A Mixed Linear Integer Programming Formulation and a Simulated Annealing Algorithm for the Mammography Unit Location Problem

Marcos Vinícius Andrade de Campos[1][a], Manoel Victor Stilpen Moreira de Sá[2][b], Patrick Moreira Rosa[2][c], Puca Huachi Vaz Penna[2][d], Sérgio Ricardo de Souza[1] and Marcone Jamilson Freitas Souza[2][e]

[1]*Departamento de Computação, Centro Federal de Educação Tecnológica, Avenida Amazonas, 7675, 30.510-000 Belo Horizonte, Brazil*

[2]*Departamento de Computação, Universidade Federal de Ouro Preto, 35.400-000 Ouro Preto, Brazil*

Keywords: Mammography Unit Location, Maximal Covering Location, Facility Location, Simulated Annealing.

Abstract: Breast cancer is the most commonly occurring one in the female population. Early diagnosis of this disease, through mammography screening, can increase the chances of cure to 95%. Studies show that Brazil has a relatively satisfactory number of mammography units, but this equipment is poorly geographically distributed. This paper focuses on the Mammography Unit Location Problem (MULP), which aims an efficient distribution of mammography units, in order to increase the covered demand. Focusing on the State of Minas Gerais, Brazil, an analysis is made considering that, in the real world, there is a difficulty in relocating equipment already installed. Therefore, it would be interesting to optimize the location of new equipment purchases. Since MULP is NP-hard, an algorithm based on the Simulated Annealing meta-heuristic is also developed to handle large instances of the problem.

1 INTRODUCTION

Breast cancer is the most common cause of death among women (Bray et al., 2018). Only in Brazil, this type of cancer led to obit 16724 women in 2017, which represents 2.9% of female deaths (INCA, 2017).

In turn, mammography screening is the primary method of early detection for the diagnosis of breast-related malignancies (Xavier et al., 2016). Combined with proper treatment, early detection via mammography screening contributes to the reduction in the mortality rate of women with breast cancer (Berry et al., 2005). When a tumor is detected in the early stages, the chance of cure is greater than 95% (Witten and Parker, 2018).

The Brazilian Ministry of Health recommends that women between the ages of 50 and 69 perform the mammography screening at least once every two years (Brasil, 2017). In addition, it is estimated that, by diagnostic indication, 8.9% women of this age group and 20% of the female population between 40 and 49 years of age need to perform the screening annually. Thus, 58,9% of women, between 50 and 69 years old, need mammography screenings annually, besides 20% of those between 40 and 49 years old.

The Brazilian healthcare system has a significant amount of mammography units to serve the population (Amaral et al., 2017). In December 2015, these authors estimated that there were 4647 units in use, of which 2083 were available for use in the Brazilian public healthcare system. This amount of equipment would be sufficient to cover the estimated demand of 12.7 million mammography screenings per year. On the other hand, also according to these authors, the Brazilian Ministry of Health determines that the maximum acceptable distance between patients and equipment is 60 km. Given this guideline, 4.5% of the required screenings would not be performed. This situation would be even more grave if only public mammography equipment were considered. Thus, the number would go from 4.5% to 17% of uncovered

[a] https://orcid.org/0000-0002-5599-8889
[b] https://orcid.org/0000-0003-2635-3157
[c] https://orcid.org/0000-0003-3897-0961
[d] https://orcid.org/0000-0001-5414-1405
[e] https://orcid.org/0000-0002-7141-357X

de Campos, M., Moreira de Sá, M., Rosa, P., Penna, P., de Souza, S. and Souza, M.
A Mixed Linear Integer Programming Formulation and a Simulated Annealing Algorithm for the Mammography Unit Location Problem.
DOI: 10.5220/0009420704280439
In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 1*, pages 428-439
ISBN: 978-989-758-423-7

women, which would represent 2.17 million screenings not performed due to the lack of available equipment within a maximum radius of 60 km.

In the current scenario, there is an inefficient geographical distribution of mammography equipment as detected by several authors, in addition to the one previously mentioned (Miranda and Patrocinio, 2018; Silva et al., 2018). Each of these papers has a different focus, however, all of them agree that, although the amount of equipment is relatively adequate, the poor distribution of mammography equipment still causes difficulty in accessing the mammography screening.

In this paper, we address the Mammography Unit Location Problem (MULP), considering the Brazilian reality. The objective is to improve the geographic distribution of mammography units available in the public healthcare system of Brazil. Initially, we propose a Mixed-Integer Linear Programming (MILP) model for the MULP. A case study of the Minas Gerais state is introduced, and two scenarios are analyzed. In the first one, we considered the impossibility of changing the location of the equipment installation, and the goal is to determine the locations of new mammography units to be acquired; in the second one, there is the freedom to move the equipment, which enables better results with the same amount of mammography units.

The problem under study can be formulated as the Maximal Covering Location Problem (Church and ReVelle, 1974) with additional constraints. It is an NP-Hard problem (Garey and Johnson, 1979), and this issue limits the use of exact methods, since they may require an unacceptable time for decision making in real instances. Thus, we developed an algorithm based on the Simulated Annealing metaheuristic to allow the treatment of large instances of the problem.

The rest of this paper is organized as follows. Section 2 shows a brief literature review dealing with Facility Locating Problems, especially concerning to health prevention equipment, such as mammography units. Section 3 describes the Mammography Unit Location Problem, considering the characteristics established by the Brazilian Ministry of Health. Section 4 introduces a MILP formulation for this problem, while Section 5 presents the proposed Simulated Annealing based-algorithm. The results of the computational experiments with these methods are reported in Section 6. Finally, the conclusions and the perspectives of future work are presented in Section 7.

2 LITERATURE REVIEW

The facility location is a critical issue, either for industry or for a healthcare facility (Daskin and Dean, 2005). In this kind of problem, there are customers, with their respective demands, and locations, where potentially the facilities will be installed. The aim is to determine in which locations the facilities will be open and, consequently, in which of those open facilities each customer will be associated with.

The first facility location study dates from 1909. The objective was to determine the location of a warehouse so that its distance to its customers was minimized (Weber, 1929). Since then, several studies have been carried out in the theme (Ahmadi-Javid et al., 2017).

A facility location model can be continuous or discrete in nature. Continuous models are those where facilities can be located anywhere in the feasible region, while discrete models only allow them to be established at predetermined locations, which eventually may also be a point of demand. Discrete models can be classified into three categories: (i) median-based problems; (ii) covering-based problems; and (iii) other problems (Ahmadi-Javid et al., 2017). Median-based models are characterized by locating facilities at candidate points to minimize the weighted average distance costs between demand points and the facilities to which they are assigned. In turn, in the covering-based location problem, given a specified level of demand coverage, which must be achieved, the goal is to find the number and location of facilities such that all demand points are within a specified travel distance (or time) of the facility that serves them. The third category includes problems that are not in either of the above two categories, for example, p-dispersion problem, maximum dispersion problem, among others.

Problems dealing with health services, such as the MULP, are often coverage type. This category is subdivided into p-center location, set covering location, and maximal covering location problems. The p-center location problem consists of locating a facility, such as a school or a hospital, so that the distance of the customer farthest from it is minimized (Hakimi, 1964). The set covering location problem aims to minimize the number of open facilities, ensuring that all demand is met and respecting a maximum distance/time between any customer and the facility that covers it (Toregas et al., 1971). Finally, given the number of facilities to be opened, the Maximal Covering Location Problem (MCLP) intends to determine the best location of them, so the covered demand is maximized.

The MCLP-based models are widely used in healthcare, especially in the public sector, due to budgetary constraints (Sathler et al., 2017).

The location of preventive healthcare facilities is addressed in several papers (Verter and Lapierre, 2002; Zhang et al., 2009; Zhang et al., 2010; Gu et al., 2010; Zhang et al., 2012; Davari et al., 2016; Dogan et al., 2019).

Verter and Lapierre (2002) proposed a model for determining the optimal configuration of a network of preventive healthcare facilities. The goal was to maximize the participation in prevention programs, including mammography screening. The model presented was based on three premises: (i) each individual would look for the nearest facility; (ii) the probability of a person participating in a prevention program decreases as his/her distance to the facility increases; and, finally, (iii) each opened facility would need to have a minimum number of customers. To model the probability of participation, a decay function was used, where the value would be 0 (zero) for distances greater than the maximum allowed, and would progressively increase until it reached 1 as the distance approached 0. The model was applied to locate mammography centers in the city of Montreal, Canada. Zhang et al. (2009) added the concept of congestion. Contrary to the previous study, where only time or distance was the factor that defined accessibility to the facility, the authors consider the total time spent, which is composed of the travel time, the waiting time to receive the service, and the service time. To solve it, an allocation heuristic and four location heuristics were developed. In Zhang et al. (2010), the authors add a decision variable in relation to the previous work. In addition, to decide whether a facility will open or not, there is concern about the number of services that each facility will offer. The model was built as a bi-level nonlinear optimization model. A Tabu Search based algorithm was developed to treat the problem. A new measure of accessibility was presented in Gu et al. (2010). Unlike previous work, where distance/time was the main factor of accessibility, the proposed measure combined regional availability, the distance between customers and facilities, and the number of customers that the facility attracts. The Huff-based competitive location model (Huff, 1964) was used to estimate the workload of facilities. The problem is solved using both exact methods and a local search-based heuristic algorithm using relocate moves. In Zhang et al. (2012), the probability of a customer to choose a feature was considered. The authors presented two models were presented. The first considers that the customer chooses a facility with a certain probability. The probability increases according to the attractiveness of the facility. The other model defines that the customer will choose the most attractive facility. The closer a facility is to the customer, the more attractive it will be to the facility. To solve the problem, a probabilistic-search based heuristic and a genetic algorithm were presented. Davari et al. (2016) enhanced the model by adding budget constraints. As a solution method, an algorithm based on the Variable Neighborhood Search metaheuristic was used. Dogan et al. (2019) presented a multi-objective model consisting of three terms: (i) minimization of the average total weighted deviation between the realized and maximum possible participation incurred over the planning period; (ii) minimization of the total deviation resulting from exceeding the expected acceptable waiting time at facilities averaged over the planning period; and (iii) minimization of the deviation resulting from exceeding the budget. The model was applied to locate Cancer Diagnostic, Screening and Training Centers in Istanbul, Turkey.

For the Brazilian reality, there are few proposals for a more efficient geographical distribution of mammography units (Corrêa et al., 2018; Souza et al., 2019).

Corrêa et al. (2018) analyzed if a more rational distribution of existing mammography units was possible, taking, as a case study, a set of 12 health regions of the Minas Gerais State, Brazil, involving 142 cities. The authors developed four formulations of integer linear programming, based on the classical p-median problem, whose objective was to minimize the sum of the distances between served locality and serving locality. The first formulation adds a maximum distance constraint between each city and the location that provides it with care. In the second formulation, the distance restriction is allowed to be violated, penalizing the positive distance deviation in the objective function. The third and fourth formulations reproduce the first and second ones, respectively. However, considering in the objective function, in addition to the distance between the city that serves and the city served, also the demand of screenings in the city that needs care. The authors concluded that, even respecting the distance restriction, the amount of mammography units in the studied region is sufficient to meet all the demand for screenings.

Souza et al. (2019) considered, as a case study, the distribution of the mammography units in the Rondônia State, Brazil. The authors proposed two integer linear programming formulations, both aiming to maximize the number of women attended, respecting the minimum demand restriction and the maximum distance restriction between each city and the city that hosts the mammography units. The two formulations differ with respect to the form of attending each city. In the first formulation, the demand of a city should be fully met, either by itself or by a neighboring city located within the service radius. In the second formulation, a partial service is possible, that is, a city could have fractions of its demand met by different cities. In order to solve the problem, a preprocessing is applied, in which every city with demand higher than the mammography attendance receives p mammography units, where p is the integer part of the division of the city demand for equipment screening capacity.

3 PROBLEM STATEMENT

The Mammography Unit Location Problem (MULP) addressed here has the following characteristics:

- There is a set S of n candidate cities to host p mammography units;
- Each mammography unit has an annual capacity of *cap* screenings;
- Each city i has an annual demand for screenings. According to the recommendations of the Brazilian Ministry of Health (Brasil, 2017), this number is composed by 58.9% of the female population aged from 50 to 69 years old and 20% of women aged from 40 to 49 years old;
- Women from a city j may be covered by mammography units installed in another city i. However, the distance between i and j can not exceed R km. In Brazil, the Ministry of Health (Brasil, 2017) recommends that R be equal to 60 km;
- In order to host a mammography unit, a city i must have the infrastructure to do so. In this paper, we consider that a city is eligible for hosting an equipment if it has an annual demand greater than or equal to *demMin* screenings;
- A city j can be covered by one or more cities;
- A city i that hosts an equipment can serve one or more cities, provided that its own demand is fully met.

The objective of MULP is to maximize the total number of women covered by the existing mammography units. In short, it is necessary to define in which cities the equipment will be installed and the number of equipment installed. Besides, it is necessary to define which cities will be served by these mammography units.

4 MILP FORMULATION

The mathematical programming formulation in the current paper is based on the work of Souza et al. (2019). In the mentioned study, it is necessary that the demand for mammography screening in each city be less than the capacity of a mammography unit. When this condition is not met, a preprocessing procedure is required. For each city with demand greater than the equipment's capacity, as many mammography units as necessary are allocated until residual demand is less than the capacity for screenings of the equipment. At the end of the procedure, all cities had lower demand than the service capacity, and the number of mammography units used for this is decreased from the total available. In this new work, we add variables and constraints that make the preprocessing unnecessary. Table 1 presents the formulation parameters and decision variables.

The proposed MILP formulation is given by the Equations (1)-(15):

$$\max \sum_{i \in N} \sum_{j \in S_i} dem_j \cdot x_{ij} \tag{1}$$

$$s.t. \sum_{i \in S_j} x_{ij} \leq 1 \quad \forall j \in N \tag{2}$$

$$\sum_{i \in N} y_i = p \tag{3}$$

$$\sum_{j \in S_i} dem_j \cdot x_{ij} \leq cap.y_i \quad \forall i \in N \tag{4}$$

$$z_i \geq y_i / p \quad \forall i \in N \tag{5}$$

$$z_i \leq y_i \quad \forall i \in N \tag{6}$$

$$t_i \leq z_i \quad \forall i \in N \tag{7}$$

$$t_i \geq dem_i \cdot x_{ii} - dem_i + 1 \quad \forall i \in N \tag{8}$$

$$t_i \leq x_{ii} \quad \forall i \in N \tag{9}$$

$$t_i \geq x_{ij} \quad \forall i, \forall j \in N \mid i \neq j \tag{10}$$

$$y_i = 0 \quad \forall i \in N \mid dem_i < demMin \tag{11}$$

$$x_{ij} \in [0,1] \quad \forall i, \forall j \in N \tag{12}$$

$$y_i \in \mathbb{Z} \quad \forall i \in N \tag{13}$$

$$z_i \in \{0,1\} \quad \forall i \in N \tag{14}$$

$$t_i \in \{0,1\} \quad \forall i \in N \tag{15}$$

The objective function (1) aims to maximize the total demand for mammography screenings. Constraints (2) ensure that the city j can only be served at most 100% of your demand. Constraint (3) forces that are used exactly p mammography units, allowing a city i hosts from 0 to p equipment. Constraints (4) ensure that the equipment's service capacity is respected. Constraints (5) and (6) make z_i to be 1 if there is at least one equipment installed in city i and

Table 1: Description of Parameters and Decision Variables.

	Problem Parameters
N	Set of candidate cities
S_i	Set of cities whose distance from city i is less or equal to R km, that is, $S_j = \{j \in N \mid d_{ij} \leq R$ and $d_{ji} \leq R\}$
d_{ij}	Distance from city i to city j
dem_j	Demand for mammography screenings from city j
cap	Annual screening capacity of each mammography unit
p	Number of mammography units to be allocated
R	Maximum distance women should travel to perform a mammography screening
$demMin$	Minimum annual screening demand that a city must have in order to host an equipment
	Decision variables
x_{ij}	Continuous variable into interval $[0,1]$ that indicates the fraction of demand from city j which is served by mammography units in the city i.
y_i	Integer variable that represents the amount of equipment installed in the city i.
z_i	Binary variable that assumes value 1 if i hosts an equipment and 0, otherwise.
t_i	Binary variable that assumes value 1 if demand of city i is fully served by the mammography units themselves and 0, otherwise.

0, otherwise. Constraints (7), (8), and (9) force t_i to be 1 if the demand of city i is fully met by its mammography units and 0, otherwise. Constraints (10) stipulate that a city i will only serve another city j if its demand just has been fully served by itself. Constraints (11) prevent a local with a demand less than the minimum demand hosts an equipment. Finally, Constraints (12), (13), (14), and (15) impose the domain of the decision variables.

The formulation earlier presented works in a scenario where there is freedom to relocate equipment. Although this strategy improves equipment distribution, in the real world it would be difficult to perform this operation. Hence, it makes sense to maintain the current distribution of already installed mammography units and to optimize the location of future equipment. In this alternative formulation, it is necessary to add the following parameter:

c_i = Number of mammography units currently installed in the city i

Finally, to conclude this new formulation, just add the following set of constraints:

$$y_i \geq c_i \quad \forall i \in \mathcal{N} \tag{16}$$

Constraints (16) determine that the number of mammography units installed in each city must be greater than or equal to the number currently installed in that one.

5 THE PROPOSED SIMULATED ANNEALING ALGORITM

This Section presents the proposed Simulated Annealing algorithm, developed for solving realistic instances of the MULP. It is organized as follows. Subsection 5.1 shows the representation of a solution for the MULP. Subsection 5.3 describes how the idle capacity of the equipment is distributed among the cities of a region. In Subsection 5.2, the procedure for creating an initial solution is presented. Subsection 5.4 shows as a solution is evaluated. The neighborhood structure for exploring the solution space of the MULP is discussed in Subsection 5.5. The details of the Simulated Annealing algorithm are presented in Subsection 5.6.

5.1 Solution Representation

A MULP solution s is represented by a pair (y,x), in which $y = (y_1, y_2, \cdots, y_n)$ is a n-vector, where y_i represents the number of mammography units installed at city i, and $x = (x_{ij})$ is a $(n \times n)$-matrix, where x_{ij} represents the fraction of the demand of city j that is covered by the city i. As an example, Figure 1 illustrates a solution $s = (y,x)$ to a problem with 4 cities and 2 mammography units.

Vector y

1	2	3	4
1	0	1	0

Matrix x

	1	2	3	4
1	1	0.3	0	0.9
2	0	0	0	0
3	0	0.6	1	0
4	0	0	0	0

Figure 1: Representation of a Solution with Two Mammography Units and Four Cities.

In Figure 1, positions 1 and 3 of the vector y indicate that the corresponding cities have one equipment each, while the others host none. In the matrix x, the cells x_{11} and x_{33} indicate that mammography units fully meet the demand of the locations where they are installed. Finally, the cells x_{14}, x_{12} and x_{32} indicate that city 1 covers 90% of the demand of city 4 and 30% of the demand of city 2. On the other hand, city 3 covers 60% of the demand of city 2.

5.2 Initial Solution

Simulated Annealing (SA) is a metaheuristic that does not depend on a good initial solution to generate good results. Therefore, we use a random initial solution for solving MULP with this metaheuristic. The process starts by generating a subset of candidate locations to host equipment, which consists of all cities that have the demand equal to or greater than the minimum demand. The next step is to choose at random one of these cities to host the equipment. Once this is done, all mammography units are assigned to this location. Finally, the idle capacity of these mammography units should be used to serve other cities within a radius of R kilometers, as described in Subsection 5.3.

5.3 Idle Screening Capacity

The idle capacity of the equipment in a given city i can be used to cover the demand of one or more cities. All cities that are no more than R kilometers from city i are candidates to have their demand, or part of it, attended by this city. Among the candidate cities, the priority will be given to those who have no other cover option than the city i. For any city to be a potential choice for another, it must be within a radius of R kilometers and have a minimum demand to host a mammography unit.

If the sum of residual demands from priority cities surpass the idle capacity of city i, the allocation preference is given to cities with lower residual demand. The residual demand of a city is the difference between the total number of examinations required by it and the portion of this demand that is already covered. Finally, if even after meeting all residual demand from priority cities, there is still idle capacity at city i, so the other candidate cities will be analyzed, giving preference again to those with lower residual demand.

5.4 Solution Evaluation

A MULP solution is evaluated according to Equation (1), where x_{ij} represents the fraction of the demand from city j covered by city i.

5.5 Neighborhood Structure

The solution space of the MULP is explored by relocation moves. This move consists of removing the equipment from a city i and installs it in a different city j, assuming that j has demand equal to or greater than $demMin$. Eventual idle screening capacity in city j is handled as described in Subsection 5.3. The set of neighbors s' of a solution s generated by this type of move forms the neighborhood $\mathcal{N}(s)$.

Relocating the equipment requires attention, since the demand of the city i, in which the equipment was removed, may be fully or partially uncovered. Consequently, the care service for its neighboring cities may be compromised. On the other hand, the city that receives the equipment may have all or part of its demand attended, and it may serve other neighboring cities, both partially and fully.

Algorithm 1 shows a relocation move from a city i to a city j in a given solution. Initially, at lines 2 and 3, the $RetHubs(i,x)$ and $RetHubs(j,x)$ functions are called, returning lists lst_i and lst_j, which represent the cities that serve the cities i and j, respectively. In line 4, the number of equipment installed in the city i is decreased by one unit. Subsequently, the $UnserveNeighborhood(i,x)$ and $UnserveNeighborhood(j,x)$ functions are called for eliminating the links of the cities i and j with cities in their regions, respectively. In line 7, the number of equipment in city j is increased by one unit. In line 8, it is verified if the list of locations that serves city i is empty. For each city in the list lst_i, the $ServeNeighborhood(i,x)$ function is called at line 11 to update the links of this city to cities in its region. In line 15, the links of city i to cities in its region are updated. The same procedure for city i started at line 8 also is applied for city j. In this case, the number of elements of lst_j is checked at line 16, the $ServeNeighborhood(j,x)$ function is called at line 19, and the links of city j to cities in its region are updated at line 23.

5.6 Simulated Annealing

The Simulated Annealing (SA) metaheuristic (Kirkpatrick et al., 1983) makes an analogy with the metal annealing process. In this process, the metal is heated to a high temperature, and then cooled slowly so that the final product is a homogeneous mass (Haeser and Ruggiero, 2008).

Unlike traditional descent methods, where neighboring solutions are accepted only if they generate an improvement in the objective function, SA also accepts worsening solutions, according to the probabil-

Algorithm 1: Relocation Move.

procedure RELOCATE(i, j, y, x)
    $lst\_i \leftarrow RetHubs(i, x)$;
    $lst\_j \leftarrow RetHubs(j, x)$;
    $y[i] \leftarrow y[i] - 1$;
    $UnserveNeighborhood(i, x)$;
    $UnserveNeighborhood(j, x)$;
    $y[j] \leftarrow y[j] + 1$;
    if $(lst\_i > 0)$ **then**
        $it \leftarrow lst\_i.begin()$;
        while $(it \neq lst\_i.end())$ **do**
            $ServeNeighborhood(it, x)$;
            $it++$;
        end while
    end if
    $ServeNeighborhood(i, x)$;
    if $(lst\_j > 0)$ **then**
        $it \leftarrow lst\_j.begin()$;
        while $(it \neq lst\_j.end())$ **do**
            $ServeNeighborhood(it, x)$;
            $it++$;
        end while
    end if
    $ServeNeighborhood(j, x)$;
end procedure

ity function (Dowsland, 1993) given by Eq. (17):

$$P(\Delta, T) = e^{-\Delta/T} \tag{17}$$

where $P(\Delta, T)$ is the probability of accepting a move, Δ is the variation in the value of the objective function (in this case, a decrease variation), and T is the current temperature.

The idea of SA is to start from a high initial temperature, and, as the method progresses, it will cool until it reaches a freezing value at the end of the procedure. The pseudocode of the Simulated Annealing metaheuristic is described in Algorithm 2.

Algorithm 2: Simulated Annealing.

1: **procedure** SA$(f(s), \mathcal{N}(s), SAmax, \alpha, T_i, T_{zero}, s)$
2:     $s^\star \leftarrow s$;
3:     $T \leftarrow T_i$;
4:     **while** $(T > T_{zero})$ **do**
5:         $IterT \leftarrow 0$;
6:         **while** $(IterT < SAmax)$ **do**
7:             $IterT++$;
8:             Choose $s' \in \mathcal{N}(s)$ at random
9:             $\Delta = f(s) - f(s')$;
10:             **if** $\Delta < 0$ **then**
11:                 $s \leftarrow s'$
12:                 **if** $f(s) > f(s^\star)$ **then**
13:                     $s^\star \leftarrow s$
14:                 **end if**
15:             **else**
16:                 Generate $x \in [0, 1]$ at random;
17:                 **if** $(x < e^{-\Delta/T})$ **then**
18:                     $s \leftarrow s'$
19:                 **end if**
20:             **end if**
21:         **end while**
22:         $T \leftarrow \alpha \cdot T$;
23:     **end while**
24:     Return $s^\star$;
25: **end procedure**

In line 2 of Algorithm 2, the current solution is saved in a variable $s^\star$, which represents the best solution obtained so far. In line 4, it is checked if the temperature has reached the freezing value. If this value has not been reached, there will be *SAmax* (input parameter) iterations (line 6), where neighbors of the current solution (line 8) will be generated at random. Each neighbor is evaluated at line 9 according to Eq. (1). In line 10, it is checked if there was an improvement and, if so, the neighboring solution becomes the current solution (line 12). If the new generated solution is the best solution obtained so far, it is stored in the variable $s^\star$ (line 13). In the case of the neighboring solution has not generated improvement, the probability test is performed at line 17. If this worsening solution is accepted, it becomes the new current solution. In line 22, the current temperature is cooled based on a factor α, received as an input parameter. Finally, the method returns the best solution $s^\star$ (line 24) found during the search.

The initial temperature (T_i) of Algorithm 2 is determined by simulation (Gomes Júnior et al., 2005). Their procedure works as follows. Given a random initial solution and a low temperature (T_i) as the initial temperature, the procedure verifies how many solutions are accepted in *SAmax* iterations. If $\gamma \times SAmax$ solutions are accepted, then this current temperature is the starting temperature; otherwise, the temperature is increased by a rate $\beta > 1$. Thus, the initial temperature is that at which $\gamma\%$ of the solutions are accepted at the beginning of the cooling process.

6 COMPUTATIONAL EXPERIMENTS

The mathematical programming model was implemented using CPLEX solver, version 12.7.1. The proposed Simulated Annealing algorithm was developed in C++ language. All tests were performed on a Intel

Core i5-4200U 1.6 GHz CPU computer with 6 GB of RAM under the Ubuntu 16.04.2 operating system.

Subsection 6.1 shows the characteristics of the instances used for testing the methods. In Subsection 6.2, a comparison between the results obtained by the Simulated Annealing algorithm and the CPLEX solver is introduced. In the following, Subsection 6.3 presents an analysis of the best location of new mammography units to be acquired. A comparison is made between a scenario where there is freedom to relocate all existing mammography units and another one where only new equipment to be acquired can have their locations optimized.

All solutions obtained through both CPLEX solver and Simulated Annealing algorithm, and the complete results are available at http://www.decom.ufop.br/prof/marcone/projects/MULP/MULP.html.

6.1 Instance Characteristics

For testing the methods, 8 instances referring to the projection of the female population in the age group indicated for mammography screenings in 2020 were used. Table 2 shows the characteristics of them.

Table 2: Instance Characteristics.

Instance Name	# Equip.	Min. Demand	Screening Capacity
RO-8-1800-5069	8	1800	5069
RO-8-1800-6758	8	1800	6758
MG-344-375-5069	344	375	5069
MG-258-375-6758	258	375	6758
MG-324-375-5069	324	375	5069
MG-324-375-6758	324	375	6758
MG-324-0-5069	324	0	5069
MG-324-0-6758	324	0	6758

In the first column of Table 2, the name of each instance indicates its properties in the form "State-Quantity of Mammography Units-Minimum Demand-Productivity". The second, third, and fourth columns represent, respectively, the number of equipment, the minimum demand (*demMin*), and the annual screening capacity of the equipment. For example, instances "RO-8-1800-5069" and "RO-8-1800-6758" refer to the Rondônia State, Brazil. In these two instances, the number of mammography units is equal to 8, and the minimum demand for screenings that a city needs to have to host equipment is 1800. The difference between them is due to the productivity of the equipment. In the first one, a mammography unit is able to perform 5069 mammography screenings per year, while in the other one, this number is 6758.

The other 6 instances refer to the Minas Gerais State. The "MG-344-375-5069" and "MG-258-375-6758" instances have dummy data on the number of existing mammography units. In these two instances, the number of mammography units varies according to the screening capacity of them. The quantities of 344 and 258 equipment are obtained by dividing the total demand, equivalent to 1739432 screenings, by the productivity of the equipment. In the last 4 instances, the number of equipment corresponds to the current scenario, totaling 324 mammography units. The difference between them is due to the screening capacity and the minimum demand considered.

Instances where the minimum demand is 0 (zero) have been created to analyze scenarios containing cities with low demand and located in remote regions, without coverage. In these cases, we simulated a scenario in which any city could host the equipment.

We carry out the projection to the female population indicated for the mammography screenings in 2020 as follows. The female population of each city was determined from Census 2010 (Brasil, 2010). We assume that the population aged from 30 to 39 years old in 2010 represents the population aged from 40 to 49 years old in 2020. The same holds true for the population aged from 50 to 69 years old, which, in 2020, is represented by the population aged from 40 to 59 years old in 2010. Hence, to estimate the demand of each city is possible, according to the percentages recommended for the two age groups of the female population (see Section 3). We observed that the first two instances refer to the database of Rondônia State, Brazil, and the last six to the database of Minas Gerais State, Brazil. In Rondônia, the estimated demand for screenings is 73900, while in Minas Gerais is 1739432 screenings.

The existing number of mammography units of the real instances was calculated using information from the DATASUS website (Brasil, 2019) for August 2019. The annual screening capacity for each equipment was set to 5069 screenings by the Brazilian National Cancer Institute (INCA), on November 1, 2015 (INCA, 2015). However, in 2017, this capacity was changed to 6758 ones (Brasil, 2017). The difference between these two estimates is related to the screening amount that can be performed per hour. To reach 5069 screenings per year, a mammography unit must be available 80% of the time for 22 working days in each of the 12 months of the year, with 8 working hours a day and 20 minutes for each screening. On the other hand, if we assumed that a mammography screening is done in 15 minutes, then 6758 mammography screenings can be performed annually.

The parameter R, which aims to set the maximum distance between two cities that have a service relationship, was set to 60 km, following recommendations from the Brazilian Ministry of Health (Brasil, 2017). The distances between cities, which are a key issue for defining eligibility for a service, were taken from the *Google Maps API*.

Finally, was established that a city needs to have a minimum demand for hosting an equipment. Analyzing the current distribution of mammography units in Rondônia and Minas Gerais, we defined, as the minimum demand parameter, the demand from the city that, among those that already have the equipment installed, has the smallest female population within the age range indicated for mammography screenings. In Minas Gerais, this value was set to 375 and in Rondônia to 1800.

6.2 Simulated Annealing based Solution

In this Subsection, we report the results of applying the CPLEX solver and the Simulated Annealing algorithm for solving the instances shown in Table 2. The CPLEX solver ran with a timeout of 3600 seconds for each instance. The results of the Simulated Annealing algorithm were based on 30 executions for each instance.

The value of the parameter *SAmax* was set to a number proportional to the instance dimension. After preliminary tests, we set $SAmax = 20 \times n$, where n is the number of cities of the instance. The other SA parameters are the cooling rate α and the freezing temperature T_{zero}. After testing, $\alpha = 0.99$ and $T_{zero} = 0.1$ were used. Finally, the initial temperature was obtained by the simulation method, in which the low temperature to start the process was set to 500. For the other two parameters of the simulation method, we assign the values $\gamma = 0.95$ and $\beta = 2$.

Table 3 presents part of the results of the experiments comparing CLPEX and Simulated Annealing. The first column indicates the name of the instance. In the sequel, the column "CPLEX" is divided into four other columns. The first one indicates the execution time, in seconds. The results obtained, the upper bound and the GAP to the upper bound, are shown in sequence. The Simulated Annealing column also has four sub-columns. The column "Best" brings the best result found in 30 runs, while the column "Average" shows the average result of these runs. The average execution time is displayed in the column "Average Time". In column "GAP (Average)", it is presented the GAP of the SA average result regarding the upper bound of each instance, obtained via CPLEX.

Regarding the instances of the Rondônia State ("RO-8-1800-5069" and "RO-8-1800-6758"), the Simulated Annealing algorithm obtained the optimal solution in all 30 executions. For the Minas Gerais instances, the Simulated Annealing algorithm found values very close to those generated by CPLEX. In none of these instances, the gap concerning the upper bound reached 1%. For the "MG-324-0-6758", the optimal value was found. Regarding the instance "MG-324-375-6758", the average GAP is 0.001%. Even though this value may be considered small, it is important to emphasize that, in the best result obtained by SA, the global optimum was achieved.

Due to the timeout set to 3600 seconds, in three instances, CPLEX did not reach the optimal value. Comparatively, in these same instances, the Simulated Annealing algorithm achieved similar quality results with runtimes of less than 600 seconds.

6.3 New Equipment Acquisition Analysis

In the real world, the relocating of equipment may be impossible. Considering this reality, an analysis was made regarding the destination of p additional mammography units. Tables 4 and 5 bring the results of the acquisition proposals, having as the difference between them only the productivity of the equipment, with values equal to 5069 and 6758, respectively.

In each line of these tables, p mammography units are added to the current scenarios (instances "MG-324-375-5069" and "MG-324-375-6758"), which contain 324 mammography units. Column "Screening Capacity" represents the maximum mammography screenings possible to be done using the original number of mammography units plus those added. To reach this number, it is necessary to multiply the total quantity of equipment by the capacity for annual screenings of them. Column "Fixed" brings the results considering that previously installed mammography units cannot be relocated. On the other hand, the results in column "Free" assume that equipment can be freely relocated. For these scenarios, the tables show the number of covered women, the percentage of use of each equipment, and the obtained coverage rate.

Analyzing the results, we can note how inefficient the current distribution of equipment is. In Table 4, the current coverage rate does not reach 80% of the demand. If there is freedom to relocate the equipment, this coverage could reach 94.40%. If the considered productivity is equal to 6758 annual screenings (see Table 5), the disparity is slightly smaller, but 170197 (that is, 1738872 - 1568675) additional screenings could be performed.

The percentage of equipment utilization is ob-

Table 3: Comparison of CPLEX versus Simulated Annealing Algorithm Results.

Instance	CPLEX				Simulated Annealing			
	Time (sec)	Result	Upper Bound	GAP	Best	Average	Average Time (sec)	GAP (Average)
RO-8-1800-5069	<1	40552	40552	0.000%	40552	40552	8	0.000%
RO-8-1800-6758	<1	50621	50621	0.000%	50621	50621	8	0.000%
MG-344-375-5069	3600	1705656	1708758	0.182%	1698617	1695581	514	0.771%
MG-258-375-6758	356	1695966	1695966	0.000%	1684046	1680693	534	0.901%
MG-324-375-5069	3600	1642031	1642356	0.020%	1640576	1639375	537	0.182%
MG-324-375-6758	16	1738872	1738872	0.000%	1738872	1738860	414	0.001%
MG-324-0-5069	3600	1642152	1642356	0.012%	1641280	1640302	556	0.125%
MG-324-0-6758	16	1739432	1739432	0.000%	1739432	1739432	412	0.000%

Table 4: Results of the Acquisition Proposal of New Mammography Units with Productivity Equal to 5069.

Quantity Added	Screening Capacity	Fixed			Free		
		Covered	Equipment Use	Coverage Rate	Covered	Equipment Use	Coverage Rate
0	1642356	1383109	84.21%	79.51%	1642031	99.98%	94.40%
1	1647425	1388178	84.26%	79.81%	1645401	99.88%	94.59%
2	1652494	1393247	84.31%	80.10%	1650650	99.89%	94.90%
3	1657563	1398316	84.36%	80.39%	1655267	99.86%	95.16%
4	1662632	1403385	84.41%	80.68%	1658962	99.78%	95.37%
5	1667701	1408454	84.45%	80.97%	1663281	99.73%	95.62%
...	...	...	...	...	...	...	...
47	1880599	1621222	86.21%	93.20%	1736474	92.34%	99.83%
48	1885668	1625997	86.23%	93.48%	1736953	92.11%	99.86%
49	1890737	1630378	86.23%	93.73%	1737639	91.90%	99.90%
50	1895806	1634478	86.22%	93.97%	1737732	91.66%	99.90%
51	1900875	1638548	86.20%	94.20%	1737849	91.42%	99.91%
...	...	...	...	...	...	...	...
112	2210084	1738468	78.66%	99.94%	1738872	78.68%	99.97%
113	2215153	1738695	78.49%	99.96%	1738872	78.50%	99.97%
114	2220222	1738872	78.32%	99.97%	1738872	78.32%	99.97%
115	2225291	1738872	78.14%	99.97%	1738872	78.14%	99.97%

tained by dividing the number of women attended by the service capacity of all installed equipment. This data is of great interest to public decision-makers since efficiency is one of the objectives of public administration. In scenarios with the freedom to relocate equipment, as new equipment is inserted, the percentage of utilization decreases. On the other hand, when there is a need to maintain the current distribution of equipment, it is necessary to have new equipment so that the maximum utilization can be obtained. Considering the productivity of 5069 annual screenings, the maximum usage percentage is 86.23%, which is achieved when 48 or 49 new equipment is added. For productivity equal to 6758, the best use is obtained by adding 9 mammography units, reaching a percentage of use of 72.34%.

We may notice there is a limit regarding the coverage rate. In the first lines, as more mammography units are added, more women are served. However, even without covering 100% of the demand, there is a limit in which adding equipment does not increase coverage. In Table 4, this occurs when 115 mammography units are added. In this case, the number of covered women is the same when 114 equipment are added. On the other hand, in Table 5, this limit is 78. This fact occurs once there are cities that do not have the infrastructure to host equipment, and they are far from other cities that can serve them.

7 CONCLUSIONS

This paper dealt with the Mammography Unit Location Problem considering the restrictions established by the Brazilian Ministry of Health. Based on the formulation presented by (Souza et al., 2019), a new Mixed-Integer Linear Programming model was proposed. The advantage of this model is that it does not require the previous processing mentioned by these authors. The CPLEX solver, version 12.8, was used to implement the mathematical programming model. Data from Rondônia state and Minas Gerais state,

Table 5: Results of the Acquisition Proposal of New Mammography Units with Productivity Equal to 6758.

Quantity Added	Screening Capacity	Fixed			Free		
		Covered	Equipment Use	Coverage Rate	Covered	Equipment Use	Coverage Rate
0	2189592	1568675	71.64%	90.18%	1738872	79.42%	99.97%
1	2196350	1575433	71.73%	90.57%	1738872	79.17%	99.97%
2	2203108	1582191	71.82%	90.96%	1738872	78.93%	99.97%
3	2209866	1588949	71.90%	91.35%	1738872	78.69%	99.97%
4	2216624	1595707	71.99%	91.74%	1738872	78.45%	99.97%
5	2223382	1602465	72.07%	92.13%	1738872	78.21%	99.97%
...	...	...	...	...	...	...	...
8	2243656	1622185	72.30%	93.26%	1738872	77.50%	99.97%
9	2250414	1627885	72.34%	93.59%	1738872	77.27%	99.97%
10	2257172	1632660	72.33%	93.86%	1738872	77.04%	99.97%
...	...	...	...	...	...	...	...
75	2696442	1738433	64.47%	99.94%	1738872	64.49%	99.97%
76	2703200	1738660	64.32%	99.96%	1738872	64.33%	99.97%
77	2709958	1738872	64.17%	99.97%	1738872	64.17%	99.97%
78	2716716	1738872	64.01%	99.97%	1738872	64.01%	99.97%

Brazil, were used as a case study. Additionally, a Simulated Annealing based algorithm was also developed for treating large instances of the problem.

Two distinct experiments were performed. In the first one, the Simulated Annealing algorithm was compared with the CPLEX solver. The results showed that the proposed algorithm was able to provide good quality solutions in significantly shorter times. In the second experiment, we considered the acquisition of new equipment for two scenarios of the current instance of Minas Gerais state. In the first scenario, there is the possibility of relocating existing equipment; on the other hand, in the second scenario, this relocation is not allowed.

We indicate two issues that may motivate future work. The first one concerns the costs involved with the operation. In addition to the equipment itself, there are expenses with physical infrastructure to host it, as well as expenses with human resources to operate it. Scenarios with low utilization of equipment may indicate a poor use of public resources. Thus, the use of mobile units for mammography screenings may prove to be a solution. For this, it would be necessary to define criteria, besides the minimum demand, for a city to receive equipment. Furthermore, the mobile unit routing would also need to be optimized. The other issue to be addressed in future work concerns the distance traveled to attendance. For this, a multi-objective optimization formulation is necessary, incorporating the objectives of maximizing the covered demand, focus of this work, and also to minimize the sum of the distances traveled by women.

ACKNOWLEDGEMENTS

The authors thank Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG), and Universidade Federal de Ouro Preto (UFOP) for supporting this research.

REFERENCES

Ahmadi-Javid, A., Seyedi, P., and Syam, S. S. (2017). A survey of healthcare facility location. *Computers & Operations Research*, 79:223–263.

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

Berry, D. A., Cronin, K. A., Plevritis, S. K., Fryback, D. G., Clarke, L., Zelen, M., Mandelblatt, J. S., Yakovlev, A. Y., Habbema, J. D. F., and Feuer, E. J. (2005). Effect of screening and adjuvant therapy on mortality from breast cancer. *New England Journal of Medicine*, 353(17):1784–1792. PMID: 16251534.

Brasil (2010). População residente por sexo, situação e grupos de idade. Available at https://sidra.ibge.gov.br/tabela/200. Accessed on August 14, 2019.

Brasil (2017). Critérios e parâmetros assistenciais para o planejamento e programação de ações e serviços de saúde no âmbito do Sistema Único de Saúde. Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Regulação, Avaliação e Controle de

Sistemas. Available at https://bit.ly/2t6WvjG. Accessed on August 14, 2019.

Brasil (2019). Tabnet win32 3.0: CNES - Recursos Físicos - Equipamentos - Minas Gerais. Available at http://tabnet.datasus.gov.br/cgi/deftohtm.exe?cnes/cnv/equipoMG.def. Accessed on September 20, 2019.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424.

Church, R. and ReVelle, C. (1974). The maximal covering location problem. *Papers in regional science*, 32(1):101–118.

Corrêa, V. H. V., Lima, B. J. C., Silva-e Souza, P. H., Penna, P. H. V., and Souza, M. J. F. (2018). Localização de mamógrafos: um estudo de caso na rede pública de saúde. In *Anais do L Simpósio Brasileiro de Pesquisa Operacional*, Rio de Janeiro, Brasil. SOBRAPO.

Daskin, M. S. and Dean, L. K. (2005). Location of health care facilities. In Brandeau, M. L., Sainfort, F., and Pierskalla, W. P., editors, *Operations research and health care*, volume 70, pages 43–76. Springer, Boston, MA.

Davari, S., Kilic, K., and Naderi, S. (2016). A heuristic approach to solve the preventive health care problem with budget and congestion constraints. *Applied Mathematics and Computation*, 276:442–453.

Dogan, K., Karatas, M., and Yakici, E. (2019). A model for locating preventive health care facilities. *Central European Journal of Operations Research*, pages 1–31. Available at https://doi.org/10.1007/s10100-019-00621-4.

Dowsland, K. A. (1993). Some experiments with simulated annealing techniques for packing problems. *European Journal of Operational Research*, 68(3):389–399.

Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York.

Gomes Júnior, A. d. C., Souza, M. J. F., and Martins, A. X. (2005). Simulated annealing aplicado à resolução do problema de roteamento de veículos com janela de tempo. *TRANSPORTES*, 13(2):5–20.

Gu, W., Wang, X., and McGregor, S. E. (2010). Optimization of preventive health care facility locations. *International journal of health geographics*, 9(1):17.

Haeser, G. and Ruggiero, M. G. (2008). Aspectos teóricos de simulated annealing e um algoritmo duas fases em otimização global. *Trends in Applied and Computational Mathematics*, 9(3):395–404.

Hakimi, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations research*, 12(3):450–459.

Huff, D. L. (1964). Defining and estimating a trading area. *Journal of marketing*, 28(3):34–38.

INCA (2015). Revisão do parâmetro para cálculo da capacidade de produção de um mamógrafo simples. Available at https://bit.ly/2ZrgvJU. Accessed on August 14, 2019.

INCA (2017). Atlas on-line de mortalidade. Available at https://mortalidade.inca.gov.br/MortalidadeWeb/pages/Modelo01/consultar.xhtml. Accessed on September 30, 2019.

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

Miranda, S. M. R. and Patrocinio, A. C. (2018). Distribuição de mamógrafos por macrorregião do Brasil. In *Anais do V Congresso Brasileiro de Eletromiografia e Cinesiologia e X Simpósio de Engenharia Biomédica*, pages 433–436, Uberlândia. Even3. Available at https://doi.org/10.29327/cobecseb.78881.

Sathler, T. M., Conceição, S. V., Almeida, J. F., Pinto, L. R., de Campos, F. C. C., and Miranda Júnior, G. (2017). Problema de localização e alocação de centros de especialidades médias no estado de minas gerais. In *Anais do XLIX Simpósio Brasileiro de Pesquisa Operacional – XLIX SBPO*, pages 2988–2999, Blumenau, Brasil. SOBRAPO.

Silva, M. T. A. d., Silva Júnior, V. B. d., Mangueira, J. d. O., Gurgel Junior, G. D., and Leal, E. M. M. (2018). Distribution of mammograms and mammography offering in relation to the parametric care of the Public Health Care System in Pernambuco. *Revista Brasileira de Saúde Materno Infantil*, 18(3):609 – 618. Available at http://dx.doi.org/10.1590/1806-93042018000300009.

Souza, M. J. F., Penna, P. H. V., Stilpen, M., Rosa, P. M., Monteiro, J. C., and Lisboa, M. R. (2019). Localização de mamógrafos: formulações e estudo preliminar de caso de Rondônia. In *LI Simpósio Brasileiro de Pesquisa Operacional*, volume 2, Limeira. Galoá. Available at https://bit.ly/39sozia.

Toregas, C., Swain, R., ReVelle, C., and Bergman, L. (1971). The location of emergency service facilities. *Operations research*, 19(6):1363–1373.

Verter, V. and Lapierre, S. D. (2002). Location of preventive health care facilities. *Annals of Operations Research*, 110(1):123–132.

Weber, A. (1929). *Theory of the Location of Industries*. University of Chicago Press.

Witten, M. and 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., and Carvalho, C. d. C. (2016). Cobertura de mamografias, alocação e uso de equipamentos nas regiões de saúde. *Saúde em Debate*, 40(110):20 – 35.

Zhang, Y., Berman, O., Marcotte, P., and Verter, V. (2010). A bilevel model for preventive healthcare facility network design with congestion. *IIE Transactions*, 42(12):865–880.

Zhang, Y., Berman, O., and Verter, V. (2009). Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, 198(3):922–935.

Zhang, Y., Berman, O., and Verter, V. (2012). The impact of client choice on preventive healthcare facility network design. *OR spectrum*, 34(2):349–370.