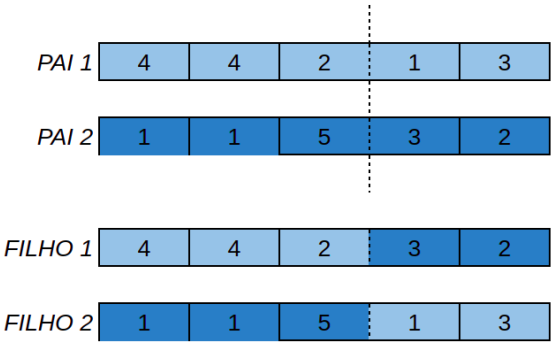

Figura 2: Cruzamento de Ponto Único

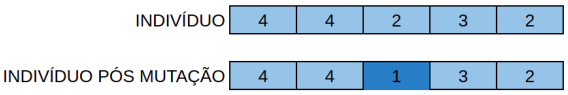

Figura 3: Mutação Uniforme

4. Experimentos Computacionais

Os algoritmos NSGA-II e SPEA2 foram desenvolvidos utilizando-se a linguagem Java, com uso do *MOEA Framework*. O experimento foi realizado em um computador Intel Core i7-10750H CPU 2.6 GHz com 16 GB de memória RAM sob um sistema operacional Ubuntu 20.04.

Os valores dos parâmetros dos algoritmos foram definidos após testes empíricos. O tamanho da população foi de 100 indivíduos. O critério de parada foi definido em 1000000 avaliações de funções objetivo. O parâmetro max_atend foi fixado em 70%. A probabilidade de realizar um cruzamento foi de $pc = 95\%$ e a de mutação, de $pm = 5\%$. Na inicialização, cada solução tem 40% de chance de ser gerada com viés de cobertura, 30% com viés de deslocamento e 30% aleatória.

Para testar os algoritmos, foi utilizada a instância "MG-324-375-5069", descrita em de Campos et al. [2020], que apresenta dados do Estado de Minas Gerais. Essa instância possui 324 mamógrafos, sendo que cada mamógrafo possui uma capacidade anual de realização de 5069 exames. Nessa instância, considera-se que uma cidade tem infraestrutura quando sua demanda é maior ou igual a 375 exames.

Para avaliar a qualidade dos resultados e comparar o desempenho dos dois algoritmos (NSGA-II e SPEA2), foram utilizados os indicadores de qualidade *Hypervolume* [Zitzler e Thiele, 1999] e *Spacing* [Schott, 1995]. O primeiro é um indicador que calcula o tamanho da região dominada pelos indivíduos contidos na fronteira de Pareto obtida, sendo que, quanto maior seu valor, melhor. Para o cálculo do *Hypervolume* é necessário um ponto de referência, sendo o ponto de Nadir o mais comum. No caso do PMLM, o ponto de Nadir é 0 para o objetivo de cobertura, enquanto para o objetivo de deslocamento ele foi definido como sendo o somatório dos produtos da demanda de cada cidade pelo limite de deslocamento R. O *Spacing* calcula o espaçamento relativo de soluções consecutivas do conjunto não-dominado, de forma que quanto menor seu valor, melhor, sendo zero um valor que representa um espaçamento uniforme. A escolha desses indicadores específicos se deu por serem indicadores unários e que não dependem de um conjunto de referência. A ideia de usar indicadores unários é facilitar a comparação de futuros trabalhos com os apresentados neste artigo. Com relação ao conjunto de referência, optamos por não usá-lo por este ser o primeiro trabalho a abordar o PMLM.

O experimento realizado consistiu em 30 execuções do algoritmo baseado no NSGA-II e outras 30 execuções do algoritmo baseado no SPEA2. A Tabela 1 reporta os resultados dos dois indicadores para cada execução de cada algoritmo. A tabela divide-se em duas partes, sendo três colunas referentes às execuções do NSGA-II e outras três colunas dedicadas ao SPEA2. Para cada algoritmo, é mostrado o número da execução e os valores de *Hypervolume* e *Spacing* do conjunto

Tabela 1: Indicadores de Qualidade por Algoritmo.

NSGA-II			SPEA2		
Execução	*Hypervolume*	*Spacing*	Execução	*Hypervolume*	*Spacing*
1	0,90607	58188,04130	1	0,90315	20286,53352
2	0,90505	54713,10768	2	0,89925	21800,74863
3	0,90720	50312,12662	3	0,90677	30822,37352
4	0,90866	68928,91163	4	0,90334	21522,89651
5	0,90039	43483,99240	5	0,90594	26047,42542
6	0,90773	58319,34234	6	0,90444	31920,08195
7	0,90432	53511,64228	7	0,90540	24031,26077
8	0,89822	48101,17174	8	0,90730	20480,80785
9	0,90456	48945,64725	9	0,90475	28403,98659
10	0,90302	45850,78802	10	0,90459	36453,84278
11	0,90384	50757,04428	11	0,90291	24380,46351
12	0,90373	60941,13604	12	0,90662	27037,58054
13	0,91014	59188,88052	13	0,90282	33435,98989
14	0,90167	49696,96997	14	0,90188	21442,90026
15	0,90834	57764,34773	15	0,90668	23283,32478
16	0,90170	43916,99959	16	0,90543	29818,80763
17	0,90941	73703,93074	17	0,90254	18550,93770
18	0,90369	45111,02232	18	0,89932	24570,68675
19	0,90475	41692,23673	19	0,90766	27057,25422
20	0,90863	63273,05786	20	0,90654	22048,49045
21	0,91115	61838,22235	21	0,90740	21196,58159
22	0,90156	43307,38564	22	0,90396	24622,20889
23	0,90787	47252,60249	23	0,90491	23515,10865
24	0,90153	44237,27217	24	0,90512	25240,32020
25	0,90672	49572,83602	25	0,90709	21788,67984
26	0,90140	41882,65054	26	0,89438	20242,49543
27	0,90601	60259,59178	27	0,90744	26403,36048
28	0,90749	57091,73822	28	0,90438	34447,08281
29	0,90747	53656,64901	29	0,90181	22113,62698
30	0,90714	75793,22649	30	0,90415	25020,85710

não-dominado resultante, respectivamente.

Pela Tabela 1, pode-se observar valores de *Hypervolume* muito parecidos, variando de $0{,}89822$ a $0{,}91115$ para o NSGA-II e de $0{,}89438$ a $0{,}90766$ para o SPEA2. Em relação ao indicador *Spacing*, aparentemente há vantagem para o SPEA2, uma vez que o maior valor do indicador encontrado foi 36453, ainda assim menor que o menor valor do NSGA-II.

As Figuras 4 e 5 mostram os gráficos *box-plot* relativos aos indicadores *Hypervolume* e *Spacing*, respectivamente, para os dois algoritmos. Os gráficos corroboram a percepção de superioridade do SPEA2 com relação ao indicador *Spacing*, ao passo que não é possível afirmar a superioridade de nenhum dos dois algoritmos em relação ao *Hypervolume*.

Ao final das 60 execuções (30 de cada algoritmo) foi apurada a fronteira de Pareto final, contendo somente as soluções não-dominadas dentre todas as soluções não-dominadas geradas durante o experimento. Para esse conjunto de soluções, o *Hypervolume* foi igual a $0{,}91$, enquanto o *Spacing* ficou em $36245{,}31$. O melhor valor para a função objetivo de cobertura foi de 1606344 exames e o menor valor de deslocamento total foi 351388 km.

5. Discussão

Neste trabalho foi apresentado o Problema Multiobjetivo de Localização de Mamógrafos. Foram considerados dois objetivos, sendo eles a cobertura e o deslocamento. No primeiro, a

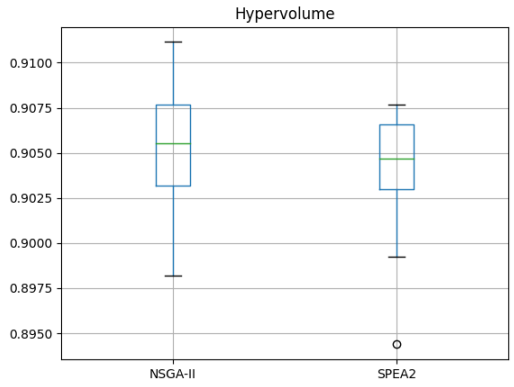

Figura 4: *Box-plot* do *Hypervolume*

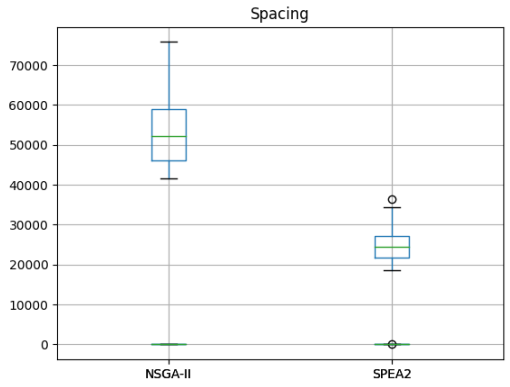

Figura 5: *Box-plot* do *Spacing*

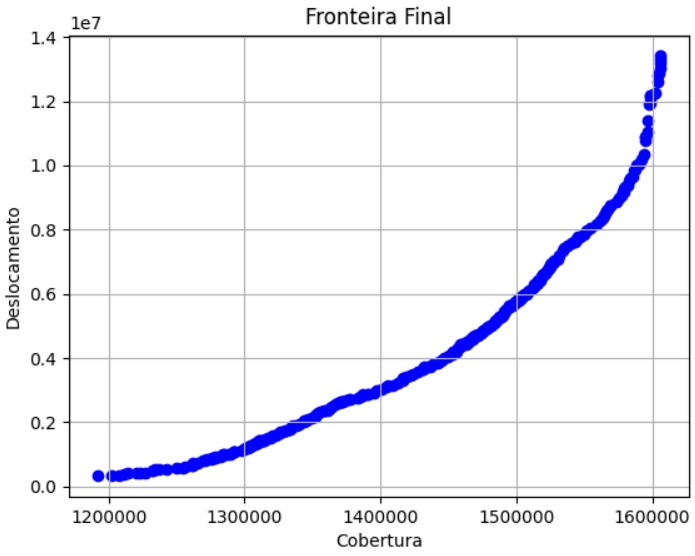

Figura 6: Fronteira de Pareto final.

intenção é maximizar o quantitativo de mulheres atendidas pelos mamógrafos disponíveis, enquanto no segundo busca-se minimizar o somatório do deslocamento dessas mulheres.

Para resolver o problema, foi utilizado o *MOEA Framework*, sendo implementados os algoritmos NSGA-II e SPEA2. Os resultados foram avaliados com base em dois indicadores de desempenho: *Spacing* e *Hypervolume*. No que se refere ao *Spacing*, o algoritmo SPEA2 mostrou-se mais eficiente, o que significa que os indivíduos ficaram mais bem distribuídos na fronteira de Pareto. Quanto ao *Hypervolume*, notou-se uma pequena vantagem do NSGA-II sobre o SPEA2, porém essa vantagem não foi provada estatisticamente conforme resultado apresentado pelo *framework*.

Na fronteira final, observou-se uma variedade de soluções de boa qualidade, propiciando ao gestor avaliar o custo-benefício entre cobertura e deslocamento. Na Figura 6, por exemplo, ilustra-se que foi possível reduzir pela metade o deslocamento total de mulheres com uma perda

de cerca de 100000 atendimentos (ou aproximadamente 6% em relação ao máximo de cobertura obtido). Com isso, o tomador de decisão pode, por exemplo, entender que vale a pena adquirir novos mamógrafos, uma vez que o volume de deslocamentos precisa ser alto para atingir a máxima cobertura.

Com relação a trabalhos futuros, sugere-se um projeto de calibração de parâmetros, uma vez que, além dos parâmetros intrínsecos aos algoritmos genéticos utilizados, outros parâmetros foram utilizados na construção da população inicial e na função de designação. Para essa tarefa pode-se usar o pacote Irace [López-Ibáñez et al., 2016]. Também espera-se que novos trabalhos busquem introduzir novos objetivos. Um desses objetivos é a minimização do desvio padrão dos percentuais de cobertura das cidades, de forma a chegar a uma cobertura mais equânime. Outra análise é com relação à existência de municípios isolados em Minas Gerais, isto é, distantes de outros municípios com infraestrutura, dado que o raio de 60 km pode ser fator inviabilizante. Em vista disso, recomenda-se o relaxamento do valor de R e, consequentemente, a introdução de um novo objetivo, a minimização do excesso de deslocamento resultante desse relaxamento. Finalmente, este estudo deveria ser replicado para outros estados brasileiros, de forma a aumentar sua contribuição para a sociedade, bem como para validar os algoritmos propostos.

6. Agradecimentos

Os autores registram seus agradecimentos ao CEFET-MG, ao IFMG, à UFOP, à UFV, ao IFB e às agências CNPq (processos 303266/2019-8 e 428817/2018-1), FAPEMIG (processo PPM-CEX 676/17) e CAPES (código de financiamento 001), pelo apoio ao desenvolvimento deste trabalho.

Referências

Amaral, P., Luz, L., Cardoso, F., e Freitas, R. (2017). Distribuição espacial de equipamentos de mamografia no brasil. *Revista Brasileira de Estudos Urbanos e Regionais (RBEUR)*, 19(2):326–341.

de Assis, R. d. S. L., Andrade de Campos, M. V., Souza, M. J. F., Silva, M. A. L., de Siqueira, E. C., e de Souza, S. R. (2021). Um algoritmo iterated greedy search para o problema de localização de mamógrafos com atendimento parcial. In *Anais do LIII Simpósio Brasileiro de Pesquisa Operacional*, volume 53, João Pessoa (PB). Galoá: Campinas. Disponível em `https://proceedings.science/sbpo-series/sbpo-2021/papers/um-algoritmo-iterated-greedy-search-para-o-problema-de-\-localizacao-de-mamografos-com-atendimento-parcial`. Acesso em: 22 jun. 2022.

de Campos, M. V. A., de Sá, M. V. S. M., Rosa, P. M., Penna, P. H. V., de Souza, S. R., e Souza, M. J. F. (2020). A mixed linear integer programming formulation and a simulated annealing algorithm for the mammography unit location problem. In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS)*, volume 1, p. 428–439, Online. Scitepress.

Deb, K., Pratap, A., Agarwal, S., e Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.

Feo, T. A. e Resende, M. G. (1995). Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6(2):109–133.

Hadka, D. (2016). *Beginner's Guide to the MOEA Framework*. Createspace Independent Publishing Platform, 2nd edition. URL `http://moeaframework.org/documentation.html`.

Kirkpatrick, S., Gelatt, C. D., e Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

López-Ibáñez, M., Dubois-Lacoste, J., Cáceres, L. P., Birattari, M., e Stützle, T. (2016). The IRACE package: Iterated racing for automatic algorithm configuration. *Operations Research Perspectives*, 3:43–58.

Miranda, S. e Patrocinio, A. (2018). Distribuição de mamógrafos por macrorregião do brasil. In *Anais do V Congresso Brasileiro de Eletromiografia e Cinesiologia e X Simpósio de Engenharia Biomédica*, p. 433–436, Ribeirão Preto. Associação Brasileira de Física Médica.

Mladenović, N. e Hansen, P. (1997). Variable neighborhood search. *Computers & Operations Research*, 24(11):1097–1100.

Ruiz, R. e Stützle, T. (2008). An iterated greedy heuristic for the sequence dependent setup times flowshop problem with makespan and weighted tardiness objectives. *European Journal of Operational Research*, 187(3):1143–1159.

Schott, J. R. (1995). *Fault tolerant design using single and multicriteria genetic algorithm optimization*. PhD thesis, Massachusetts Institute of Technology.

Souza, M. J. F., Penna, P. H. V., Stilpen, M., Rosa, P. M., Monteiro, J. C., e Lisboa, M. R. (2019). Localização de mamógrafos: formulações e estudo preliminar de caso de Rondônia. In *Anais do LI Simpósio Brasileiro de Pesquisa Operacional, Limeira, SP*, volume 2, Limeira. Galoá. URL `https://proceedings.science/sbpo-series/sbpo-2019/papers/localizacao-de-mamografos--formulacoes-e-estudo-preliminar-de-\-caso-de-rondonia`. Acesso em: 22 jun. 2022.

Souza, M. J. F., Penna, P. H. V., Moreira de Sá, M. V. S., e Rosa, P. M. (2020). A VNS-based algorithm for the mammography unit location problem. In Benmansour, R., Sifaleras, A., e Mladenović, N., editors, *Variable Neighborhood Search: 7th International Conference, ICVNS 2019, Rabat, Morocco, October 3–5, 2019, Revised Selected Papers*, volume 12010 of *Lecture Notes in Computer Science*, chapter 3, p. 37–52. Springer, Cham, Switzerland.

Witten, M. e Parker, C. C. (2018). Screening mammography recommendations and controversies. *Surgical Clinics of North America*, 98(4):667–675.

Xavier, D. R., Oliveira, R. A. D. d., Matos, V. P. d., Viacava, F., e Carvalho, C. d. C. (2016). Cobertura de mamografias, alocação e uso de equipamentos nas regiões de saúde. *Saúde em Debate*, 40(110):20 – 35. ISSN 0103-1104.

Zitzler, E., Laumanns, M., e Thiele, L. (2001). SPEA2: Improving the strength pareto evolutionary algorithm. *TIK-report*, 103.

Zitzler, E. e Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271.