

# MÉTODOS QUANTITATIVOS

PESQUISA OPERACIONAL  
VOLUME 3



Editora Poisson

Editora Poisson

Métodos Quantitativos  
- Pesquisa Operacional –  
Volume 3

1ª Edição

Belo Horizonte  
Poisson  
2018

Editor Chefe: Dr. Darly Fernando Andrade

Conselho Editorial

Dr. Antônio Artur de Souza – Universidade Federal de Minas Gerais  
Dra. Cacilda Nacur Lorentz – Universidade do Estado de Minas Gerais  
Dr. José Eduardo Ferreira Lopes – Universidade Federal de Uberlândia  
Dr. Otaviano Francisco Neves – Pontifícia Universidade Católica de Minas Gerais  
Dr. Luiz Cláudio de Lima – Universidade FUMEC  
Dr. Nelson Ferreira Filho – Faculdades Kennedy

Dados Internacionais de Catalogação na Publicação (CIP)

**M593**

**Métodos Quantitativos- Pesquisa Operacional  
Volume 3/ Organização: Editora Poisson -  
Belo Horizonte - MG : Poisson, 2018  
317p**

**Formato: PDF**

**ISBN: 978-85-7042-010-7**

**DOI: 10.5935/978-85-7042-010-7.2018B001**

**Modo de acesso: World Wide Web**

**Inclui bibliografia**

**1. Métodos Quantitativos 2. Pesquisa  
Operacional.3. Estatística I. Título**

**CDD-658**

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos seus respectivos autores.

[www.poisson.com.br](http://www.poisson.com.br)

[contato@poisson.com.br](mailto:contato@poisson.com.br)

# Sumário

|   |     |
|---|-----|
| <b>Capítulo 1:</b> Parameter tuning of the teaching-learning based optimization algorithm applied to troubleshooting.....   | 7   |
| Lucas Sousa de Oliveira, Takashi Yoneyama , Leonardo Ramos Rodrigues  |     |
| <b>Capítulo 2:</b> Otimização robusta multi-objetivo da potência de compressão e do Coeficiente de Desempenho (COP) de um sistema de refrigeração por compressão de vapor ..... | 18  |
| Patrícia Agnes Pereira da Silva, José Henrique de Freitas Gomes, Renata de Cássia Ferreira Silva, Flávio Vasconcelos da Silva   |     |
| <b>Capítulo 3:</b> Um estudo sobre a utilização de modelos multi-elos para gestão de peças de reposição .....   | 29  |
| João Paulo de Moraes Paludo, Takashi Yoneyama, Leonardo Ramos Rodrigues   |     |
| <b>Capítulo 4:</b> Identification of nonlinear systems using the teaching-learning based optimization algorithm .....   | 41  |
| Leonardo Ramos Rodrigues, Takashi Yoneyama, Lucas Sousa de Oliveira   |     |
| <b>Capítulo 5:</b> Um algoritmo híbrido baseado em otimização por nuvem de partículas e busca tabu para o problema da diversidade máxima.....                                   | 50  |
| Edison Luiz Bonotto, Lucídio dos Anjos Formiga Cabral   |     |
| <b>Capítulo 6:</b> Auxílio multicritério na priorização de riscos de projetos.....  | 65  |
| Ricardo Luiz Fernandes Bella  |     |
| <b>Capítulo 7:</b> O problema do caixeiro viajante com passageiros e lotação.....   | 76  |
| Ranmsés E. M. Bastos, Marco C. Goldberg, Elizabeth F. G. Goldberg   |     |
| <b>Capítulo 8:</b> Algoritmo genético para o problema de roteamento em arcos capacitado e aberto .....  | 88  |
| Rafael Kendy Arakaki, Fábio Luiz Usberti  |     |
| <b>Capítulo 9:</b> Meta-heurística ILS para o problema de posicionamento automático de pontos de acesso em ambientes internos .....   | 100 |
| Anderson Rufino, Thiago Gouveia da Silva, Eduardo Vieira Queiroga, Nailson dos Santos Cunha, Lucidio dos Anjos F. Cabral  |     |

# Sumário

|   |     |
|---|-----|
| <b>Capítulo 10:</b> Temperature field forecast in concrete dam with the use of Arima Models and the Finite Element Method.....  | 112 |
| Tásia Hickmann, Luiz Albino Teixeira Júnior, Alvaro Eduardo Faria, Samuel Bellido Rodrigues, Jairo Marlon Córrea, Everton Luís Garcia   |     |
| <b>Capítulo 11:</b> Aplicação da técnica de análise envoltória de dados e regressão logística: Análise da eficiência dos tribunais estaduais do Brasil.....                                   | 124 |
| Magno dos Santos Neto, Leandro André Cardoso de Souza, Pedro Paulo Gondim Simmer, Angela Cristina Gomes   |     |
| <b>Capítulo 12:</b> Improving the computational efficiency of teaching-learning based optimization for job shop scheduling problems by eliminating unnecessary objective function calls ..... | 137 |
| Leonardo Ramos Rodrigues, João Paulo Pordeus Gomes, Saulo Anderson Freitas de Oliveira , Lucas Sousa de Oliveira, Takashi Yoneyama  |     |
| <b>Capítulo 13:</b> Utilização do método multicritério Electre I como auxílio na seleção de fornecedores de uma empresa do ramo calçadista.....   | 149 |
| Jordana Ramalho de Sousa, Thiago Yuji Eguchi, Maria Madalena Guerra Ferreira, Samuel Natividade Ferreira Junior, Ruy Gomes Silva  |     |
| <b>Capítulo 14:</b> Abordagem estocástica com recurso em dois estágios aplicada a um problema de programação linear sujeito a incertezas .....  | 161 |
| Solange Regina dos Santos, Geovana Aparecida França dos Santos  |     |
| <b>Capítulo 15:</b> A comparative study of metaheuristics applied to troubleshooting optimization problems .....  | 170 |
| Lucas Sousa de Oliveira, Takashi Yoneyama , Leonardo Ramos Rodrigues  |     |
| <b>Capítulo 16:</b> Otimização do roteiro de distribuição de materiais no sistema único de saúde em Joinville/SC .....  | 182 |
| Flavia Hawerorth, Vanina Macowski Durski Silva, Ana Paula Nunes Duarte  |     |
| <b>Capítulo 17:</b> Modelagem de um processo produtivo utilizando redes neurais artificiais .....   | 190 |
| Simone Massulini Acosta, Anderson Levati Amoroso, Ângelo Márcio Oliveira Sant'Anna  |     |
| <b>Capítulo 18:</b> Desenvolvimento do menor caminho para a melhoria das linhas de ônibus em um bairro no município de Aracaju-SE .....   | 201 |
| Tayane Magalhaes Alvaia, Hellen Mariany Santos, Marcos Wandir Nery Lobao, José Ricardo Menezes Oliveira , Glaucia Regina de Oliveira Almeida  |     |

# Sumário

|  |     |
|--|-----|
| <b>Capítulo 19:</b> Otimização de escalas de tripulação de transporte público urbano com a meta heurística variable neighborhood search.....           | 210 |
| Eduardo Luis Festa, Christiane Wenck Nogueira Fernandes, Silvia Lopes de Sena Taglialenha  |     |
| <b>Capítulo 20:</b> Modelos matemáticos para otimização do processo de montagem da cabine de pintura .....   | 223 |
| Aline Maki Igari, Diogo Kenji Nomura, Jessica Ingrid Sales, Michele Martins da Silva, Priscila Harumi Souza  |     |
| <b>Capítulo 21:</b> Roteirização de veículos versus performance térmica na distribuição de sorvetes .....  | 237 |
| William Daiti Kobo Takeshita, Gustavo Henrique Moresco, Vanina Macowski Durski Silva   |     |
| <b>Capítulo 22:</b> Aplicação da abordagem estocástica baseada em simulação de Monte Carlo [SMC]: um estudo de caso .....                              | 248 |
| Paulo Henrique dos Santos, Solange da Silva, Juliano Resende Bucchianeri, Nilson Carvalho da Mata, Thiago Moura de Moraes                              |     |
| <b>Capítulo 23:</b> Análise de eficiência das policlínicas do estado do ceará utilizando análise envoltória de dados com ponderação de resultados..... | 259 |
| Camila Campos Colares das Dores, Gerardo Valdisio Rodrigues Viana, Antônio Clécio Fontelles Thomaz   |     |
| <b>Capítulo 24:</b> Análise do compartilhamento de vias entre o modal rodoviário e ferroviário por meio de simulação de eventos discretos.....         | 268 |
| Laura Luiza Pereira David, Gisele Tessari Santos   |     |
| <b>Capítulo 25:</b> Aplicação da técnica de simulação computacional em uma empresa do setor de panificação .....                                       | 279 |
| Jarbas Ancelmo da Silva Júnior, Natália Cibele de Sousa Santos, Nayara Felício de Oliveira, Lara Cristine Pereira dos Santos                           |     |
| <b>Capítulo 26:</b> Abordagens grasp aplicadas ao problema Quota CaRS .....  | 291 |
| Matheus da Silva Menezes, Marco César Goldberg, Elizabeth Ferreira Goldberg, Vanessa Elionara Souza Ferreira, Gabriela Colaço                          |     |
| <b>Autores:</b> .....  | 302 |

# Capítulo 1

## *PARAMETER TUNING OF THE TEACHING-LEARNING BASED OPTIMIZATION ALGORITHM APPLIED TO TROUBLESHOOTING*

*Lucas Sousa de Oliveira*

*Takashi Yoneyama*

*Leonardo Ramos Rodrigues*

**Abstract:** Teaching-Learning Based Optimization (TLBO) is a novel metaheuristic, population-based optimization method that simulates the teaching environment. It has recently shown great results in different applications due to its efficiency and accuracy. Parameter tuning is recognized as a crucial step when designing an optimization algorithm. The performance of optimization algorithm is very sensitive to parameter values. Therefore, finding a good parameter setting can be the difference between success and failure. In this context, TLBO arises as a good alternative because it simplifies the parameter tuning and control processes due to its reduced number of parameters. The aim of this paper is to propose a parameter tuning method for the TLBO algorithm, considering both the accuracy and the speed. Numerical experiments are carried out to illustrate the proposed parameter tuning method in troubleshooting optimization problems.

**Keywords.** TLBO. Parameter Tuning. Combinatorial Optimization. Troubleshooting.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2016)

## 1. INTRODUCTION

Evolutionary algorithms (EA) are powerful in the way they efficiently provide good solutions for real-world problems [Simon, 2013]. They have been successfully used when the search space is too large, the problem is too complex, the evaluation function is too complicated, the solution is too constrained, and/or there is no hint of how to solve the problem [Michalewicz and Fogel, 2004]. These algorithms though suffer from their generality, i.e., their lack of restrictions on their parameters. Therefore, it is always necessary to finely tune the parameters such that computational resources and algorithm accuracy are not affected. The Genetic Algorithm, for example, has a high sensitivity to mutation probability, crossover probability and population size [Pinel et al., 2012; Srinivas et al., 2014]. The selection of good parameters is usually done by one of these two classes of methods: parameter tuning, or parameter control [Eiben and Smit, 2012]. The former denotes the set of methods which define the parameters before the optimization is run, while the latter denotes the methods in which the parameters are dynamically defined during the execution of the optimization algorithm. Parameter control methods can still be categorized into deterministic, adaptive, or self-adaptive. It is important to note that it can hardly be said that a set of good parameters for a given problem will be generally good. Parameters are usually mutually sensitive and very dependent on the models and problems at hand. Those reasons strengthen the need for parameter control, even though it is harder to achieve [Wolpert and Macready, 1997; Nannen and Eiben, 2006]. An algorithm which has been recently proposed by Rao et al. [2011], the Teacher-Learner Based Optimization algorithm (TLBO), needs considerably less parameters: only the population size, besides the stop criteria. It is based on the teaching-learning process observed in a classroom and simulates the influence of a teacher on the output of a group of students in a class. TLBO has achieved remarkable performances in different types of problems such as con-strained [Rao and Patel, 2012], and unconstrained [Rao and Patel, 2013] optimization problems. TLBO has also been successfully used in other combinatorial optimization problems such as the flow shop and the job shop scheduling

problems [Baykasoglu et al., 2014], the set covering problem [Crawford et al., 2015] and the disassembly sequence planning problem Xia et al. [2013]. This paper aims to explore the optimization capabilities of the TLBO algorithm and the relationship between its parameters and computational time and accuracy. The optimization problem used to guide this study is the troubleshooting optimization problem described in section 3. Troubleshooting optimization problem is a combinatorial optimization problem which is known to be NP-hard for most of real applications [Vomlelova, 2003]. The remaining sections of this paper are organized as follows. Section 2 presents the basic principles of the TLBO algorithm. Section 3 describes the troubleshooting problem under consideration. Section 4 illustrates the application of the proposed method in three different troubleshooting models. Concluding remarks are presented in section 5.

## 2. TEACHING-LEARNING BASED OPTIMIZATION ALGORITHM

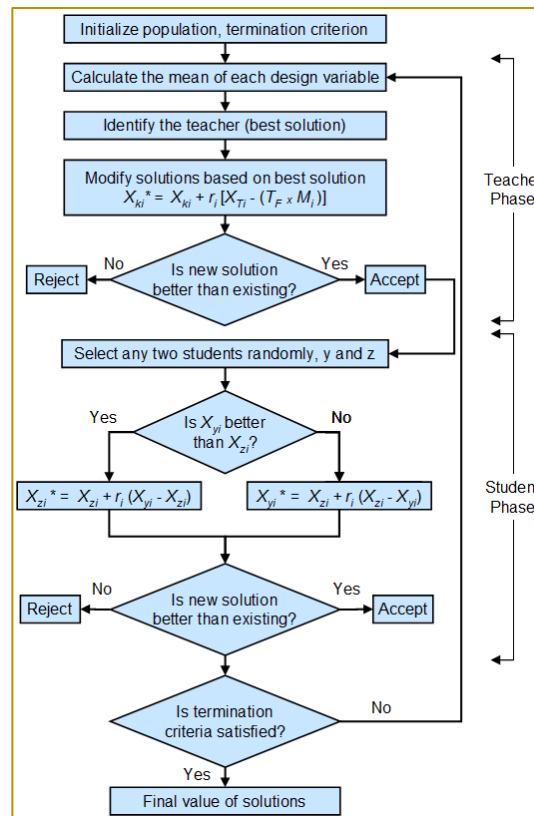
The TLBO algorithm is divided into two main parts: the Teacher Phase and the Student Phase, which is also known as the Learner Phase [Rao and Patel, 2013]. During the Teacher Phase, students learn from the teacher, while in the Learner Phase students learn through the interaction among themselves.

There is a solution  $X$  associated with each student, which corresponds to a possible solution to the optimization problem under consideration. Also, there is a result  $f(X)$  associated with each solution (or student), which can be obtained by evaluating the solution  $X$  using the objective function  $f$ . In the troubleshooting problem considered in this paper, a solution  $X$  corresponds to a troubleshooting strategy  $S$  and the associated result  $f(X)$  corresponds to its Expected Cost of Repair, denoted by  $ECR(S)$  [Vianna et al., 2016]

The student with the best solution in the population is considered as the Teacher. The steps for implementing the TLBO algorithm are presented in Figure 1. The Teacher Phase and the Student Phase are described in the next sections.



Figure 1: TLBO algorithm



### Teacher Phase

In this phase, the algorithm simulates the learning of the students from the teacher (best solution). During this phase, the teacher makes an effort to increase the mean result of the class.

Consider a group of  $n$  students. Let  $M_i$  be the mean solution of the students and  $T_i$  be the

teacher at iteration  $i$ . The teacher  $T_i$  will make an effort to move  $M_i$  to its own level. Knowledge is gained based on the quality of the teacher and the quality of students in the class. The difference  $D_i$  between the solution of the teacher,  $X_{Ti}$ , and the mean solution of the students,  $M_i$ , can be expressed according to Equation (1):

$$D_i = r_i (X_{Ti} - T_F \cdot M_i) \quad (1)$$

where  $r_i$  is a random number in the range  $[0,1]$  for iteration  $i$  and  $T_F$  is a teaching factor for

iteration  $i$ , which is randomly set to either 1 or 2 according to Equation (2):

$$T_F = \text{round}(1 + \text{rand}(0,1)) \quad (2)$$

Based on the difference  $D_i$ , the existing solution of student  $k$  in iteration  $i$ ,  $X_{ki}$ , with  $k \in$

$\{1, 2, \dots, n\}$ , is updated in the teacher phase according to Equation (3):

$$X_{ki}^* = X_{ki} + D_i \quad (3)$$

where  $X_{ki}^*$  is the updated value of  $X_{ki}$ . If  $f(X_{ki}^*)$  is better than  $f(X_{ki})$ ,  $X_{ki}^*$  is accepted

and replaces  $X_{ki}$ . Otherwise,  $X_{ki}^*$  is discarded and  $X_{ki}$  is not changed for the next iteration.

### Student Phase

In this phase, the algorithm simulates the learning of the students through interaction with one another. During this phase, students gain knowledge by discussing with other students who have more knowledge [Rao and Patel, 2013].

Consider a pair of students  $y$  and  $z$ . Let  $X_{yi}$  and  $X_{zi}$  be the solutions of students  $y$  and  $z$  at iteration  $i$ , respectively. If  $f(X_{yi})$  is better than  $f(X_{zi})$ , the solution of student  $z$  is updated

$$X_{zi}^* = X_{zi} + r_i(X_{yi} - X_{zi}) \quad (4)$$

$$X_{yi}^* = X_{yi} + r_i(X_{zi} - X_{yi}) \quad (5)$$

At the end of each iteration, the stop criteria must be checked. Different stop criteria maybe adopted. Some of the most commonly adopted stop criteria are the maximum number of iterations, the maximum number of successive iterations without any improvement and the maximum simulation time. In this paper, the maximum number of iterations is adopted as the stop criterion, as in other applications of TLBO in combinatorial problems [Baykasoglu et al., 2014], [Crawford et al., 2015], [Patil, 2016].

### 3. TROUBLESHOOTING PROBLEM

Troubleshooting is the name given for the sequence of actions performed in order to fix a given system. These actions are commonly separated into two categories: diagnostic and repair [Vomlelova, 2003].

The first type of action, usually less expensive and time consuming, can be inconclusive, i.e.,

according to Equation (4). Then, if  $f(X_{zi}^*)$  is better than  $f(X_{zi})$ ,  $X_{zi}^*$  is accepted and replaces  $X_{zi}$ . Otherwise,  $X_{zi}^*$  is discarded and  $X_{zi}$  is not changed for the next iteration.

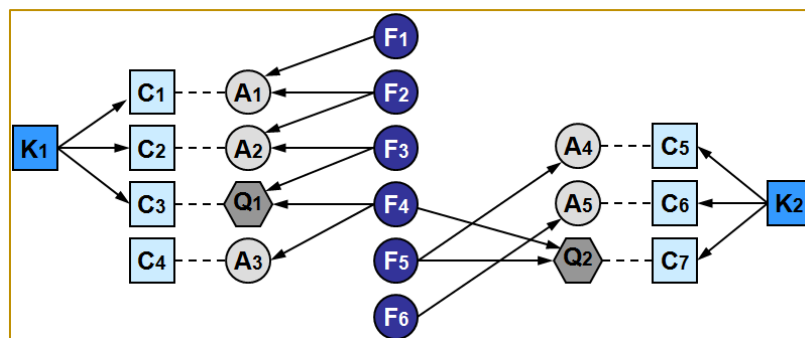
Similarly, if  $f(X_{zi})$  is better than  $f(X_{yi})$ , the solution of student  $y$  is updated according to Equation (5). Then, if  $f(X_{yi}^*)$  is better than  $f(X_{yi})$ ,  $X_{yi}^*$  is accepted and replaces  $X_{yi}$ . Otherwise,  $X_{yi}^*$  is discarded and  $X_{yi}$  is not changed for the next iteration.

it may represent some cost for no advancement toward an actual solution. Repair actions, on the other hand, are a step in the direction of solving the problem. Several actions might be required to fix a single fault, as well as several faults might have occurred simultaneously.

These types of problems are very practical and common in our daily lives. The decision-theoretic troubleshooting comes as a way to model and study our way of making decisions to optimize the troubleshooting process and thus minimize the costs involved.

A common way to represent this problem is through oriented graphs, such as the one presented in Figure 2. Nodes represent the possible failure modes ( $F_i$ ), diagnose questions ( $Q_i$ ) and repair actions ( $A_i$ ), while the oriented arcs represent the connections between failure modes and actions [Vianna et al., 2016].

Figure 1: Troubleshooting model 1 - Bayesian network representing a simple troubleshooting problem.



Costs are represented by  $C_i$ , while cost clusters are represented by  $K_i$ . A fault caused by failure mode  $F_i$  is only fixed when all repair actions connected to it are executed. A diagnostic question isolates a subset of possible failure modes. Each action and question has an associated cost, which is incurred if the action or question is executed. A cluster cost is incurred if at least one action or question connected to it is executed.

Some assumptions need to be made in order to ensure the optimality of the sequence found [Langseth and Jensen, 2001]:

1. Single Fault: only one fault can be present at the system at a time;
2. Perfect Repair: repair actions are always effective;
3. Fixed Cost: cost does not vary with time;
4. No Questions: no diagnose actions are allowed; and
5. Independent Actions: each action addresses exactly one fault.

Without these restrictions, there is no polynomial approximation algorithm for such problem [Lan, 2014]. Hence, the usage of a metaheuristic optimization algorithm is justified. In this paper, only the first three assumptions are considered such as to limit

the complexity of the problem at hand without distancing it too much from real world premises.

The solution for a troubleshooting problem is given in terms of a troubleshooting strategy. There is a sequence of actions (either repair or diagnostic) that minimize the Expected Cost of Repair (ECR) of a system according to the probability of each fault and the cost of each action. Figure 3 shows an example of a troubleshooting strategy.

4. NUMERICAL EXPERIMENTS

4.1. EXPERIMENT SETTING

This section presents numerical experiments to illustrate the application of the proposed method in three different troubleshooting models. Table 1 summarizes the main characteristics of each model used in the experiments, which are similar to those employed in a previous work involving the use of a simulated annealing strategy for optimization [Vianna et al., 2016]. The topology of troubleshooting models 1, 2 and 3 are shown in Figures 2, 4 and 5, respectively.

Table 2 presents the probability associated with each failure mode in each troubleshooting model. Table 3 shows the costs associated with each action, question and cluster.

Figure 3: Troubleshooting strategy for the model presented in Figure 2. Edges indicate whether the action was successful (true) or unsuccessful (false) and the probability this path is taken according to Table 2. The costs at the path ends are computed based on Table 3. The Estimated Cost of Repair for this strategy is  $ECR(\text{solution}) = 45 \cdot 30\% + 65 \cdot 20\% + 25 \cdot 10\% + 115 \cdot 15\% + 125 \cdot 20\% + 145 \cdot 5\% = 78.5$ .

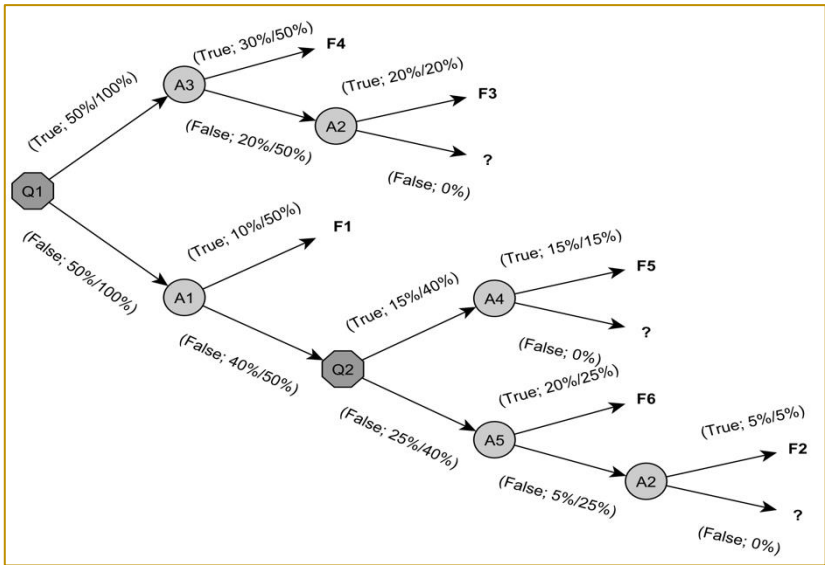


Table 1: Main characteristics of troubleshooting models

|           | Model 1 | Model 2 | Model 3 |
|-----------|---------|---------|---------|
| Failures  | 6       | 10      | 15      |
| Actions   | 5       | 10      | 16      |
| Questions | 2       | 4       | 6       |
| Clusters  | 2       | 4       | 6       |

Figure 4: Troubleshooting model 2

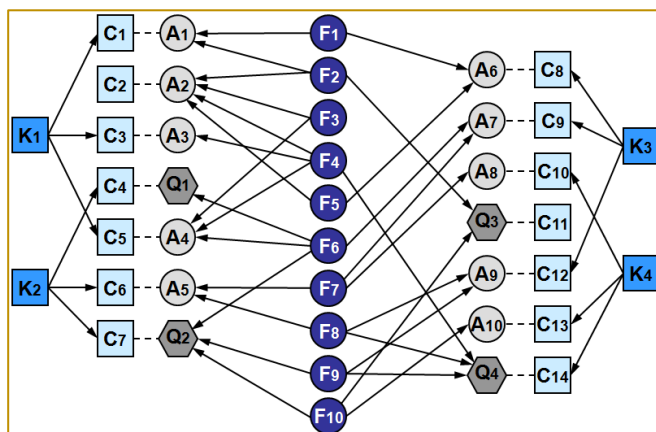


Figure 5: Troubleshooting model 3

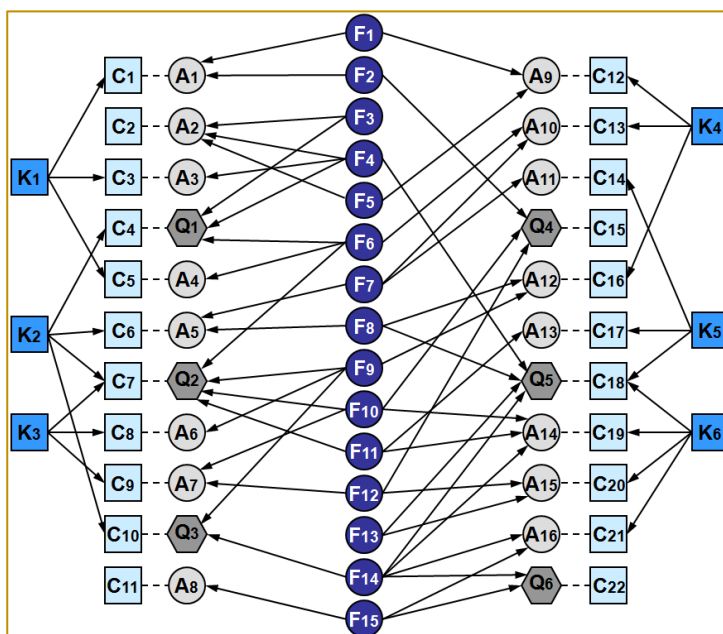


Table 2: Failure modes probabilities for each model

| Failure  | Model 1 | Model 2 | Model 3 |
|----------|---------|---------|---------|
| $F_1$    | 10%     | 1%      | 1%      |
| $F_2$    | 5%      | 5%      | 5%      |
| $F_3$    | 20%     | 10%     | 10%     |
| $F_4$    | 30%     | 15%     | 15%     |
| $F_5$    | 15%     | 8%      | 2%      |
| $F_6$    | 20%     | 4%      | 4%      |
| $F_7$    | -       | 13%     | 13%     |
| $F_8$    | -       | 17%     | 14%     |
| $F_9$    | -       | 16%     | 1%      |
| $F_{10}$ | -       | 11%     | 11%     |
| $F_{11}$ | -       | -       | 2%      |
| $F_{12}$ | -       | -       | 7%      |
| $F_{13}$ | -       | -       | 10%     |
| $F_{14}$ | -       | -       | 1%      |
| $F_{15}$ | -       | -       | 4%      |

Table 3: Actions, questions and clusters costs

| $i$      | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8 | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|----------|----|----|----|----|----|----|----|---|----|----|----|----|----|----|----|----|
| $C(A_i)$ | 10 | 20 | 30 | 40 | 50 | 40 | 35 | 5 | 20 | 15 | 20 | 30 | 70 | 10 | 20 | 15 |
| $C(Q_i)$ | 5  | 10 | 30 | 10 | 30 | 20 | -  | - | -  | -  | -  | -  | -  | -  | -  | -  |
| $C(K_i)$ | 10 | 40 | 20 | 50 | 10 | 30 | -  | - | -  | -  | -  | -  | -  | -  | -  | -  |

#### 4.2. OPTIMIZATION CONDITIONS

These models were optimized through a Matlab® implementation of TLBO. The experiments were run on a computer with a 4-core, 8-thread Intel Core i7 2.4 GHz processor and 12 GB of RAM. Both of these critical resources were monitored during the execution and did not reach 100%. Six parallel threads were used to speed up the simulations.

The population size was set to one of 15 values varying from 2 to 250 in a logarithmic, non-repeating progression. The maximum number of iterations allowed was set to 5.000, as a very large condition, so as to simulate an infinite amount of iterations. The statistic stability of the experiment was verified by a Monte Carlo approach with 50 repetitions for each parameter combination. Their ECR and simulation time for each iteration were registered for each repetition, allowing the analysis to be made post-simulation.

#### 4.3. RESULTS

As expected, the results presented in Figure 6 show that the algorithm converges faster for larger populations. This behavior was observed in all three models. Model 1 converges to the final solution before 10 iterations for populations larger than 89, for example. Also, note from Figure 6 that the

curves become closer to each other as the model size increases. It can be assumed from this result that population size keeps having a smaller impact on algorithm accuracy as population increases.

Figure 7 shows that the probability of the optimization to converge for model 1 is very low for population of 2, but reaches high levels of success (greater than 80%) for more than 8 learners through 1000 iterations. The same level of results are reached when 89 learners were used for model 2. Model 3 was too complex to return any statistically meaningful results for any setup proposed during the study, but it was kept as a limit case.

Figure 8 shows how the simulation time grows with population size for each model and how its standard deviation grows. The R2 (for linear regressions) shows a high probability that the time grows linearly with population size. The standard deviation also shows the algorithm's simulation time becomes more unpredictable even though its envelope remains linearly contained.

Combining all information in Figures 6, 7 and 8, it becomes obvious the influence the parameters have on the optimization outcome. Setting the parameters poorly may not only reduce the chances of getting to the optimal problem solution, but also consume a lot of time. Figure 9 summarizes the results of this paper, and provides a way to find better

parameters for models in the range considered. Note that the algorithm is nonlinear, and as such, the contour plots

might present noisy results. Its purpose as a guidance is not reduced by this fact though.

Figure 6: Mean Absolute Error of the Expected Cost of Repair for each model and each population size. Log-log axis used on the left to make it possible to visualize all models on same axis ranges. Linear-linear axis used on the right to show the exponential convergence of the algorithm.

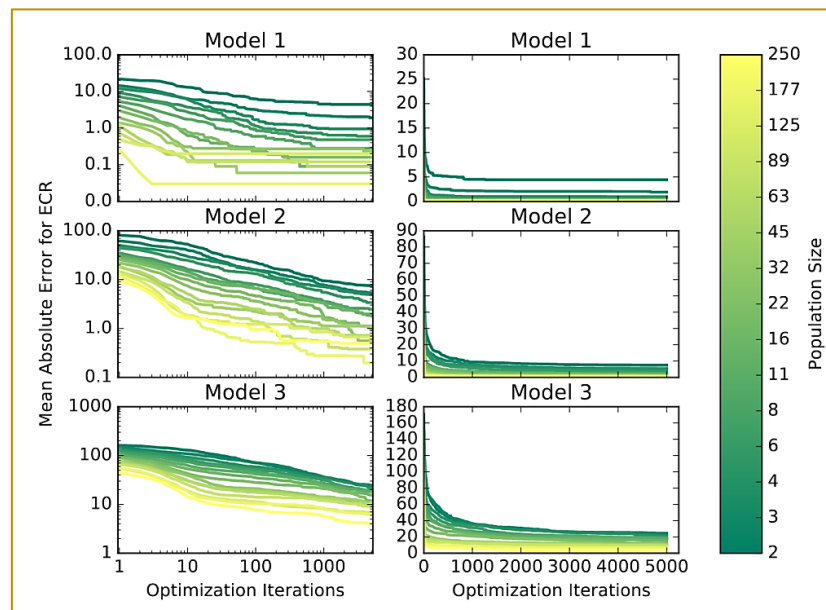
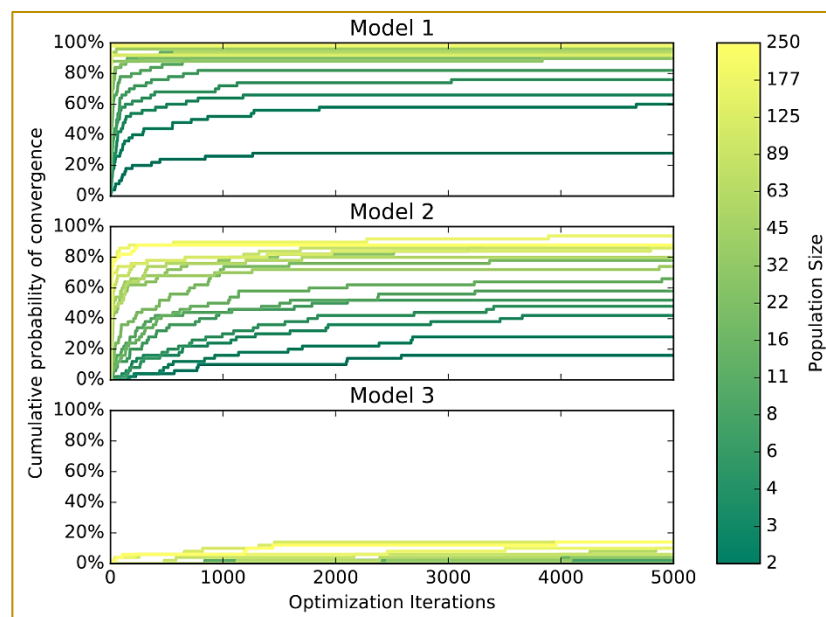


Figure 7: Cumulative probability that the simulation converged for a given amount of iterations and population size.



## 5. CONCLUSIONS

There is an optimal parameter setting for each troubleshooting model that runs through the TLBO algorithm. In this paper, it was shown

that the time needed to find a stable solution for a troubleshooting optimization problem varies linearly with the population size. The Mean Absolute Error (MAE) decays exponentially with the same parameter.

Figure 8: Computational time vs population size. Linear R2 are 0.9689, 0.9892 and 0.9898 for models 1, 2 and 3 respectively. Linear R2 for standard deviation is 0.9211, 0.9749, and 0.9995 for models 1, 2 and 3 respectively. Log-log axis used to make it possible to show full time range for all models.

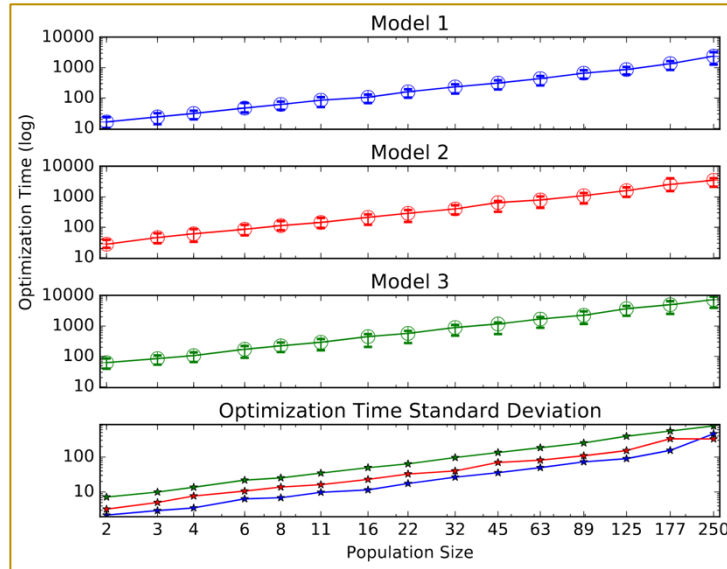
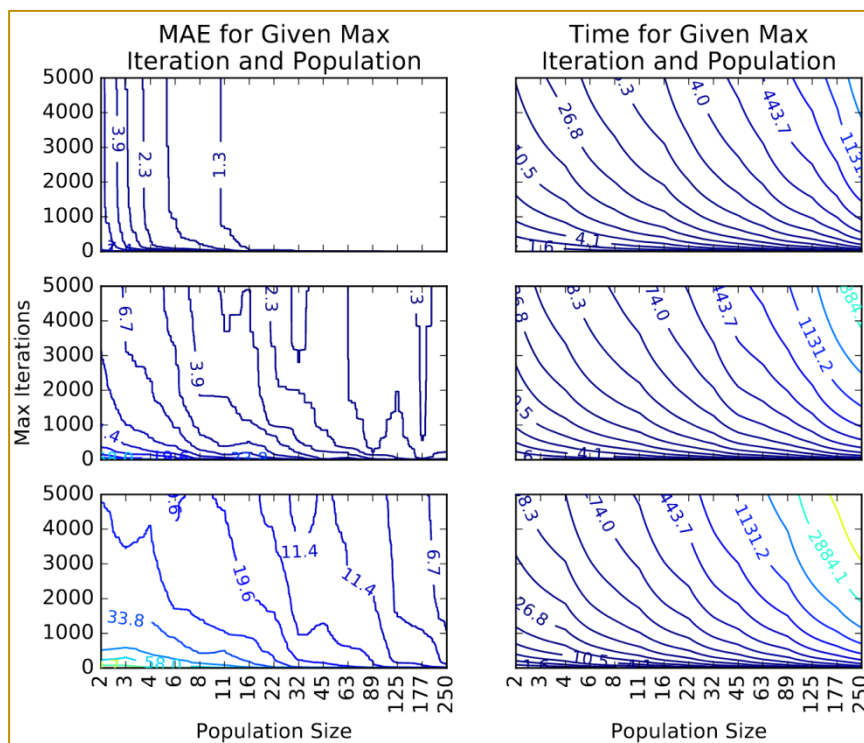


Figure 9: Relationship of maximum iteration and population size with mean absolute error and simulation time.



The use of an optimization algorithm to solve troubleshooting problems is justified only to models with dozens of fault nodes, i.e. bigger than model 2, since smaller models can be solved in a feasible time through extensive

search. It became clear how the troubleshooting problem becomes harder to solve as the model comes closer to real-life scales, to the point in which its accuracy cannot be guaranteed.

An extension of this work will focus on increasing the ranges up to large scales, while comparing the results with well-known evolutionary algorithms in order to investigate whether the limitations are in the TLBO algorithm or in the complexity of the problem under consideration. Other approaches such as hybrid evolutionary algorithms will also be considered to ensure the quality of the solution.

## REFERENCES

- [1]. Baykasoglu, A., Hamzadayi, A., and Köse, S. Y. (2014). Testing the performance of teaching learning based optimization (TLBO) algorithm on combinatorial problems: Flow shop and jobshop scheduling cases. *Information Sciences*, 276:204–218.
- [2]. Crawford, B., Soto, R., Leiva, F. A., Johnson, F., and Paredes, F. (2015). Problema del conjunto de cobertura resuelto mediante el algoritmo binario de optimización basado en enseñanza-aprendizaje. In *Conferencia Ibérica de Sistemas e Tecnologías de Informação*, p. 106–109, Agueda. IEEE.
- [3]. Eiben, A. E. and Smit, S. K. (2012). Autonomous Search, chapter Evolutionary Algorithm Parameters and Methods to Tune Them, p. 15–36. Springer Berlin Heidelberg, Berlin, Heidelberg. ISBN 978-3-642-21434-9. URL [http://dx.doi.org/10.1007/978-3-642-21434-9\\_2](http://dx.doi.org/10.1007/978-3-642-21434-9_2).
- [4]. Langseth, H. and Jensen, F. V. (2001). Heuristics for two extensions of basic troubleshooting. *Frontiers in Artificial Intelligence and Applications*, 66:80–89.
- [5]. Lan, V. (2014). Decision-theoretic troubleshooting: Hardness of approximation. *International Journal of Approximate Reasoning*, 55(4):977 – 988. ISSN 0888-613X. URL <http://www.sciencedirect.com/science/article/pii/S0888613X1300162X>. Special issue on the sixth European Workshop on Probabilistic Graphical Models.
- [6]. Michalewicz, Z. and Fogel, D. B. (2004). *How to Solve It: Modern Heuristics*. Springer, enlarged 2nd edition. ISBN 9783540224945. URL <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/3540224947>.
- [7]. Nannen, V. and Eiben, A. (2006). A method for parameter calibration and relevance estimation in evolutionary algorithms. In *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, GECCO '06*, p. 183–190, New York, NY, USA. ACM. ISBN 1-59593-186-4. URL <http://doi.acm.org/10.1145/1143997.1144029>.
- [8]. Patil, A. B. (2016). Teaching learning based optimization: Application and variation. In *Proceedings of the International Conference on Computing, communication and Energy Systems*, p. 1–5, Sangli.
- [9]. Pinel, F., Danoy, G., and Bouvry, P. (2012). Security and Intelligent Information Systems: International Joint Conferences, SIIS 2011, Warsaw, Poland, June 13-14, 2011, Revised Selected Papers, chapter Evolutionary Algorithm Parameter Tuning with Sensitivity Analysis, p. 204–216. Springer Berlin Heidelberg, Berlin, Heidelberg. ISBN 978-3-642-25261-7. URL [http://dx.doi.org/10.1007/978-3-642-25261-7\\_16](http://dx.doi.org/10.1007/978-3-642-25261-7_16).
- [10]. Rao, R. V. and Patel, V. (2012). An elitist teaching-learning based optimization algorithm for solving complex constrained optimization problems. *International Journal of Industrial Engineering Computations*, 3:535–560.
- [11]. Rao, R. V. and Patel, V. (2013). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. *Scientia Iranica*, 20:710–720.
- [12]. Rao, R. V., Vakharia, D. P., and Savsani, V. J. (2011). Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43:303–315.



- [13]. Simon, D. (2013). *Evolutionary Optimization Algorithms*. Wiley. ISBN 9781118659502. URL <https://books.google.com.br/books?id=gwUwIEPqk30C>.
- [14]. Srinivas, C., Reddy, B. R., Ramji, K., and Naveen, R. (2014). Sensitivity analysis to determine the parameters of genetic algorithm for machine layout. *Procedia Materials Science*, 6:866–876. ISSN 2211-8128. URL <http://www.sciencedirect.com/science/article/pii/S2211812814004696>. 3rd International Conference on Materials Processing and Characterisation (ICMPC 2014).
- [15]. Vianna, W. O. L., Rodrigues, L. R., Yoneyama, T., and Mattos, D. I. (2016). Troubleshooting optimization using multi-start simulated annealing. In *10th Annual IEEE Systems Conference*, Orlando, FL, USA, 18-21 April 2016.
- [16]. Vomlelov, M. (2003). Complexity of decision-theoretic troubleshooting. *International Journal of Intelligent Systems*, 18(2):267–277.
- [17]. Wolpert, D. H. and Macready, W. G. (1997). No free lunch theorems for optimization. *Trans. Evol. Comp.*, 1(1):67–82. ISSN 1089-778X. URL <http://dx.doi.org/10.1109/4235.585893>.
- [18]. Xia, K., Gao, L., Wang, L., Li, W., and Chao, K. M. (2013). A simplified teaching-learning-based optimization algorithm for disassembly sequence planning. In *Business Engineering (ICEBE), 2013 IEEE 10th International Conference on*, p. 393–398.

# Capítulo 2

## OTIMIZAÇÃO ROBUSTA MULTI-OBJETIVO DA POTÊNCIA DE COMPRESSÃO E DO COEFICIENTE DE DESEMPENHO (COP) DE UM SISTEMA DE REFRIGERAÇÃO POR COMPRESSÃO DE VAPOR

*Patrícia Agnes Pereira da Silva*

*José Henrique de Freitas Gomes*

*Renata de Cássia Ferreira Silva*

*Flávio Vasconcelos da Silva*

**Resumo:** Este trabalho buscou encontrar os parâmetros ótimos de um sistema de refrigeração por compressão de vapor para se obter os melhores resultados da potência de compressão e do coeficiente de desempenho do ciclo. Para isso, realizou-se a otimização simultânea de média e variância dessas respostas, utilizando o conceito de Projeto de Parâmetros Robusto. Os modelos matemáticos foram desenvolvidos através de um arranjo combinado que considerou três variáveis de controle (frequência do motor do compressor e as frequências dos motores das bombas do evaporador e condensador) e um ruído (carga térmica). Após a aglutinação das médias e variâncias através da equação do Erro Quadrático Médio, realizou-se otimização utilizando o Método do Critério Global e o método Gradiente Reduzido Generalizado.

**Palavras-Chave:** Sistema de refrigeração por compressão de vapor, Arranjo combinado, Otimização robusta

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2016)

## 1. INTRODUÇÃO

Equipamentos de refrigeração industrial são utilizados por diversas empresas industriais e comerciais, desde pequenos restaurantes e lanchonetes para conservação de alimentos e bebidas em câmaras frias até grandes indústrias de alimentos e bebidas ou químicas. Para Stanfield e Skaves (2010), suas aplicações podem ser divididas em quatro grupos: produção e distribuição de alimentos, processos químicos e industriais, aplicação especial e conforto por ar condicionado.

Atualmente, existem quatro sistemas principais de refrigeração: os sistemas por compressão de vapor, por absorção, por gás e por efeitos termoelétricos. Nos sistemas de absorção, o fornecimento de energia ocorre por um fluxo de calor; nos sistemas por compressão de vapor e por gás, o fornecimento de energia é por energia mecânica e nos sistemas termoelétricos é por energia elétrica [Vián e Astrain 2009].

Entre todos os tipos de sistemas de refrigeração, sistemas de refrigeração por compressão de vapor ainda são dominantes no mercado atual [Zhao et al. 2013a, Brito et al. 2014]. Esses sistemas são amplamente usados, em especial, nas instalações industriais para processamento e armazenagem de alimentos e em equipamentos de pequeno porte tais como refrigeradores, freezers e condicionadores de ar compactos [Salvador 1999].

É comumente observado que os sistemas de refrigeração são responsáveis por uma grande quantidade de consumo de energia elétrica em muitas plantas industriais [Widell e Eikevik 2010, Brito et al. 2014, Dincer e Kanoglu 2010, Mendes 2012]. Em alguns setores, principalmente alimentos, bebidas, produtos químicos é o processo que representa uma parte significativa dos custos globais de energia local (até 90% no caso de algumas instalações de armazenamento a frio) [Dincer 2003]. Em supermercados, segundo Qureshi e Zubair (2012), os sistemas de refrigeração constituem a maior área de uso de energia, sendo responsáveis por aproximadamente 50-60% do consumo total de energia.

De acordo com Manske (1999), a indústria de refrigeração industrial, historicamente, tem prestado muito pouca atenção na energia necessária para alcançar os objetivos dos processos de refrigeração. Assim, uma vez

que o ciclo de vapor de refrigeração de compressão consome a maior parte da energia em qualquer sistema de refrigeração, o esforço para reduzir o consumo de energia por meio de sistema de controle e otimização no sistema de refrigeração de compressão de vapor é de significado prático devido tanto à escassez de energia quanto às preocupações do aquecimento global [Zhao et al. 2013b].

Nas últimas décadas, a limitação dos recursos energéticos disponíveis e o contínuo aumento do custo da energia, bem como a crescente preocupação ambiental, têm levado a uma necessidade mundial de redução da demanda de energia que se traduz na pesquisa de meios cada vez mais eficientes para o uso da mesma [Salvador 1999].

Devido a esta constante e atual preocupação do consumo de energia pelas indústrias, pretende-se, através do método experimento, estudar o efeito das variáveis de controle e do ruído de um sistema de refrigeração por compressão de vapor sobre a eficiência energética. Este trabalho busca também identificar a combinação ótima dos parâmetros que oferecerão os melhores resultados quanto ao menor consumo de energia elétrica pelo sistema de refrigeração. Para tanto, foi utilizada a técnica DOE (*Design of experiments*) para o planejamento e análise dos experimentos e o software Minitab para análise dos resultados.

## 2. OTIMIZAÇÃO ROBUSTA BASEADA NA MODELAGEM DO ERRO QUADRÁTICO MÉDIO

Por otimização robusta entende-se como sendo a otimização simultânea de média e variância, utilizando o conceito de Projeto de Parâmetros Robusto (PPR). Desta forma, Montgomery (2005) define o Projeto de Parâmetros Robustos como uma abordagem para a análise e melhoria de processos que enfatiza a escolha dos níveis de fatores controláveis (ou parâmetros) em um processo para atingir dois objetivos: (1) para garantir que a média da resposta de saída esteja no nível ou alvo desejado e (2) para assegurar que a variabilidade em torno deste valor alvo seja tão pequena quanto possível.

Montgomery (2005) afirma que a Metodologia de Superfície de Resposta tem se mostrado uma técnica eficiente para modelagem e análise de dados em se tratando de Projeto

de Parâmetros Robustos (PPR), podendo seu método de análise ser desenvolvido a partir de arranjos cruzados ou arranjos combinados. Nesse sentido, o presente trabalho baseou sua execução experimental no arranjo combinado.

A utilização do arranjo combinado permite

analisar a influência das variáveis de controle e também das variáveis de ruído nas respostas estudadas. Para isso, as variáveis de ruído, para fins de experimentação, são tomadas como variáveis de controle e ambas são combinadas em um único arranjo experimental.

A Eq. (1) descreve o modelo desenvolvido a partir de um arranjo combinado [Montgomery 2005]:

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \sum_{i=1}^r \gamma_i z_i + \sum_{i=1}^k \sum_{j=1}^r \delta_{ij} x_i z_j + \varepsilon \quad (1)$$

Sendo:

- $y$  – Resposta de interesse
- $x_i$  – Variáveis de controle
- $z_i$  – Variáveis de ruído
- $\beta_0, \beta_i, \beta_{ii}, \beta_{ij}, \gamma_i, \delta_{ij}$  – Coeficientes a serem estimados
- $k$  – Número de variáveis de controle
- $r$  – Número de variáveis de ruído
- $\varepsilon$  – Erro experimental

Na Eq. (1), os coeficientes são estimados pelo Método dos Mínimos Quadrados Ordinários (*Ordinary Least Squares – OLS*). Uma vez estabelecido o modelo de superfície de

resposta representado na Eq. (1), a equação para a média da resposta  $y$  pode ser extraída diretamente do modelo combinado. Para isso, utiliza-se o seguinte critério:

$$\mu(y) = f(\mathbf{x}) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j \quad (2)$$

O modelo de variância é desenvolvido utilizando o princípio da propagação de erro, pela derivação:

$$\sigma^2(y) = \sum_{i=1}^r \left[ \frac{\partial y(\mathbf{x}, \mathbf{z})}{\partial z_i} \right]^2 \sigma_{z_i}^2 + \sigma^2 \quad (3)$$

Para a Eq. (3), Montgomery (2005) sugere adotar  $\sigma_{z_i}^2 = 1$  e a variável  $\sigma^2$  corresponde ao erro residual (*MS*) do modelo de superfície de resposta descrito na Eq. (1).

Com as equações de média e variância

desenvolvidas é possível combiná-las em uma única função conhecida como erro quadrático médio (*EQM*), apresentada por Köksoy (2006) como a soma da variância com a diferença quadrática entre a média da resposta e o seu valor alvo (Eq. 4).

$$EQM(y_j) = \left[ \mu(y_j) - T_{y_j} \right]^2 + \sigma^2(y_j) \quad (4)$$

Sendo:

- $EQM(y_j)$  – Erro quadrático médio para a  $j$ -ésima resposta
- $\mu(y_j)$  – Modelo estabelecido para a média da  $j$ -ésima resposta
- $T_{y_j}$  – Alvo da  $j$ -ésima resposta, obtido pela otimização individual de  $\mu(y_j)$
- $\sigma^2(y_j)$  – Modelo estabelecido para a variância da  $j$ -ésima resposta

A aglutinação das funções média e variância pela função  $EQM$  é feita para cada resposta, gerando novas equações que, por sua vez, serão aglutinadas para compor a função global do problema, que deverá ser minimizada. Assim, a otimização simultânea de médias e variâncias fornecerão valores próximos aos alvos com as menores variabilidades possíveis.

Para definir a formulação multi-objetivo do problema, o presente trabalho optou pelo Método do Critério Global (MCG), apresentado por Rao (1996) como um método eficaz quando se deseja a otimização de múltiplas respostas. Com isso, formulando os modelos para o erro quadrático médio de acordo com o MCG, a otimização robusta multi-objetivo pode ser obtida por meio da seguinte expressão:

$$\text{Minimizar } EQM_G = \sum_{j=1}^n \left[ \frac{T_{EQM(y_j)} - EQM(y_j)}{T_{EQM(y_j)}} \right]^2 \quad (5)$$

Sujeito a:  $x^T x \leq \alpha^2$

Sendo:

$EQM_G$  – Erro quadrático médio global

$EQM(y_j)$  – Erro quadrático médio para a  $j$ -ésima resposta

$T_{EQM(y_j)}$  – Alvo para o erro quadrático médio da  $j$ -ésima resposta, obtido pela minimização individual do modelo desenvolvido para  $EQM(y_j)$

$n$  – Número de respostas consideradas

$x^T x \leq \alpha^2$  – Restrição esférica para o espaço experimental

A determinação dos parâmetros ótimos é encontrada pela resolução do problema definido pela Eq. (5) através do emprego de métodos de otimização. Neste trabalho optou-se pelo uso do método Gradiente Reduzido Generalizado (GRG) [Rao 1996] para este propósito.

### 3. PROCEDIMENTO EXPERIMENTAL

Para a modelagem e otimização robusta do sistema de refrigeração por compressão de vapor, experimentos foram planejados e realizados através de um arranjo composto central (*Central Composite Design* - CCD)

baseado em um arranjo combinado, criado para quatro variáveis: três variáveis de controle ( $FC$ ,  $FBE$ ,  $FBC$ ) e uma variável de ruído ( $CT$ ).

Considerando que os níveis extremos das variáveis de ruído (pontos axiais) foram desconsiderados, conforme recomendado por Montgomery (2005), e adotando 7 experimentos para os pontos centrais, o arranjo combinado ficou então composto por 29 experimentos.

As variáveis de controle e de ruído com seus respectivos níveis de operação são apresentadas nas Tabelas 1 e 2.

Tabela 1 – Variáveis de controle e seus níveis

| Parâmetros                         | Unidade | Notação | Níveis |    |    |    |    |
|------------------------------------|---------|---------|--------|----|----|----|----|
|                                    |         |         | -2     | -1 | 0  | +1 | +2 |
| Frequência do compressor           | Hz      | $FC$    | 43     | 48 | 53 | 58 | 63 |
| Frequência da bomba do evaporador  | Hz      | $FBE$   | 33     | 40 | 47 | 54 | 61 |
| Frequência da bomba do condensador | Hz      | $FBC$   | 33     | 40 | 47 | 54 | 61 |

Tabela 2 – Variável de ruído e seus níveis

| Ruídos        | Unidade | Notação | Níveis |      |      |
|---------------|---------|---------|--------|------|------|
|               |         |         | -1     | 0    | +1   |
| Carga térmica | W       | CT      | 3000   | 3500 | 4000 |

As respostas a serem otimizadas são o coeficiente de desempenho do ciclo (*COP*) e a potência consumida pelo compressor (*W<sub>c</sub>*), representadas pelas suas respectivas médias e variâncias. O compressor é o componente mecanicamente mais complexo de um sistema de refrigeração por compressão de vapor [Salvador 1999]. O compressor possui essencialmente as funções de aumentar a pressão do fluido refrigerante e promover a

circulação desse fluido no sistema. Assim, o compressor comprime o fluido refrigerante de um estado de vapor superaquecido em baixa pressão e temperatura para um estado de alta pressão e temperatura. A quantidade de energia por unidade de tempo fornecida para o fluido refrigerante no compressor para elevar sua temperatura é denominada de potência de compressão, expressa pela Eq. 6.

$$\dot{W}_c = \dot{m}f (h_2 - h_1) \quad (6)$$

Sendo:

$\dot{W}_c$  – Potência de compressão [kW]

$\dot{m}f$  – Fluxo mássico do fluido refrigerante [kg/s]

$h_1, h_2$  – Entalpia do fluido refrigerante nos pontos 1 (entrada do compressor) e 2 (saída do compressor) respectivamente [kJ/kg]

A eficiência de ciclo é normalmente definida como a relação entre a energia útil que é o objetivo do ciclo, e a energia consumida que deve ser paga para a obtenção do efeito desejado [Pinelli 2008]. Embora o COP do ciclo real seja sempre menor que o do ciclo teórico para as mesmas condições de operação, pode-se, com o ciclo teórico, verificar que parâmetros influenciam no desempenho do sistema [Venturini e Pirani 2005]. O COP é definido pela Eq. 7.

$$COP = \frac{\text{Energia útil}}{\text{Energia gasta}} = \frac{\dot{Q}_0}{\dot{W}_c} \quad (7)$$

Sendo:

$\dot{Q}_0$  – Capacidade frigorífica [kW]

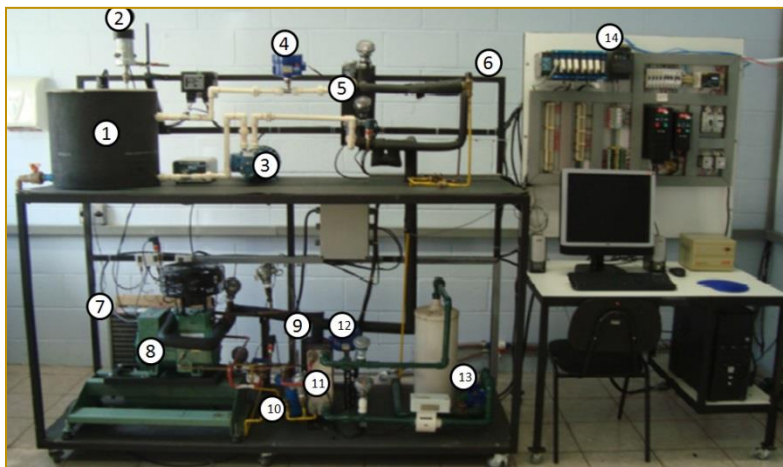
A planta de refrigeração utilizada neste trabalho pertence ao Laboratório de Controle e Automação da Faculdade de Engenharia Química (FEQ) da Unicamp. A planta é constituída de um ciclo primário de refrigeração com o R404 como fluido refrigerante e dois ciclos secundários, um no condensador e outro no evaporador, utilizando uma mistura de propilenoglicol e água (mesma proporção de volume) para a troca de calor com o fluido refrigerante.

Os processos de compressão, condensação, expansão e evaporação do sistema para

realização dos experimentos foram realizados por um compressor alternativo, um trocador de calor tipo placas, uma válvula de expansão termostática e um trocador de calor tipo placas, respectivamente. A simulação da carga térmica foi feita por meio de uma resistência elétrica, considerada nesse trabalho como uma variável não controlável (ruído) do processo.

A Figura 1 apresenta os principais componentes da planta utilizada e uma breve descrição dos mesmos.

Figura 1 - Planta experimental e os principais componentes



**1 Tanque de aquecimento de propilenoglicol:** onde é simulada a carga térmica do sistema. Uma resistência elétrica é responsável pelo aquecimento do propilenoglicol, que é então resfriado no evaporador.

**2 Agitador do tanque de propilenoglicol:** tem como objetivo deixar o fluido aquecido no tanque o mais homogêneo possível.

**3 Bomba de fluido do tanque:** responsável pelo fluxo do propilenoglicol que será resfriado no condensador e aquecido no tanque.

**4 Medidor de vazão tipo turbina:** responsável pela medição da vazão volumétrica do fluido que é resfriado no evaporador

**5 Evaporador:** onde o fluido refrigerante troca calor com o fluido do tanque.

**6 Válvula de expansão termostática:** onde ocorre a queda de pressão (expansão) do fluido refrigerante.

**7 Condensador a ar:** onde o fluido refrigerante é resfriado, trocando calor com o ar. Este componente não foi utilizado nos experimentos deste trabalho.

**8 Compressor:** responsável pelo aumento de pressão do sistema. É do tipo semi-hermético, um estágio, um êmbolo e de simples efeito.

**9 Separador de óleo:** durante a operação do compressor, parte do óleo lubrificante pode ir junto com o fluido refrigerante. O separador é responsável por retirar o óleo do fluido.

**10 Separador de líquido:** separador que retira qualquer partícula líquida do vapor

proveniente do evaporador antes de entrar no compressor.

**11 Condensador de placas:** onde o fluido refrigerante é resfriado. É onde ocorre a troca calor com a água que vem diretamente da torre de resfriamento.

**12 Medidor de vazão tipo turbina:** mede a vazão da água proveniente da torre de resfriamento.

**13 Bomba de condensação:** bomba responsável por deslocar a água proveniente da torre de resfriamento.

**14 Sistema de aquisição e controle:** todo aparato responsável pela coleta de dados dos medidores da planta (pressão, temperatura, vazão e frequência dos inversores) e pelo controle da planta.

A planta de refrigeração conta ainda com sensores de temperatura e pressão permitindo uma melhor avaliação e monitoramento do comportamento do sistema.

Depois da execução dos experimentos configurando o equipamento de acordo com os parâmetros estabelecidos pelo arranjo experimental, os dados foram tratados buscando obter apenas informações no regime permanente. Após o tratamento dos dados, foi possível obter a potência de compressão ( $W_c$ ) e o coeficiente de desempenho ( $COP$ ) pelo programa *Coolpack*®. A matriz experimental (Tabela 3) foi, então, construída e utilizada como fonte de dados para a modelagem e otimização do

processo. O planejamento de experimentos foi feito pelo software *Minitab*® através dos parâmetros fornecidos para cada nível de cada variável de controle e de ruído. Assim, as cinco primeiras colunas da matriz foram geradas pelo próprio software. A cada

experimento o sistema era desligado e religado de acordo com os parâmetros estabelecidos pelo planejamento experimental. As respostas foram, então, acrescentadas a matriz.

Tabela 3 – Matriz experimental

| Testes | Parâmetros        |                    |                    | Ruído            | Respostas                       |                 |
|--------|-------------------|--------------------|--------------------|------------------|---------------------------------|-----------------|
|        | <i>FC</i><br>(Hz) | <i>FBE</i><br>(Hz) | <i>FBC</i><br>(Hz) | <i>CT</i><br>(W) | <i>W<sub>comp</sub></i><br>(kW) | <i>COP</i><br>- |
| 1      | 48                | 40                 | 40                 | 3000             | 1,409                           | 2,38            |
| 2      | 58                | 40                 | 40                 | 3000             | 1,520                           | 2,19            |
| 3      | 48                | 54                 | 40                 | 3000             | 1,520                           | 2,19            |
| 4      | 58                | 54                 | 40                 | 3000             | 1,541                           | 2,15            |
| 5      | 48                | 40                 | 54                 | 3000             | 1,414                           | 2,34            |
| 6      | 58                | 40                 | 54                 | 3000             | 1,548                           | 2,13            |
| 7      | 48                | 54                 | 54                 | 3000             | 1,340                           | 2,44            |
| 8      | 58                | 54                 | 54                 | 3000             | 1,543                           | 2,16            |
| 9      | 48                | 40                 | 40                 | 4000             | 1,352                           | 2,58            |
| 10     | 58                | 40                 | 40                 | 4000             | 1,579                           | 2,13            |
| 11     | 48                | 54                 | 40                 | 4000             | 1,428                           | 2,37            |
| 12     | 58                | 54                 | 40                 | 4000             | 1,563                           | 2,17            |
| 13     | 48                | 40                 | 54                 | 4000             | 1,416                           | 2,38            |
| 14     | 58                | 40                 | 54                 | 4000             | 1,558                           | 2,16            |
| 15     | 48                | 54                 | 54                 | 4000             | 1,430                           | 2,39            |
| 16     | 58                | 54                 | 54                 | 4000             | 1,549                           | 2,19            |
| 17     | 43                | 47                 | 47                 | 3500             | 1,328                           | 2,55            |
| 18     | 63                | 47                 | 47                 | 3500             | 1,578                           | 2,12            |
| 19     | 53                | 33                 | 47                 | 3500             | 1,456                           | 2,31            |
| 20     | 53                | 61                 | 47                 | 3500             | 1,513                           | 2,22            |
| 21     | 53                | 47                 | 33                 | 3500             | 1,508                           | 2,23            |
| 22     | 53                | 47                 | 61                 | 3500             | 1,503                           | 2,24            |
| 23     | 53                | 47                 | 47                 | 3500             | 1,495                           | 2,25            |
| 24     | 53                | 47                 | 47                 | 3500             | 1,495                           | 2,25            |
| 25     | 53                | 47                 | 47                 | 3500             | 1,483                           | 2,25            |
| 26     | 53                | 47                 | 47                 | 3500             | 1,483                           | 2,25            |
| 27     | 53                | 47                 | 47                 | 3500             | 1,480                           | 2,26            |
| 28     | 53                | 47                 | 47                 | 3500             | 1,481                           | 2,25            |
| 29     | 53                | 47                 | 47                 | 3500             | 1,481                           | 2,25            |

#### 4. RESULTADOS

O modelo genérico de superfície de resposta para as três variáveis de controle (*FC*, *FBE*,

*FBC*) e a variável de ruído (*CT*), correspondente às respostas *COP* e *W<sub>c</sub>*, pode ser expresso conforme indica a Eq. (8):

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \beta_1 FC + \beta_2 FBE + \beta_3 FBC + \beta_{11} FC^2 + \beta_{22} FBE^2 + \beta_{33} FBC^2 + \beta_{12} FC \cdot FBE + \beta_{13} FC \cdot FBC + \beta_{23} FBE \cdot FBC + \gamma_1 CT + \delta_{11} FC \cdot CT + \delta_{21} FBE \cdot CT + \delta_{31} FBC \cdot CT \quad (8)$$

Os modelos de *COP* e *W<sub>c</sub>* são expressos com as variáveis de controle e de ruído codificadas. Os coeficientes dos modelos foram estimados pelo Método dos Mínimos Quadrados Ordinários (OLS) através do software estatístico *Minitab*®. Os ajustes (*R*<sup>2</sup> *adj.*) encontrados inicialmente para os

modelos de *W<sub>c</sub>* e *COP* foram de 82,47% e 84,95% respectivamente. Dessa forma, visando aumentar a capacidade de representação dos dados, o modelo foi reduzido retirando os termos menos significativos. Após esta redução, foram obtidos os valores apresentados pelas Eqs.



(9) e (10), sendo os novos ajustes iguais a 84,63% e 87,74%, respectivamente.

$$Wc(x, z) = 1,4898 + 0,0663 FC + 0,0097 FBE - 0,0052 FBC + 0,0025 CT - 0,0088 FC * FC - 0,0085 FC * FBE + 0,0096 FC * CT - 0,0166 FBE * FBC + 0,0110 FBC * CT \quad (9)$$

$$COP(x, z) = 2,2509 - 0,1104 FC - 0,0171 FBE + 0,0021 FBC + 0,0244 CT + 0,0210 FC * FC + 0,0219 FC * FBE - 0,0219 FC * CT + 0,0356 FBE * FBC - 0,0181 FBC * CT \quad (10)$$

Os modelos de média e variância das respostas  $Wc$  e  $COP$ , a seguir, foram extraídos das Eqs. (9) e (10), respectivamente, de acordo com as Eqs. (2) e (3):

$$\mu(Wc) = 1,4898 + 0,0663 FC + 0,0097 FBE - 0,0052 FBC - 0,0088 FC * FC - 0,0085 FC * FBE - 0,0166 FBE * FBC \quad (11)$$

$$\sigma^2(Wc) = 0,0000062 - 0,000048 FC + 0,000055 FBC - 0,000093 FC * FC + 0,000121 FBC * FBC + 0,000212 FC * FBC \quad (12)$$

$$\mu(COP) = 2,2509 - 0,1104 FC - 0,0171 FBE + 0,0021 FBC + 0,0210 FC * FC + 0,0219 FC * FBE + 0,0356 FBE * FBC \quad (13)$$

$$\sigma^2(COP) = 0,000595 - 0,001069 FC - 0,000883 FBC + 0,000480 FC * FC + 0,000328 FBC * FBC + 0,000793 FC * FBC \quad (14)$$

Pode-se observar que os modelos de média e variância desenvolvidos por um arranjo combinado foram escritos como função apenas das variáveis de controle, embora a variável de ruído fosse testada em diferentes níveis durante os experimentos.

Uma vez obtidas as equações de média e variância, é possível aglutiná-las através do modelo do erro quadrático médio ( $EQM$ ). Primeiramente, para a construção dos

modelos de acordo com a Eq. (4), é necessário especificar os valores alvo das respostas. Esses valores podem ser obtidos pela otimização individual da função de média de  $Wc$  e  $COP$ . Assim, os valores alvo de  $Wc$  e  $COP$  definidos foram 1,315 kW e 2,575, respectivamente. Os modelos do erro quadrático médio de cada resposta podem ser então representados pelas Eqs. (15) e (16):

$$EQM(Wc) = [\mu(Wc) - 1,315]^2 + \sigma^2(Wc) \quad (15)$$

$$EQM(COP) = [\mu(COP) - 2,575]^2 + \sigma^2(COP) \quad (16)$$

Com as equações do  $EQM$  estabelecidas, realizou-se a otimização robusta minimizando a média e a variância das respostas simultaneamente. Para construir a formulação de otimização (Eq. 5), foi necessário definir antes os valores alvo das Eqs. (15) e (16) através da otimização individual de cada

função. Os alvos para os erros quadráticos médios foram fixados em 0,000160 para  $EQM(Wc)$  e 0,003715 para  $EQM(COP)$ . Assim, a otimização robusta multi-objetivo, formulada com base no Método do Critério Global, foi definida como:

$$\text{Minimizar } EQM_G = \left[ \frac{0,000160 - EQM(Wc)}{0,000160} \right]^2 + \left[ \frac{0,003715 - EQM(COP)}{0,003715} \right]^2 \quad (17)$$

$$\text{Sujeito a: } FC^2 + FBE^2 + FBC^2 \leq 4,0$$

Sendo:

$EQM_G$  – Erro quadrático médio global

$EQM(Wc), EQM(COP)$  – Modelos do erro quadrático médio representados pelas Eqs. (15) e (16)

Para a solução do problema de otimização, a formulação anterior foi devidamente programada em uma planilha do *Microsoft Excel*® e o suplemento *Solver* do programa foi utilizado, empregando o método GRG. Assim, o resultado ótimo obtido apresentou os seguintes parâmetros robustos dispostos na Tabela 4. Tais resultados foram obtidos com um nível de confiança de 95%.

O resultado ótimo mostra que as médias se estabeleceram próximos aos alvos, já que uma diferença de 0,024 para o coeficiente de desempenho e 0,001 kW para a potência de compressão foram verificados em relação aos valores desejados. Quanto à variabilidade no ponto de ótimo, as Figs. 2 e 3 apresentam uma comparação dos resultados com as melhores condições experimentais para o coeficiente de desempenho (experimento 9) e para a potência de compressão (experimento 17). Tais gráficos foram construídos a partir da simulação dos dados obtidos com os modelos de superfície de resposta desenvolvidos. Como pode ser observado, no experimento com melhor coeficiente de desempenho (Fig. 2) foi observada uma alta potência de compressão e uma maior variabilidade em torno desta característica.

Da mesma forma, o experimento com melhor potência de compressão (Fig. 3) apresentou um coeficiente de desempenho próximo do ótimo, porém, com uma maior variabilidade. O resultado ótimo, por sua vez, foi capaz de combinar as vantagens dos dois experimentos em uma única condição, ou seja, apresentar ao mesmo tempo um coeficiente de desempenho alto e uma potência de compressão baixa, com baixa variabilidade para ambas as respostas.

Em comparação com o melhor resultado individual, o coeficiente de desempenho no ponto de ótimo, foi levemente inferior ao valor obtido no experimento 9, porém com menor variabilidade. Em relação a potência de compressão, o resultado ótimo foi inferior e apresentou uma variabilidade significativamente menor em relação ao experimento 17. Levando em consideração à existência de conflitos de interesse característicos de problemas de otimização multi-objetiva e tendo em vista os resultados observados nas Figs. 2 e 3, julgou-se como satisfatória a otimização robusta desenvolvida neste trabalho para um sistema de refrigeração por compressão de vapor.

Tabela 4 - Parâmetros robustos ótimos para um sistema de refrigeração por compressão de vapor

|                 | Variáveis de controle |       |       | Respostas |                 |            |                 |
|-----------------|-----------------------|-------|-------|-----------|-----------------|------------|-----------------|
|                 | $FC$                  | $FBE$ | $FBC$ | $\mu(Wc)$ | $\sigma^2(Wc)$  | $\mu(COP)$ | $\sigma^2(COP)$ |
| Resultado ótimo | 44                    | 44    | 51    | 1,316     | 0,000099        | 2,551      | 0,003165        |
| Unidade         | Hz                    | Hz    | Hz    | kW        | kW <sup>2</sup> | -          | -               |

Figura 2 – Comparação do resultado ótimo com o experimento de melhor coeficiente de desempenho

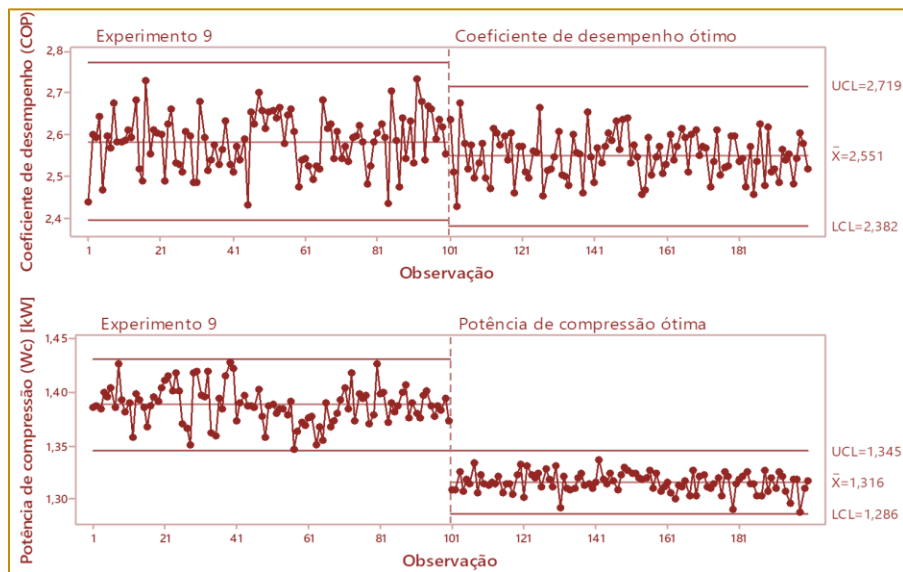
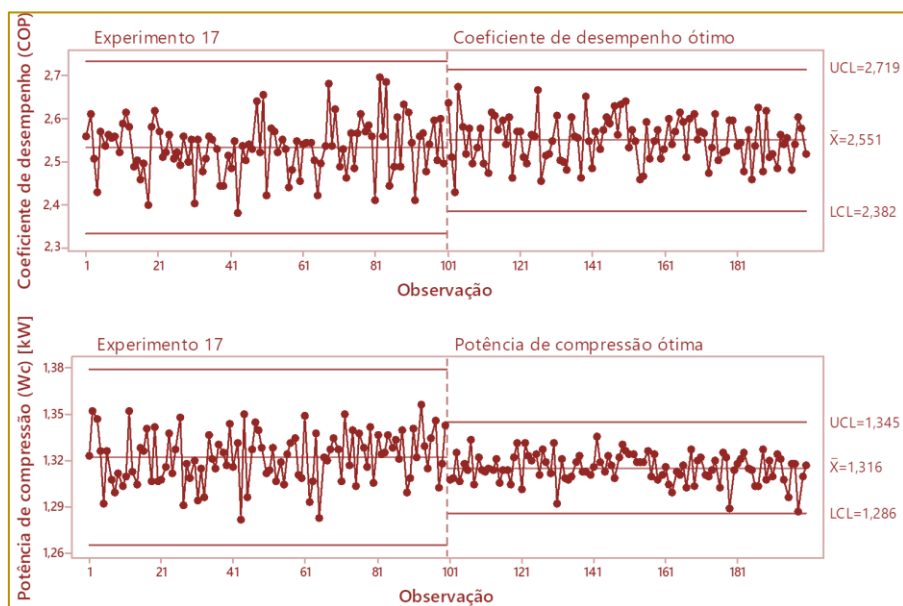


Figura 3 – Comparação do resultado ótimo com o experimento de melhor potência de compressão



## 5. CONCLUSÃO

O presente trabalho apresentou os melhores parâmetros obtidos para uma otimização robusta da potência consumida pelo compressor e o coeficiente de desempenho do ciclo de um sistema de refrigeração por compressão de vapor. A modelagem das respostas levou em consideração um ruído do processo (carga térmica), variável que não se tem controle, mas que pode influenciar o resultado final. Assim, a modelagem foi feita pelo arranjo combinado das variáveis de controle (frequência do motor do compressor e as frequências dos motores das bombas do

evaporador e condensador) e a de ruído. Os modelos desenvolvidos foram caracterizados como expressões confiáveis para a representação das respostas, já que os ajustes foram superiores a 80% (84,63% para  $Wc$  e 87,74% para  $COP$ ). Os modelos adquiridos foram então remodelados pela equação do erro quadrático médio que aglutinou em uma única equação a média e a variância de cada resposta.

Para realizar a otimização robusta multi-objetiva, optou-se por utilizar o Método do Critério Global e o método GRG. A solução ótima obtida para o sistema de refrigeração

apresentou os seguintes parâmetros para as variáveis de controle:  $FC = 44$  Hz,  $FBE = 44$  Hz,  $FBC = 51$  Hz. Com este ajuste, obtêm-se valores médios para  $Wc$  e  $COP$  iguais a 1,316 kW e 2,551, e variâncias de 0,000099 kW<sup>2</sup> (desvio-padrão 0,010 kW) e 0,003165 (desvio-padrão 0,056), respectivamente. Tais resultados se mostraram devidamente próximos aos valores ótimos individuais, em uma condição simultânea de mínima variabilidade para ambas as respostas. Em relação à variabilidade, a simulação e comparação dos dados no ponto de ótimo com os melhores resultados obtidos

experimentalmente mostraram que a solução ótima foi capaz de atingir um alto coeficiente de desempenho, uma baixa potência de compressão com baixa variabilidade para ambas as respostas, o que não ocorreu quando elas foram verificadas de forma isolada.

## 6. AGRADECIMENTOS

Os autores agradecem a FAPESP, FAPEMIG, CAPES e CNPq pelo apoio à realização deste trabalho.

## REFERÊNCIAS

- [1] Brito, P., Lopes, P., Reis, P., Alves, O. (2014). Simulation and optimization of energy consumption in cold storage chambers from the horticultural industry. *International Journal of Energy and Environmental Engineering*, v.88, p. 1-15.
- [2] Dincer, I. (2003). *Refrigeration Systems and Applications*. New Jersey: John Wiley & Sons.
- [3] Dincer, I., Kanoglu, M. (2010). *Refrigeration Systems and Applications*. New Jersey: John Wiley & Sons.
- [4] Köksoy, O. (2006). Multiresponse robust design: Mean square error (MSE) criterion. *Applied Mathematics and Computation*, v. 175, p. 1716-1729.
- [5] Manske, K. A. (1999). *Performance Optimization of Industrial Refrigeration Systems*. Master Science, Thesis, Mechanical Engineering, University of Wisconsin- Madison.
- [6] Mendes, T. (2012). Diagnóstico Termodinâmico Aplicado a um Sistema de Refrigeração por Compressão de Vapor. Dissertação (Mestrado em Conversão de Energia) - Instituto de Engenharia Mecânica, Universidade Federal de Itajubá, UNIFEI, Itajubá.
- [7] Montgomery, D.C. (2005). *Design and Analysis of Experiments*. Ed. John Wiley, New York.
- [8] Qureshi, B. A., Zubair, S. M. (2012). The effect of refrigerant combinations on performance of a vapor compression refrigeration system with dedicated mechanical sub-cooling. *International Journal of Refrigeration*, v.35, p. 47-57.
- [9] Rao, S.S. (1996). *Engineering optimization: theory and practice*. Ed. John Wiley & Sons, New Jersey.
- [10] Salvador, F. (1999). Projeto de um sistema de refrigeração industrial com 'set-point' variável. Dissertação (Mestrado em Engenharia de Sistemas) - Escola Politécnica da Universidade de São Paulo, USP, São Paulo.
- [11] Stanfield, C., Skaves, D. (2009). *Fundamentals of HVAC/R*. New Jersey: Prentice Hall.
- [12] Venturini, O. J., Pirani, M. J. (2005). Eficiência Energética em Sistemas de Refrigeração Industrial e Comercial. Rio de Janeiro: Ed. Eletrobrás.
- [13] Vián, J. G., Astrain, D. (2009). Development of a hybrid refrigerator combining thermoelectric and vapor compression technologies. *Applied Thermal Engineering*, v.29, p. 3319-3327.
- [14] Widell, K.N., Eikevik, T. (2010). Reducing power consumption in multi-compressor refrigeration systems. *International Journal of Refrigeration*, v.33, p. 88-94.
- [15] Zhao, L., Cai, W., Ding, X., Chang, W. (2013a). Decentralized optimization for vapor compression refrigeration cycle. *Applied Thermal Engineering*, v.51, p. 753-763.
- [16] Zhao, L., Cai, W., Ding, X., Chang, W. (2013b). Model-based optimization for vapor compression refrigeration cycle. *Energy*, v.55, p. 392-402.

# Capítulo 3

## UM ESTUDO SOBRE A UTILIZAÇÃO DE MODELOS MULTI-ELOS PARA GESTÃO DE PEÇAS DE REPOSIÇÃO

*João Paulo de Moraes Paludo*

*Takashi Yoneyama*

*Leonardo Ramos Rodrigues*

**Resumo:** Na indústria aeronáutica, o investimento em peças de reposição possui valores altos em função dos elevados custos das peças e dos altos níveis de disponibilidade exigidos. Portanto, a implementação de uma estratégia eficiente de controle de estoques torna-se crucial. O uso de modelos multi-elos pode proporcionar uma redução considerada nos custos de estoques para uma determinada disponibilidade. Este artigo apresenta um estudo sobre o uso de um modelo multi-elos para gestão de estoques de peças de reposição. Simulações foram realizadas e os resultados obtidos com o modelo multi-elos foram comparados com os resultados obtidos com um modelo *Single Site*. Os resultados mostram que o uso de modelo multi-elos pode proporcionar reduções significativas de custos.

## 1. INTRODUÇÃO

Considerando a operação das companhias aéreas, em que é crucial manter uma elevada taxa de utilização das aeronaves, o modo de gerenciar as peças de reposição pode afetar de forma significativa os diversos custos envolvidos, bem como a disponibilidade da frota [Kilpi et al., 2009]. Um nível elevado de peças de reposição pode trazer uma alta disponibilidade. Contudo, implicará em grandes custos em função do alto capital paralisado e das despesas gerais de armazenamento. Por outro lado, um estoque subdimensionado pode levar a uma baixa disponibilidade da frota, o que também incorre em custos [Kennedy et al., 2002].

Nesse contexto, a utilização de estoques compartilhados de peças de reposição tem se mostrado uma forma de aumentar a disponibilidade das aeronaves, sem incorrer em custos muito elevados de estoques [Wong et al., 2005; Kutanoglu, 2008].

O objetivo deste trabalho é apresentar um estudo sobre o uso de um modelo multi-elos para gestão de estoques de peças de reposição. Simulações são realizadas e os resultados obtidos com o modelo multi-elos são comparados com os resultados obtidos com um modelo *Single Site*.

As demais seções deste trabalho estão organizadas como segue. A seção 2 descreve o problema de compartilhamento de peças de reposição. A seção 3 apresenta uma breve revisão da literatura. Na seção 4 são apresentados os modelos usados neste

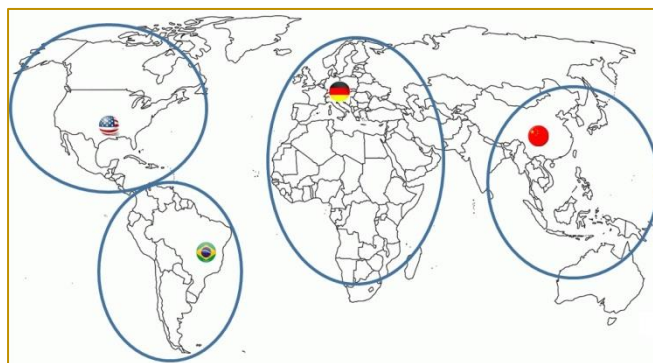
trabalho. A seção 5 apresenta os cenários e os dados usados nas simulações. Os resultados das simulações são apresentados na seção 6. As conclusões do trabalho são apresentadas na seção 7.

## 2. O PROBLEMA DE COMPARTILHAMENTO DE PEÇAS DE REPOSIÇÃO

Um importante fator de desempenho para operadores de aeronaves é a alta utilização das mesmas. Mesmo com a utilização de componentes modulares e de alta confiabilidade, um serviço de substituição rápida é fundamental em caso de falhas. Manter unidades funcionais de substituição em mãos diminui o tempo de reparo das peças falhadas. Os custos deste serviço dependem da disponibilidade desejada e são causados, principalmente, pelo capital investido na reposição de componentes [Kilpi e Vepsalainen, 2004].

A Figura 1 mostra um exemplo de como podem ser criadas redes de compartilhamento de peças. Nela pode-se observar um sistema de compartilhamento por região, no qual alguns países possuem um estoque que é responsável por atender as demandas de uma determinada região. O dimensionamento de cada estoque é feito em função da disponibilidade desejada para a região correspondente. No exemplo, todos os operadores de aeronaves localizados na América do Sul seriam atendidos por um estoque de peças localizado no Brasil.

Figura 1: Exemplo de um sistema de compartilhamento de peças



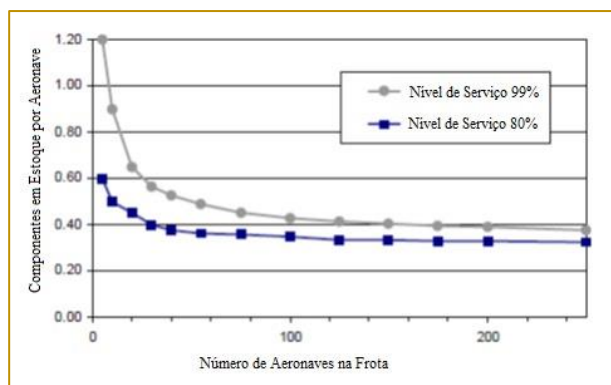
O principal objetivo da utilização desta abordagem de compartilhamento de peças de reposição é a redução dos custos com estoques. Se considerarmos um sistema em que cada operador seja responsável por gerir

seus próprios estoques, cada um teria de escalonar seus estoques conforme a tamanho de sua frota e a disponibilidade requerida. Como pode ser visto na Figura 2, um operador com poucas aeronaves necessitaria

de um maior número de peças por aeronave para manter a mesma disponibilidade. A partir do momento em que diversos operadores passam a compartilhar suas peças de reposição, o número de aeronaves atendidas por este estoque aumenta e, conforme

demonstra a Figura 2, tem-se um número menor de peças por aeronave, ou seja, um menor custo com estoques por aeronave para um mesmo nível de serviço [Kilpi e Vepsalainen, 2004].

Figura 2: Economia de escala nos custos de estoque



A ideia da utilização do compartilhamento de estoques entre operadores não é uma prática nova no meio aeronáutico. De acordo com Lombardo [2000], nos anos 60 vários operadores com aeronaves similares fundaram consórcios de manutenção com o objetivo de reduzir custos. Alguns dos consórcios formados foram a KSSU e a Atlas. A KSSU era formada pela KLM, SAS, Swissair e UTA. A Atlas era formada pela Air France, Alitalia, Ibeira e Lufthansa.

### 3. REFERENCIAL TEÓRICO

Nos tempos atuais, é de conhecimento geral que a maior parte das organizações possui um local destinado a estocar as peças de reposição, além de um sistema de gerenciamento dessas peças. Desse modo, a forma com que o gerenciamento do estoque é realizado pode se caracterizar como um dos potenciais diferenciais competitivos de uma organização.

No setor aeronáutico, o gerenciamento de estoques se torna ainda mais crítico pois, em geral, as peças de reposição são caras, aumentando os custos com estoques. De acordo com Kinnison [2004], a gestão de estoques é a atividade onde as companhias aéreas gastam a maior parte de seus recursos, sendo esta área objeto de grande interesse da alta administração.

Sabendo da importância do gerenciamento de estoques, diversos modelos foram criados

para auxiliar no controle e dar maior previsibilidade na gestão dos mesmos.

#### 3.1. SISTEMAS MULTI-ELOS

Sistemas multi-elos podem ser estruturados de diversas formas, e a determinação da topologia de cada sistema é determinante para otimização e análise do mesmo. A Figura 3 apresenta as diferentes topologias de sistemas multi-elos [Snyder e Shen, 2011]. Cada estágio de um sistema multi-elo pode representar uma localidade, um item ou até mesmo uma atividade.

O sistema serial é a topologia mais simplificada dos sistemas multi-elos pois cada estágio possui no máximo um predecessor e um sucessor.

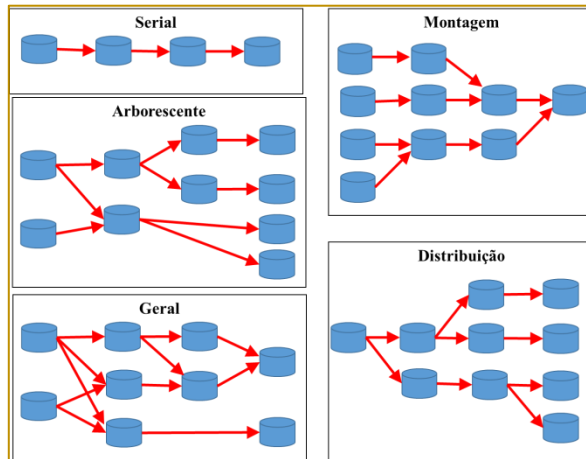
No sistema de montagem cada estágio contém no máximo um sucessor, enquanto no sistema de distribuição cada estágio tem no máximo um predecessor. Sistemas de montagem são geralmente utilizados para representar como um produto final é concebido a partir de uma lista de materiais intermediários. Sistemas de distribuição buscam representar como é feito fluxo de abastecimento de estoque intermediários até a entrega do produto ao consumidor final.

Sistemas arborescentes são sistemas híbridos de montagem e distribuição. Os mesmos não possuem limitações quanto ao número de predecessores e sucessores. Contudo, o grafo de um sistema arborescente não

permite a criação de ciclos. Na teoria dos grafos, um ciclo é uma parte do grafo cujas ligações permitem que a partir de um nó de partida seja feita uma movimentação através de uma sequência de outros nós, voltando ao nó de partida sem repetir quaisquer as ligações de nós já utilizadas.

Por fim existem os sistemas gerais onde não existem restrições quanto ao número de sucessores e predecessores, sendo permitida a existência de ciclos no grafo, permitindo maior flexibilidade ao mesmo tempo que aumenta a complexidade da análise.

Figura 3: Topologias de sistemas multi-elos



A literatura relacionada à gestão de estoques em sistemas multi-elos teve início nos anos 60. Um método analítico para solução de problemas de gerenciamento de estoques em um sistema multi-elos serial foi apresentado por Clark e Scarf [1960]. Esse modelo considera uma demanda estocástica e não estacionária, os tempos de ressuprimento são determinísticos e o custo é composto pelos custos de armazenagem, pedido e entrega.

Outro modelo foi desenvolvido por Ehrhardt et al. [1981], onde foi proposta uma política estacionária de revisão contínua do tipo  $(s, S)$ . Nessa política, assim que o estoque atingir um nível igual ou inferior a  $s$ , um novo pedido é feito para elevar o nível do estoque para  $S$ . Nesse trabalho, os autores consideram apenas o estoque principal, cujas demandas vêm de estoques avançados que utilizam a mesma política de revisão contínua com demandas independentes e identicamente distribuídas. Como custos atribuídos a este modelo, foram considerados os custos relativos a armazenagem, entrega e pedido.

Um modelo de revisão contínua com política do tipo  $(s, Q)$  foi proposto por Ganeshan

[1999]. Nessa política, um pedido de compra é disparado sempre que o nível do estoque for igual ou inferior a  $s$ , e o tamanho do pedido é fixo e igual a  $Q$ . O modelo proposto é composto por um estoque principal e controla apenas um tipo de item que é fornecido para diversos estoques avançados. As demandas provenientes dos estoques avançados comportam-se segundo uma distribuição de Poisson e os tempos de ressuprimento são estocásticos.

### 3.2. NÍVEL DE ESTOQUE VIRTUAL

O nível de estoque é um dos principais parâmetros a serem otimizados na teoria de estoques. Sendo assim, pode-se definir dois tipos de nível de estoque, o nível de estoque em mãos e o nível de estoque virtual. O nível de estoque em mãos é a quantidade de peças que estão fisicamente armazenadas no estoque. O nível de estoque virtual mostra qual a real quantidade de peças no sistema, incluindo pedidos que ainda não foram recebidos e pedidos em aberto. O cálculo do estoque virtual é feito pela Equação (1).

$$S = OH + DI - BO \quad (1)$$



onde  $S$  é o nível de estoque virtual,  $OH$  é o número de peças em mãos,  $DI$  é o número de pedidos enviados que ainda não foram recebidos e  $BO$  é o número de pedidos em aberto.

O estoque virtual é muito útil pois mostra quantas peças prontas para uso estão disponíveis. Se o estoque virtual é superior ao estoque em mãos, isso significa que existem pedidos a receber. Se o estoque virtual é menor do que no estoque em mãos, isso significa que existem pedidos em aberto.

### 3.3. ESTIMATIVA DE PEDIDOS EM ABERTO

Uma etapa importante na modelagem de um sistema de gerenciamento de estoques é prever a demanda de cada um dos itens pelo qual o estoque é composto. Desta maneira, pode-se estimar o número de peças necessário para atingir o patamar de disponibilidade ou custo desejado, considerando a diversificação de cada estoque.

No presente trabalho, adotou-se a premissa de que as demandas seguem uma distribuição de Poisson. Segundo o teorema de Palm, se a demanda por um item é um processo de Poisson com média anual  $m$  e os tempos de reparo das unidades falhadas são independentes e identicamente distribuídos de acordo com qualquer distribuição com média  $T$  anos, então a distribuição de probabilidade do estado estacionário para o número de unidades de reparo tem uma

distribuição de Poisson com média  $mT$  [Sherbrooke, 2004].

O valor de  $mT$  corresponde ao valor médio de peças em reparo. Portanto, o número estimado de pedidos em abertos ou EBO (*expected backorders*) pode ser calculado a partir de parâmetros de demanda de cada item em cada local de operação da aeronave, dos tempos de reparo nas oficinas, da probabilidade do item ser reparado no próprio local de operação e do tempo de remessa entre o armazém central e o armazém local (*lead-time*) [Sherbrooke, 2004].

### 3.4. ANÁLISE MARGINAL

Uma vez que a adição de uma nova peça de reposição ao estoque resulta em uma redução do número de pedidos em aberto, em sistemas de múltiplos componentes é necessário utilizar alguma técnica para decidir qual dos componentes deve ser comprado para se atingir a maior redução de *expected backorders* (EBO) com relação ao custo desta compra.

A Análise Marginal [Sherbrooke, 2004] recebe este nome porque em cada passo do algoritmo necessita-se avaliar apenas o impacto no custo e na disponibilidade causado pelo próximo item a ser adquirido. Ou seja, o valor marginal ou incremental fornece todas as informações necessárias sobre cada item. O valor calculado pela Análise Marginal,  $\Delta EBO$ , é obtido com base na Equação (2).

$$\Delta EBO = (EBO(s-1) - EBO(s)) / C \quad (2)$$

onde  $EBO(s-1)$  é o número esperado de pedidos em aberto com  $s-1$  peças,  $EBO(s)$  é o número esperado de pedidos em aberto com  $s$  peças e  $C$  é o custo unitário da peça.

Considere, por exemplo, o item do tipo "A", de custo 30, e o item do tipo "B", de custo 65. O número de pedidos em abertos é fornecido

pela Tabela 1. Se for utilizada a técnica de análise marginal, pode-se decidir se, a partir de um estoque vazio, ( $s=0$ ), deve-se comprar uma peça do item "A" ou do item "B" para que com uma peça em estoque tenha-se a maior redução de pedidos em aberto por custo de compra.

Tabela 1: Exemplo de aplicação da Análise Marginal

| s | Item A |              | Item B |              |
|---|--------|--------------|--------|--------------|
|   | EBO(s) | $\Delta EBO$ | EBO(s) | $\Delta EBO$ |
| 0 | 3,780  | -            | 6,220  | -            |
| 1 | 2,800  | 0,033        | 5,220  | 0,015        |
| 2 | 1,910  | 0,030        | 4,230  | 0,015        |
| 3 | 1,180  | 0,024        | 3,290  | 0,015        |

Tendo como base as colunas de EBO de cada item, observa-se que com um estoque zero ter-se-ia  $EBO(0) = 3,78 + 6,22 = 10$  pedidos em aberto. As colunas de  $\Delta EBO$  permitem comparar o quanto os pedidos em abertos iriam ser reduzidos por cada unidade monetária gasta. Desta maneira, o item "A" é o escolhido, uma vez que seu  $\Delta EBO = (3,78 - 2,80)/30 = 0,033$  e o item "B"  $\Delta EBO = (6,22 - 5,22)/65 = 0,015$ . Portanto, com um estoque de uma peça, chega-se ao número de pedidos em aberto de  $2,80 + 6,22 = 9,02$ . A análise marginal é usada nos dois modelos estudados neste trabalho: o modelo *Single Site* e o modelo *METRIC*, que serão descritos na próxima seção.

#### 4. MODELOS PROPOSTOS

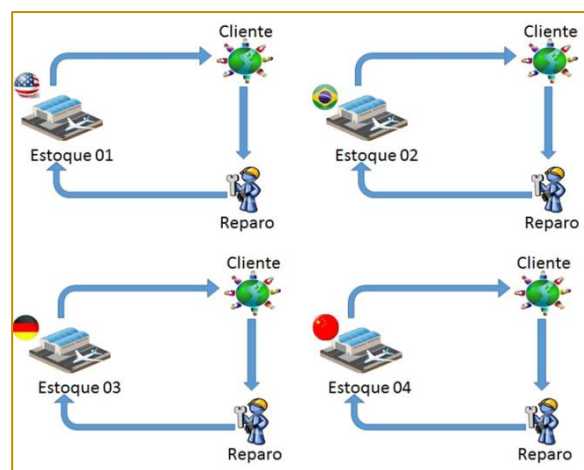
##### 4.1. MODELO *SINGLE SITE*

O modelo *Single Site* tem como objetivo modelar a dinâmica presente no sistema de gerenciamento de peças com apenas um estoque e diversos tipos de peças a serem administrados. O cenário modelado para o caso *Single Site* buscou modelar diversas

bases em diferentes localizações, sendo que em cada base é possível estocar diferentes itens. A Figura 4 mostra os ciclos percorridos pelas peças. Uma vez que a falha ocorre em alguma das aeronaves da frota do cliente, a mesma é retirada da aeronave e enviada para a oficina de reparo. Considerando que o estoque é capaz de atender a demanda deste tipo de peça, outra peça em condições operacionais é enviada para o cliente. Na oficina de reparo a peça falhada é reparada. Após o reparo, a peça é enviada para o estoque do qual o cliente em questão demandou a peça para substituição. Desta maneira, o ciclo é fechado e a peça reparada volta a ficar disponível para consumo dos clientes.

Conforme apresentado neste modelo, as peças que falham no ciclo de clientes da base 1 serão consertadas e repostas novamente na base 1. Neste cenário, não é possível mandar uma peça para reparo em uma região diferente da região de origem da peça.

Figura 4: Cenário de simulação para o modelo *Single Site*



Para um sistema de estoques de múltiplos tipos de componentes, o cálculo da

disponibilidade da frota no modelo *Single Site* é feito usando a Equação (3).

$$A = 100 \prod_{i=1}^I \left\{ (1 - EBO_i(s_i) / NZ_i) \right\}^{Z_i} \quad (3)$$

onde  $A$  é a disponibilidade da frota,  $I$  é o número de tipos de componentes,  $s_i$  é a quantidade de peças do componente  $i$ ,

$EBO(s_i)$  é o número estimado de pedidos em aberto do componente  $i$  se  $s_i$  peças forem compradas,  $N$  é o tamanho da frota e  $Z_i$  é o

número de peças do componente  $i$  instalados em cada aeronave.

#### 4.2. MODELO METRIC

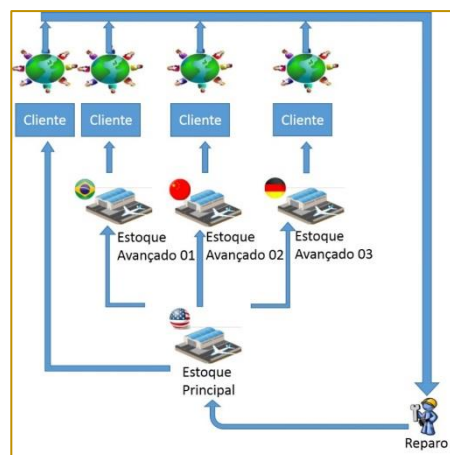
O modelo *METRIC* (*Multi-Echelon Technique for Recoverable Item Control*) foi originalmente proposto por Sherbrooke [2004] e tem como objetivo representar a dinâmica do gerenciamento de estoques em uma rede de compartilhamento de peças em que as diversas bases compartilham peças entre si e também com um estoque principal, denominado *Hub*. Para a modelagem *METRIC*, também é considerado o cenário com diversos tipos de peças.

Em sua concepção original, o modelo *METRIC* envolve o emprego de um estoque principal (*Hub*) de onde os estoques

avançados (bases) demandam itens. O principal objetivo da introdução de um *Hub* onde não existem demandas diretas de clientes é poder concentrar as peças em um só local e eliminar a necessidade de duplicar estoques de segurança.

O cenário modelado neste estudo trata-se de uma adaptação do modelo *METRIC*. O modelo trabalhado busca analisar um sistema de gerenciamento de estoques conforme apresentado na Figura 5, em que as demandas geradas pelos clientes são atendidas pelo estoque da base correspondente. As bases, por sua vez, são atendidas pelo *Hub*. As peças são sempre enviadas para um único reparador e, uma vez reparadas, são enviadas para o *Hub*.

Figura 5: Cenário de estudo do modelo *METRIC*



Para o modelo *METRIC*, considerando que uma parcela das peças falhadas será reparada na própria base e as demais peças serão reparadas no *Hub*, deve-se atentar para o cálculo das demandas de cada base e

principalmente do *Hub*. O cálculo da demanda média anual de peças no *Hub* para o modelo *METRIC* é feito usando a Equação (4).

$$m_{os} = \sum_{j=1}^J m_j (1 - r_j) \quad (4)$$

onde  $m_{os}$  é a demanda média anual no *Hub* proveniente das bases,  $m_j$  é a demanda média anual na base  $j$  e  $r_j$  é a probabilidade de reparo na base  $j$ .

Considerando a adaptação utilizada no modelo proposto, a demanda média anual total do *Hub*,  $m_0$ , é obtida pela Equação (5):

$$m_0 = m_{os} + m_{Hub} \quad (5)$$

onde  $m_{Hub}$  é a demanda média anual no *Hub* proveniente dos clientes diretos.

O valor médio de peças em reparo no *Hub* é calculado de acordo com o teorema de Palm

$$\mu_j = m_j (r_j T_j + (1 - r_j) \{O_j + EBO(S_0 | m_0 T_0) / m_0\}) \quad (6)$$

onde  $T_j$  é o tempo médio de reparo na base  $j$ ,  $O_j$  é o *lead-time* entre o *Hub* e a base  $j$ ,  $S_0$  é o nível de estoque do *Hub* e  $T_0$  é o tempo médio de reparo no *Hub*.

Para o item que é sempre reparado na base, o valor médio de peças em reparo será dado por  $m_j T_j$ , assim como apresentado para o modelo *Single Site*. Para os itens que em parte das vezes são reparados no *Hub*, nota-se que a parte correspondente aos reparos feitos na base são calculados por  $m_j r_j T_j$  e o

[Sherbrooke, 2004]. O cálculo do valor médio de peças em reparo em cada base é feito usando a Equação (6).

restante da equação envolve as peças que serão reparadas no *Hub*, onde são levados em consideração o *lead-time* do reparo no *Hub* mais o número de pedidos em aberto colocados pela base avançada que ainda não foram finalizados.

Para realização do cálculo da disponibilidade do modelo *METRIC* é utilizada a Equação (7), que é uma média das disponibilidades de cada uma das bases avançadas ponderada pelo tamanho das respectivas frotas.

$$A = \frac{\sum_{j=1}^J A_j N_j}{\sum_{j=1}^J N_j} \quad (7)$$

onde  $A_j$  é a disponibilidade na base  $j$ ,  $N_j$  é o tamanho da frota atendida pela base  $j$  e  $J$  é o número total de bases.

Com os valores das disponibilidades, dos pedidos em aberto e dos custos das diversas configurações de estoque para cada nível total de estoque  $S$ , pode-se fazer a construção da curva de pedidos em aberto por custo e de disponibilidade por custo.

## 5. SIMULAÇÃO

Considerando os modelos *Single Site* e *METRIC* apresentados, serão feitas simulações nas quais o objetivo é minimizar o custo total das peças adquiridas, sujeito a uma restrição de disponibilidade. Será feito um estudo comparativo dos modelos *Single Site* e *METRIC* em termos de custo de aquisição de peças. Para o modelo *Single Site*, uma simulação é feita para cada base e

os custos são somados para compor o custo total. Para o modelo *METRIC*, é feita uma simulação para cada localização possível do *Hub*, e o menor valor é considerado.

Neste trabalho foram usados dados sintéticos. Quatro bases hipotéticas são consideradas. A quantidade de aeronaves atendidas por cada uma das bases é apresentada na Tabela 2. As características dos itens e dos estoques associados são apresentadas nas Tabelas 3 e 4. Nessas tabelas, são apresentados a demanda média anual  $m_j$ , o tempo médio de reparo  $T_j$  (em ano), a probabilidade de reparo do item na própria base  $r_j$  e o custo unitário de cada peça  $C_j$ . O *lead time* entre as bases é apresentado na Tabela 5. O *lead-time* representa o tempo de deslocamento (em ano) entre as bases e o *Hub*.

Tabela 2: Tamanho das frotas por base

| Base 1 | Base 2 | Base 3 | Base 4 |
|--------|--------|--------|--------|
| 80     | 10     | 60     | 25     |

Tabela 3: Dados dos itens 1 e 2

| Dado  | Item 1 |        |        |        | Item 2 |        |        |        |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
|       | Base 1 | Base 2 | Base 3 | Base 4 | Base 1 | Base 2 | Base 3 | Base 4 |
| $m_j$ | 45,2   | 6,1    | 23,4   | 12,5   | 36,7   | 8,3    | 17,7   | 15,5   |
| $T_j$ | 0,077  | 0,144  | 0,114  | 0,069  | 0,103  | 0,135  | 0,142  | 0,070  |
| $r_j$ | 0,40   | 0,10   | 0,30   | 0,20   | 0,30   | 0,15   | 0,25   | 0,20   |
| $C_j$ | 30     | 30     | 30     | 30     | 150    | 150    | 150    | 150    |

Tabela 4: Dados dos itens 3 e 4

| Dado  | Item 3 |        |        |        | Item 4 |        |        |        |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
|       | Base 1 | Base 2 | Base 3 | Base 4 | Base 1 | Base 2 | Base 3 | Base 4 |
| $m_j$ | 80,8   | 13,6   | 72,1   | 10,7   | 28,3   | 11,3   | 53,2   | 10,2   |
| $T_j$ | 0,077  | 0,123  | 0,130  | 0,075  | 0,113  | 0,101  | 0,085  | 0,086  |
| $r_j$ | 0,20   | 0,20   | 0,20   | 0,20   | 0,60   | 0,10   | 0,40   | 0,20   |
| $C_j$ | 85     | 85     | 85     | 85     | 65     | 65     | 65     | 65     |

Tabela 5: *Lead-time* entre as bases

| Item   | Bases 1 e 2 | Bases 1 e 3 | Bases 1 e 4 | Bases 2 e 3 | Bases 2 e 4 | Bases 3 e 4 |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|
| Item 1 | 0,060       | 0,057       | 0,042       | 0,059       | 0,073       | 0,058       |
| Item 2 | 0,110       | 0,079       | 0,074       | 0,090       | 0,118       | 0,083       |
| Item 3 | 0,071       | 0,050       | 0,047       | 0,056       | 0,071       | 0,074       |
| Item 4 | 0,094       | 0,076       | 0,075       | 0,076       | 0,122       | 0,102       |

Para as simulações utilizando o modelo *METRIC*, uma das bases deve ser escolhida para ser o *Hub*. Portanto, quando se utiliza a base 1 como *Hub*, por exemplo, a probabilidade de reparo na base 1 ( $r_1$ ) é alterada para 1, e o tempo médio de reparo no *Hub*, é alterado para  $T_1$ .

## 6. RESULTADOS

Foram simulados os casos em que cada um dos itens é controlado separadamente, além de todas as combinações possíveis envolvendo dois, três e quatro itens. Em todas as simulações, buscou-se atender uma disponibilidade mínima de 99%. Primeiramente, é realizada uma análise para definir o melhor posicionamento do *Hub* no *METRIC*. Em seguida, é apresentado um comparativo entre os resultados dos modelos.

### 6.1. POSICIONAMENTO DO HUB NO MODELO *METRIC*

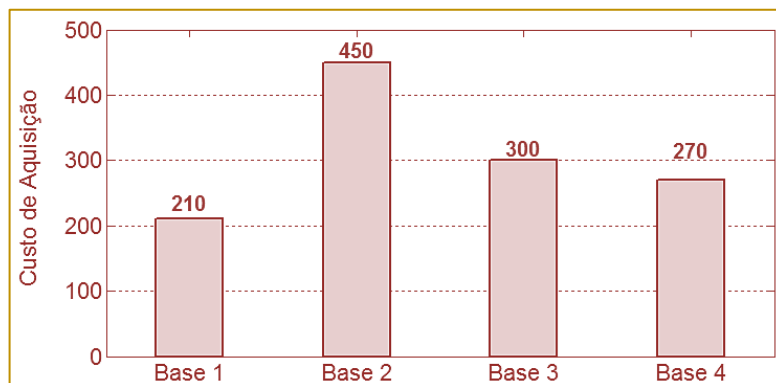
Para identificar a melhor posição para o *Hub*, deve ser feita uma simulação para cada uma das possíveis localizações. A Figura 6 mostra os custos obtidos para as diferentes localizações do *Hub*, considerando apenas o item 1. Nota-se que a configuração que resulta no menor custo é aquela na qual o *Hub* é colocado na base 1. Esse resultado se repetiu para todos os cenários simulados. Os resultados também ilustram o impacto que a escolha da localização do *Hub* tem no resultado final do modelo *METRIC*.

O fato de a base 1 possuir a maior demanda entre as bases contribui para ela seja a melhor escolha para a localização do *Hub*. Isto ocorre pois, quando as peças necessitam ser reparadas no *Hub*, os tempos de transporte entre os estoques afetam significativamente a disponibilidade das peças nas bases avançadas. Obviamente,

dependendo das relações entre custo da peça, probabilidade de reparo na base e tempos de transporte, uma base que não

possui a maior demanda pode, ainda assim, ser a mais indicada para o posicionamento do *Hub*.

Figura 6: Avaliação da melhor localização do *Hub* para o item 1



## 6.2. COMPARAÇÃO ENTRE OS MODELOS *SINGLE SITE E METRIC*

A Tabela 6 mostra o número de peças obtido na simulação do modelo *Single Site* para cada uma das bases, considerando

novamente o cenário no qual apenas o item 1 é controlado. O custo total neste exemplo é 450. É possível notar que as bases com maior demanda são as que necessitam de maior número de peças para atingir a disponibilidade requerida.

Tabela 6: Número de peças em cada base no modelo *Single Site*

| Parâmetro       | Base 1 | Base 2 | Base 3 | Base 4 |
|-----------------|--------|--------|--------|--------|
| Número de Peças | 5      | 3      | 4      | 3      |
| Custo           | 150    | 90     | 120    | 90     |

Para a simulação do modelo *METRIC* foi usada a configuração na qual o *Hub* é posicionado na base 1, por ser a configuração de menor custo. A Tabela 7 mostra o número de peças calculado para cada base (incluindo a base 1, que atua como *Hub*) no cenário no qual apenas o item

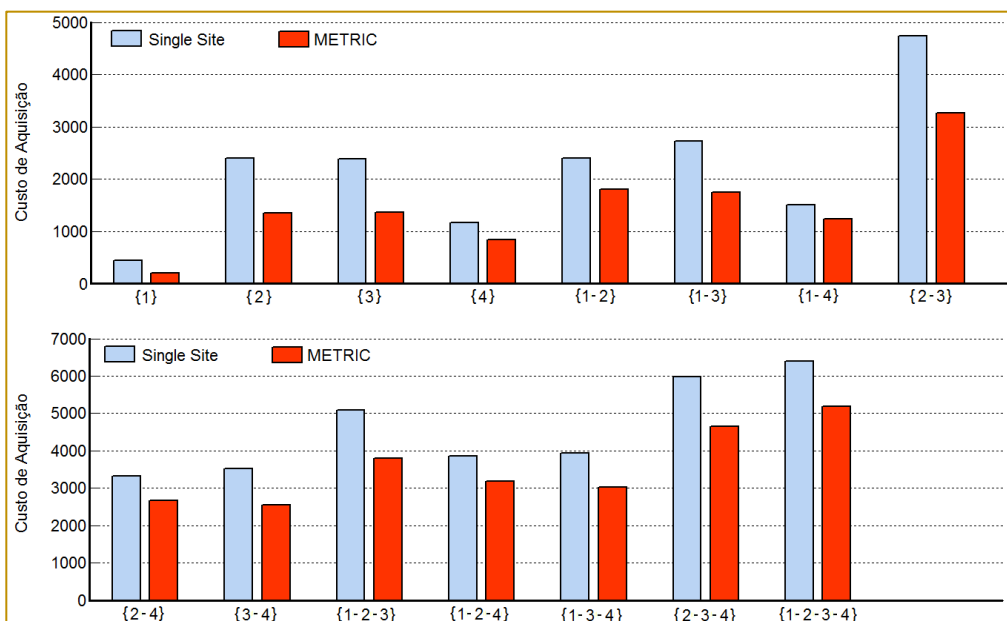
1 é controlado. O custo total neste exemplo é 210. Enquanto no modelo *METRIC* a disponibilidade mínima foi atingida com apenas 7 peças, para o modelo *Single Site* seriam necessárias 15 peças. Esta diferença no número de peças acarreta no aumento dos custos.

Tabela 7: Número de peças em cada base no modelo *METRIC*

| Parâmetro       | <i>Hub</i> | Base 2 | Base 3 | Base 4 |
|-----------------|------------|--------|--------|--------|
| Número de Peças | 4          | 0      | 2      | 1      |
| Custo           | 120        | 0      | 60     | 30     |

A Figura 7 mostra os custos totais obtidos com os dois modelos para todas as combinações possíveis de itens. Em todas as simulações do modelo *METRIC* o *Hub* está

localizado na base 1. A Tabela 8 apresenta a redução de custo de aquisição de peças proporcionada pelo modelo *METRIC* para todas as simulações.

Figura 7: Comparação de custos entre os modelos *Single Site* e *METRIC*Tabela 8: Redução de custo obtida com o modelo *METRIC*

| Itens | Redução de custo | Itens | Redução de custo | Itens     | Redução de custo |
|-------|------------------|-------|------------------|-----------|------------------|
| {1}   | 53,3%            | {1-3} | 35,8%            | {1-2-3}   | 25,3%            |
| {2}   | 43,8%            | {1-4} | 18,2%            | {1-2-4}   | 17,7%            |
| {3}   | 42,9%            | {2-3} | 31,0%            | {1-3-4}   | 23,3%            |
| {4}   | 27,8%            | {2-4} | 19,5%            | {2-3-4}   | 22,2%            |
| {1-2} | 25,0%            | {3-4} | 27,4%            | {1-2-3-4} | 19,0%            |

## 7. CONCLUSÕES

Neste trabalho foi realizado um estudo sobre o uso de um modelo multi-elos para gestão de estoques de peças de reposição. Simulações foram realizadas com o objetivo de comparar os resultados obtidos com o modelo multi-elos, que é uma variação do modelo *METRIC*, com os resultados obtidos com um modelo *Single Site*.

Observou-se que ambos os modelos têm suas vantagens e desvantagens. Por isso, é de extrema importância a utilização de uma ferramenta computacional capaz fornecer as informações necessárias para uma tomada de decisão correta. A partir dos modelos desenvolvidos, é possível fazer uma primeira análise de qual seria a melhor abordagem e a respectiva configuração ideal para o posicionamento das peças de reposição para que se possa reduzir os custos sem deixar de atender o requisito de disponibilidade.

No exemplo que considera apenas o item 1, usando o modelo *METRIC* foi possível atender o requisito de disponibilidade com 7 peças, enquanto usando o modelo *Single Site* foram necessárias 15 peças. Nesse exemplo, o uso do modelo *METRIC* proporcionou uma redução de cerca de 53% nos custos relativos à aquisição de peças de reposição. Constatou-se que a escolha da localização do *Hub* tem grande influência no resultado final.

Observou-se ainda que, nas simulações com o modelo *METRIC*, apesar da disponibilidade global atender o requisito de disponibilidade, a disponibilidade observada em cada base pode variar. Em alguns casos, algumas bases poderão ter disponibilidades menores que a disponibilidade mínima exigida. Em geral, valores mais baixos de disponibilidade foram observados em bases com frotas pequenas e em bases distantes do *Hub*.

## REFERÊNCIAS

- [1] Clark, A. e Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6:475–490.
- [2] Ehrhardt, R. A., Schultz, C. R. e Wagner, H. M. (1981). (S, s) Policies for a wholesale inventory system. *TIMS Studies in the Management Science*, 16:145–161.
- [3] Ganeshan, R. (1999). Managing supply chain inventories: a multiple retailer, one warehouse, multiple supplier model. *International Journal of Production Economics*, 59:341–354.
- [4] Kennedy, W. J., Patterson, J. W. e Fredendall, L. D. (2002). An overview of recent literature on spare parts inventories. *International Journal of Production Economics*, 76:201–215.
- [5] Kilpi, J., Töyli, J. e Vepsäläinen, A. (2009). Cooperative strategies for the availability service of repairable aircraft components. *International Journal of Production Economics*, 117:360–370.
- [6] Kilpi, J. e Vepsäläinen, A. P. J. (2004). Pooling of spare components between airlines. *Journal of Air Transport Management*, 10:137–146.
- [7] Kinnison, H. A. (2004). *Aviation Maintenance Management*. McGraw-Hill, Nova York.
- [8] Kutanoglu, E. (2008). Insights into inventory sharing in service parts logistics systems with time-based service levels. *Computers & Industrial Engineering*, 54:341–358.
- [9] Lombardo, D. A. (2000). European airlines: third-party maintenance providers. *Aviation Maintenance*, 19(7):22–25.
- [10] Sherbrooke, C. C. (2004). *Optimal inventory modeling of systems: multi-echelon techniques*. Klumer Academic Pub., Nova York.
- [11] Snyder, L. V. e Shen, Z. J. M. (2011). *Fundamentals of Supply Chain Theory*. John Wiley & Sons, Hoboken, New Jersey.
- [12] Wong, H., Cattrysse, D. e Van Oudheusden, D. (2005). Inventory pooling of repairable spare parts with non-zero lateral transshipment time and delayed lateral transshipments. *European Journal of Operational Research*, 165:207–218.



# Capítulo 4

## IDENTIFICATION OF NONLINEAR SYSTEMS USING THE TEACHING-LEARNING BASED OPTIMIZATION ALGORITHM

*Leonardo Ramos Rodrigues*

*Takashi Yoneyama*

*Lucas Sousa de Oliveira*

**Abstract:** This paper uses the recently proposed Teaching-Learning Based Optimization (TLBO) algorithm to solve the nonlinear system parameter identification problem. It is assumed that the structure of the system is known. Numerical experiments are performed to evaluate the performance of the TLBO algorithm applied to two nonlinear models already studied in the literature, namely the Output Error Polynomial (OEP) and the Output Error Rational (OER) models. Results obtained in a previous work using the Cuckoo Search via Lévy Flights algorithm are considered to establish a comparative performance baseline. The results obtained in the numerical experiments show that the TLBO algorithm provided good accuracy in both models, outperforming the baseline for the OER model. The TLBO algorithm is easy to implement, and the reduced number of parameters simplifies the parameter tuning process. Based on the results, the TLBO may be considered as an attractive alternative for the nonlinear system parameter identification problem.

## 1. INTRODUCTION

The parameter identification problem for linear systems is a well-established area of study [Ljung, 2010]. A set of tools is available for linear system identification such as modal testing [Ewins, 2000] and modal analysis [Maia and Silva, 1997]. For linear systems, the transfer function which relates the inputs with the outputs of the system remains constant. Thus, the mathematical model obtained in the system identification process can be used for any input level [Gondhalekar, 2009].

For nonlinear systems, however, in most cases it is not possible to obtain a mathematical model of the system by performing the system identification process at a single input level. The model obtained at a given operating point can, at best, provide information about the equivalent linear system at that operating point [Gondhalekar, 2009]. The nonlinear system identification problem can be divided into three different stages [Gondhalekar, 2009]:

- 1 - Nonlinearity detection;
- 2 - Nonlinearity characterization; and
- 3 - Nonlinear parameter estimation.

In the first stage, the presence of nonlinearities in the system is detected. Many methods are available in the literature for this stage. A common method used to detect the presence of nonlinearities in the system is the verification of fundamental principles of linear systems such as linear superposition and reciprocity. A summary of nonlinearity detection techniques is presented in Vanhoenacker et al. [2002].

In the second stage, the main goal is to identify the location, the type, and the functional form of all nonlinearities in the system [Kerschen et al., 2006]. Nonlinearity characterization is an important stage in nonlinear system identification problems. Some of the methods used in the first stage may be extended in order to be used in the second stage. For simple models, the location of nonlinearity can be identified by looking at the model structure. For complex models, however, mathematical techniques need to be applied in order to identify the location of nonlinearities. Some works have been published with methods for nonlinearities characterization such as Adams and Allemang [1999], Tanrikulu and Ozguven [1991]. There is still a lack of validated techniques for nonlinearity characterization of

complex systems, and subjective judgments are often made about the location and type of nonlinearities for complex systems.

In the third stage, the goal is to estimate the parameter values of the nonlinear model. Nonlinear parameter extraction techniques based on experimental measurements are commonly used in this stage. The techniques used in this stage can be divided into spatial methods and modal methods. Most of spatial methods are in time-domain, such as the methods proposed by Masri and Caughey [1979] and by Billings and Leontaritis [1985]. Some modal methods can be found in Gilbert [2003] and Kerschen et al. [2009].

This paper is focused on the third stage (nonlinear parameter estimation). Some works have been proposed using metaheuristic methods to estimate parameters of nonlinear systems [Souza et al., 2014]. Good results have been achieved by methods using Artificial Neural Networks [Samad and Mathur, 1992], Genetic Algorithms [Yao and Sethares, 1994] and Particle Swarm Optimization [Schwaab et al., 2008], among other metaheuristics. In this paper, the recently proposed TLBO algorithm is used in order to estimate the parameters of nonlinear systems. Numerical experiments with nonlinear models already discussed in the literature are carried out in order to evaluate the performance of the TLBO algorithm. Results obtained in a previous work using the Cuckoo Search via Lévy Flights algorithm [Souza et al., 2014] are used in order to establish a comparative performance baseline.

The remaining sections of this paper are organized as follows. Section 2 describes the parameter identification problem under consideration. Section 3 presents the basic principles of the TLBO algorithm. Section 4 presents the methodology used in this paper to tune the parameters of the TLBO algorithm. Section 5 shows the results obtained in the numerical experiments in which the TLBO algorithm was applied to two nonlinear problems already studied in the literature. Concluding remarks are presented in section 6.

## 2. PROBLEM DEFINITION

The parameter identification problem addressed in this paper can be stated as follows. Consider a discrete nonlinear system  $A$  with input  $u(k)$  and output  $y(k)$ . A

mathematical model of system A is available. The output  $y(k)$  is a function of past inputs, past outputs and measurements noise. The system A can be represented by a mathematical model based on the nonlinear difference equation model presented in Equation (1).

$$y(k) = g(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u)) + e(k) \quad (1)$$

where  $g(\cdot)$  is a nonlinear function,  $n_y$  and  $n_u$  are the maximum lags in the output and input, respectively, and  $e(k)$  is the system noise, which is assumed to be a white noise sequence.

Consider a parameter vector  $\theta$  containing the set of parameters of the mathematical model of system A which are unknown. Each element of  $\theta$  is a real number. Therefore, the problem consists in estimating the values of each element of parameter vector  $\theta$ .

A dataset containing an input signal  $u(k)$  and the corresponding output signal  $y(k)$  is available. Thus, it is possible to define an objective function J which quantifies, for a given  $\theta$ , the difference between the output obtained with  $u(k)$  and the real output, when  $u(k)$  is used as input. When the objective function J is minimized, an estimation for  $\theta$  can be obtained.

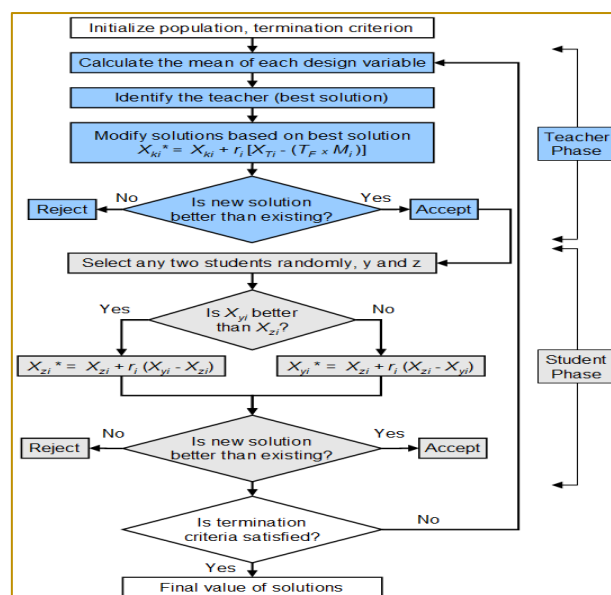
### 3. TEACHING-LEARNING BASED OPTIMIZATION ALGORITHM

The TLBO algorithm has been recently proposed in the literature as a novel population oriented metaheuristic based on the teaching-learning process observed in a classroom [Rao et al., 2011]. This algorithm simulates the influence of a teacher on the output of a group of students in a class. The TLBO algorithm is divided into two main parts: the Teacher Phase and the Student Phase, which is also known as the Learner Phase [Rao and Patel, 2013]. During the Teacher Phase, students learn from the teacher, while in the Learner Phase students learn through the interaction among themselves.

There is a solution X associated with each student, which corresponds to a possible solution for the optimization problem under consideration. Also, there is a result  $f(X)$  associated with each solution (or student), which can be obtained by evaluating the solution X using the objective function J. In the parameter estimation problem considered in this paper, a solution X corresponds to a candidate parameter vector  $\theta$ .

The student with the best solution in the population is considered as the Teacher. The steps for implementing the TLBO algorithm are presented in Figure 1 [Rao et al., 2011]. The Teacher Phase and the Student Phase are described in the next sections.

Figure 1: TLBO flowchart



### 3.1. TEACHER PHASE

In this phase, the algorithm simulates the learning of the students from the teacher (best solution). During this phase, the teacher makes an effort to increase the mean result of the class.

Consider a group of  $n$  students. Let  $M_i$  be the mean solution of the students and  $T_i$  be the

teacher at iteration  $i$ . The teacher  $T_i$  will make an effort to move  $M_i$  to its own level. Knowledge is gained based on the quality of the teacher and the quality of students in the class. The difference  $D_i$  between the solution of the teacher,  $X_{Ti}$ , and the mean solution of the students,  $M_i$ , can be expressed according to Equation (2):

$$D_i = r_i(X_{Ti} - T_F \cdot M_i) \quad (2)$$

where  $r_i$  is a random number in the range  $[0, 1]$  for iteration  $i$  and  $T_F$  is a teaching factor for

iteration  $i$ , which is randomly set to either 1 or 2 according to Equation (3):

$$T_F = \text{round}(1 + \text{rand}(0,1)) \quad (3)$$

Based on the difference  $D_i$ , the existing solution of each student  $k$  in iteration  $i$ ,  $X_{ki}$ ,

with  $k \in \{1, 2, \dots, n\}$ , is updated in the teacher phase according to Equation (4):

$$X_{ki}^* = X_{ki} + D_i \quad (4)$$

where  $X_{ki}^*$  is the updated value of  $X_{ki}$ .

If  $f(X_{ki}^*)$  is better than  $f(X_{ki})$ ,  $X_{ki}^*$  is accepted and replaces  $X_{ki}$ . Otherwise,  $X_{ki}^*$  is discarded and  $X_{ki}$  is not changed for the next iteration.

in iteration  $i$ , respectively. If  $f(X_{yi})$  is better than  $f(X_{zi})$ , the solution of student  $z$  is updated according to Equation (5). Then, if  $f(X_{zi}^*)$  is better than  $f(X_{zi})$ ,  $X_{zi}^*$  is accepted and replaces  $X_{zi}$ . Otherwise,  $X_{zi}^*$  is discarded and  $X_{zi}$  is not changed for the next iteration. Similarly, if  $f(X_{zi})$  is better than  $f(X_{yi})$ , the solution of student  $y$  is updated according to Equation (6). Then, if  $f(X_{yi}^*)$  is better than  $f(X_{yi})$ ,  $X_{yi}^*$  is accepted and replaces  $X_{yi}$ . Otherwise,  $X_{yi}^*$  is discarded and  $X_{yi}$  is not changed for the next iteration.

### 3.2. STUDENT PHASE

In this phase, the algorithm simulates the learning of the students through interaction with one another. During this phase, students gain knowledge by discussing with other students who have more knowledge [Rao and Patel, 2013].

Consider a pair of students  $y$  and  $z$ . Let  $X_{yi}$  and  $X_{zi}$  be the solutions of students  $y$  and  $z$

$$X_{zi}^* = X_{zi} + r_i(X_{yi} - X_{zi}) \quad (5)$$

$$X_{yi}^* = X_{yi} + r_i(X_{zi} - X_{yi}) \quad (6)$$

At the end of each iteration, the stop criteria are checked. Different stop criteria may be adopted. Some of the most commonly adopted stop criteria are the maximum number of iterations, the maximum number of successive iterations without any improvement, the maximum simulation time

and the maximum number of function evaluations.

In this paper, the maximum number of function evaluations is adopted as the stop criterion, as in other applications of TLBO in combinatorial

problems [Baykasoglu et al., 2014], [Crawford et al., 2015], [Patil, 2016].

#### 4. PARAMETERS TUNING FOR THE TLBO

A metric commonly used for comparing the performance of different metaheuristics for solving continuous optimization problems is the number of function evaluations.

In this paper, the results obtained with TLBO will be compared with the results obtained by Souza et al. [2014]. In their work, the authors proposed a modified version of the cuckoo search optimization algorithm, which was originally proposed by Yang and Deb [2009]. The stop criterion adopted by [Yang and Deb, 2009] is based on the number of generations without a significant improvement in the

objective function. They used a population size  $p$  of 15, a discovery rate  $r$  of 25% and the algorithm stopped when no significant improvement was observed in the last 1,000 generations.

Ong [2014] tested the convergence of the cuckoo search algorithm for various optimization problems. The average number of generations needed for the algorithm to converge was about 1,500. Thus, the average total number of generations,  $g$ , needed for the stop criterion adopted in Souza et al. [2014] to be reached is estimated to be approximately 2,500.

The total number of functions evaluations in the cuckoo search,  $n_{CS}$ , can be estimated using Equation (7).

$$n_{CS} = p + g[1 + (r \times p)] \quad (7)$$

Therefore, based on the parameter values, the estimation of the number of function evaluations in the cuckoo search algorithm is 11,890. In this paper, in order to provide a fair comparison, the stop criterion adopted for the TLBO algorithm is also the maximum number of function evaluations. It was decided to use 10,000 as the maximum number of function evaluations allowed.

In order to define the population size for the TLBO algorithm,  $PS$ , ten different possible values were considered: [10, 20, 30, ..., 100]. For each value of  $PS$ , a Monte Carlo method with 30 iterations was carried out. The best results for both the OEP and the OER models (that will be discussed in more details in section 5) were obtained for  $PS = 50$ . This value was then adopted to run the experiments that will be presented in section 5.

#### 5. NUMERICAL EXPERIMENTS

This section presents a set of numerical experiments to investigate the performance of

the TLBO algorithm applied to the parameter identification of nonlinear systems problem. Two different models are used in the experiments: the Output Error Polynomial (OEP) model and the Output Error Rational (OER) model. The OEP model is linear in parameters, while the OER model is not.

Both the OEP and the OER models were used by Souza et al. [2014]. As mentioned earlier, the results obtained with TLBO will be compared with the results obtained by Souza et al. [2014] using a modified version of the cuckoo search optimization algorithm.

##### 5.1. OUTPUT ERROR POLYNOMIAL (OEP) MODEL

In this experiment, the Output Error Polynomial (OEP) model originally presented in Piroddi and Spinelli [2003] is considered. The OEP model is described by Equations (8) and (9). Figure 2 shows the Matlab Simulink implementation of the OEP model used to run the experiments.

$$w(k) = 0.75w(k-2) + 0.25u(k-1) - 0.2w(k-2)u(k-1) \quad (8)$$

$$y(k) = w(k) + e(k) \quad (9)$$

Figure 1: OEP model

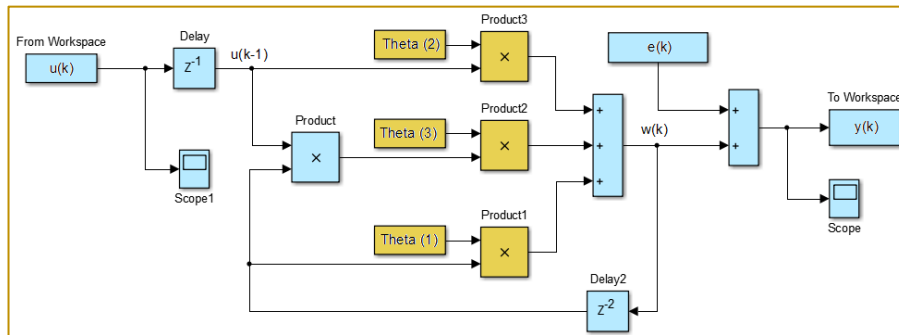
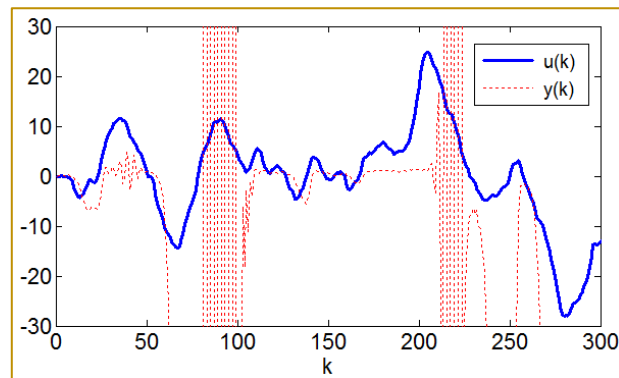


Figure 3 shows the signal  $u(k)$  used as input for the OEP model, as well as the corresponding output signal,  $y(k)$ . One characteristic of the OEP model is that it presents an unstable behavior, which brings an additional difficulty to the identification

problem because it may generate multiple local minima in the objective function. In this example, the minimum and the maximum values of  $y(k)$  are  $-3.67 \times 10^{12}$  and  $6.93 \times 10^3$ , respectively.

Figure 2: Unstable behavior of the OEP model



A series of 300 measurements was used in the experiments. The error  $e(k)$  is a white Gaussian noise with mean zero and variance 0.25. A Monte Carlo method with 1,000

iterations was used. Tables 1 and 2 show the results obtained with the Cuckoo Search and the TLBO algorithm, respectively.

Tabela 1: OEP model results with cuckoo search

| Parameter  | Reference | Minimum Value | Maximum Value | Mean Value | Standard Deviation |
|------------|-----------|---------------|---------------|------------|--------------------|
| $\theta_1$ | 0.75      | 0.7497        | 0.7500        | 0.7500     | 0.000              |
| $\theta_2$ | 0.25      | 0.0316        | 0.3408        | 0.2336     | 0.0402             |
| $\theta_3$ | -0.20     | -0.2001       | -0.2000       | -0.2000    | 0.0000             |

Tabela 2: OEP model results with TLBO

| Parameter  | Reference | Minimum Value | Maximum Value | Mean Value | Standard Deviation |
|------------|-----------|---------------|---------------|------------|--------------------|
| $\theta_1$ | 0.75      | 0.7490        | 0.7515        | 0.7499     | 0.0009             |
| $\theta_2$ | 0.25      | 0.0557        | 0.3342        | 0.2424     | 0.0508             |
| $\theta_3$ | -0.20     | -0.2016       | -0.1994       | -0.2004    | 0.0006             |

It can be noted from the results presented in Tables 1 and 2 that, for the OEP model, the TLBO algorithm provided an excellent accuracy in the estimates of parameters  $\theta_1$  and  $\theta_3$ . It can also be noted that estimating the value of parameter  $\theta_2$  is the most challenging task for the OEP model. This fact was also observed in the results obtained by Souza et al. [2014]. For parameter  $\theta_2$ , the TLBO algorithm provided a better estimate. Although the standard deviation obtained with the TLBO algorithm was higher than the one obtained with the cuckoo search algorithm, the TLBO provided a more accurate estimate of the mean value. Additionally, the interval defined by the minimum and the maximum

estimated values was narrower for the TLBO algorithm.

### 5.1. OUTPUT ERROR RATIONAL (OER) MODEL

In this experiment, the Output Error Rational (OER) model originally presented in Zhu [2005] is considered. The OER model is described by Equations (10) and (11). Figure 4 shows the Matlab Simulink implementation of the OER model used to run the experiments.

$$w(k) = \frac{0.3w(k-1)w(k-2)+0.7u(k-1)}{1+w(k-1)^2+u(k-1)^2} \quad (10)$$

$$y(k) = w(k) + e(k) \quad (11)$$

For the OER model, a random number generation block was used to generate an uniformly distributed input signal with mean zero and variance 0.33. A series of 1,000 measurements were used in the experiments. The error  $e(k)$  is a white Gaussian noise with

mean zero and variance 0.01. A Monte Carlo method with 1,000 iterations was used. Tables 3 and 4 show the results obtained with the Cuckoo Search and the TLBO algorithm, respectively.

Figure 1: OER simulink model

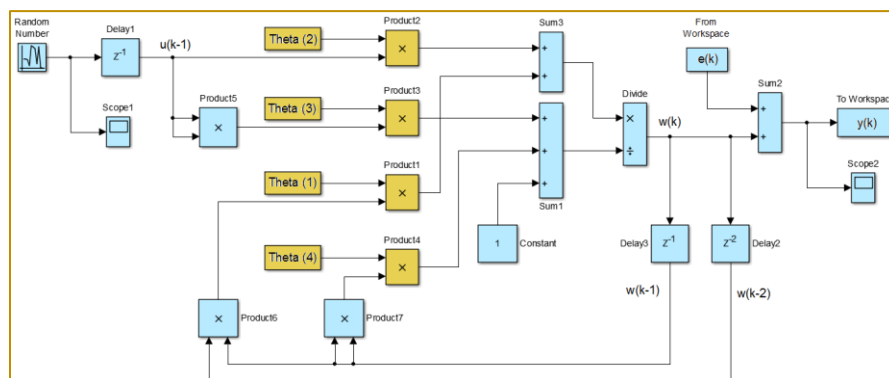


Tabela 1: OER model results with cuckoo search

| Parameter  | Reference | Minimum Value | Maximum Value | Mean Value | Standard Deviation |
|------------|-----------|---------------|---------------|------------|--------------------|
| $\theta_1$ | 0.3       | 0.0874        | 0.5552        | 0.3008     | 0.0745             |
| $\theta_2$ | 0.7       | 0.6182        | 0.7882        | 0.6981     | 0.0266             |
| $\theta_3$ | 1.0       | -0.4115       | 2.5834        | 0.9693     | 0.5121             |
| $\theta_4$ | 1.0       | 0.7476        | 1.2415        | 0.9971     | 0.0837             |

Tabela 2: OER model results with TLBO

| Parameter  | Reference | Minimum Value | Maximum Value | Mean Value | Standard Deviation |
|------------|-----------|---------------|---------------|------------|--------------------|
| $\theta_1$ | 0.3       | 0.2572        | 0.3228        | 0.2981     | 0.0184             |
| $\theta_2$ | 0.7       | 0.6526        | 0.7106        | 0.6962     | 0.0092             |
| $\theta_3$ | 1.0       | 0.6084        | 1.5208        | 0.9795     | 0.1587             |
| $\theta_4$ | 1.0       | 0.8721        | 1.1368        | 0.9923     | 0.0440             |

Based on the results presented in Table 3 and Table 4, it can be said that, for the OER model, the TLBO algorithm outperformed the cuckoo search method. The estimated mean values obtained with the TLBO algorithm were very close to the real values. Also, when compared with the results obtained with the cuckoo search method, the standard deviation of all parameters were smaller, and the interval defined by the minimum and the maximum estimated values were narrower for all parameters. Also, it can be observed that the results are consistent with the results obtained in Souza et al. [2014] in terms of the relative difficulty to estimate the parameters of the OER model, with  $\theta_3$  being the most challenging parameter to be estimated.

## 6. CONCLUSIONS

In this paper, the performance of the recently proposed Teaching-Learning Based Optimization (TLBO) algorithm to solve the nonlinear system parameter identification problem was evaluated. It was assumed that the structure of the system is known.

Numerical experiments were carried out using two nonlinear models already studied in the literature, namely the Output Error Polynomial (OEP) and the Output Error Rational (OER) models. The OEP model has an unstable behavior, while the OER model is nonlinear in parameters. Those characteristics increase the complexity of the parameter identification

process. The results obtained with the TLBO algorithm were compared with the results obtained by Souza et al. [2014] with the Cuckoo Search via Lévy Flights algorithm.

The TLBO algorithm provided an excellent accuracy in the estimates of parameters  $\theta_1$  and  $\theta_3$  for the OEP model. For parameters  $\theta_2$ , which is the most challenging in the OEP model, the TLBO algorithm presented a better performance. For the OER model, the TLBO algorithm outperformed the Cuckoo Search via Lévy Flights algorithm method for all parameters, providing good estimates with smaller standard deviation and narrower intervals for all parameters.

The TLBO algorithm was easy to test and implement. Also, the reduced number of parameters simplified the parameter tuning process. Based on the results, it can be concluded that the TLBO may be considered as an attractive alternative to solve the nonlinear system parameter identification problem.

In this paper, the TLBO algorithm was used in its basic version. One possible suggestion for extending this paper is to evaluate the performance of the TLBO algorithm, considering the improvements that have been proposed in the literature. Numerical experiments to compare the results obtained with the TLBO with other metaheuristics could also be conducted.

## REFERENCES

- [1] Adams, D. E. and Allemang, R. J. (1999). Characterization of nonlinear vibrating systems using internal feedback and frequency response modulation. *Journal of Vibration and Acoustics*, 121: 495–500.
- [2] Baykasoğlu, A., Hamzadayı, A., and Köse, S. Y. (2014). Testing the performance of teaching learning based optimization (TLBO) algorithm on combinatorial problems: Flow shop

and job shop scheduling cases. *Information Sciences*, 276:204–218.

- [3] Billings, S. A. and Leontaritis, I. J. (1985). Input-output parametric models for nonlinear systems part I - deterministic nonlinear systems. *International Journal of Control*, 41(2):303–328.
- [4] Crawford, B., Soto, R., Leiva, F. A., Johnson, F., and Paredes, F. (2015). Problema del conjunto de cobertura resuelto mediante el algoritmo binario de optimización basado en



enseñanza aprendizaje. In Conferência Ibérica de Sistemas e Tecnologias de Informação, p. 106–109, Águeda. IEEE.

- [5] Ewins, D. J. (2000). *Modal Testing: Theory, Practice, and Application*. Research Studies Press, 2nd edition.
- [6] Gilbert, C. (2003). Fitting measured frequency response using non-linear modes. *Mechanical Systems and Signal Processing*, 17(1):211–218.
- [7] Gondhalekar, A. C. (2009). *Strategies for Non-linear Systems Identification*. PhD thesis, University of London.
- [8] Kerschen, G., Peeters, M., Golinval, J. C., and Vakasis, A. F. (2009). Nonlinear normal modes, part I: A useful framework for the structural dynamicist. *Mechanical Systems and Signal Processing*, 23:170–194.
- [9] Kerschen, G., Worden, K., Vakasis, A. F., and Golinval, J. C. (2006). Past, present and future of nonlinear identification in structural dynamic. *Mechanical Systems and Signal Processing*, 20:505–592.
- [10] Ljung, L. (2010). Approaches to identification of nonlinear systems. In *Proceedings of the 29th Chinese Control Conference*, p. 1–5, Beijing. IEEE.
- [11] Maia, N. M. M. and Silva, J. M. M. (1997). *Theoretical and Experimental Modal Analysis*. Taunton, Somerset, England, Research Studies Press Ltd.
- [12] Masri, S. F. and Caughey, A. (1979). A non parametric identification technique for nonlinear dynamic problems. *The Scientific World Journal*, 46.
- [13] Ong, P. (2014). Adaptive cuckoo search algorithm for unconstrained optimization. *The Scientific World Journal*, 2014.
- [14] Patil, A. B. (2016). Teaching learning based optimization: Application and variation. In *Proceedings of the International Conference on Computing, Communication and Energy Systems*, p. 1–5, Sangli.
- [15] Piroddi, L. and Spinelli, W. (2003). An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781.
- [16] Rao, R. V. and Patel, V. (2013). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. *Scientia Iranica*, 20:710–720.
- [17] Rao, R. V., Vakharia, D. P., and Savsani, V. J. (2011). Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43: 303–315.
- [18] Samad, T. and Mathur, A. (1992). Parameter estimation for process control with neural networks. *International Journal of Approximate Reasoning*, 3:149–164.
- [19] Schwaab, M., Biscaia, E. C., Monteiro Jr., J. L., and Pinto, J. C. (2008). Nonlinear parameter estimation through particle swarm optimization. *Chemical Engineering Science*, 63(3):1542–1552.
- [20] Souza, D. K., Leandro, G. V., Oliveira, G., and Ribeiro, E. P. (2014). Estimação de parâmetros de sistemas não lineares por busca de cuco via voos de Lévy. In *Anais do XX Congresso Brasileiro de Automática*, p. 769–776, Belo Horizonte.
- [21] Tanrikulu, O. and Ozguven, H. N. (1991). A new method for the identification of nonlinearities in vibrating structures. In *Proceedings of the 4th International Conference on Structural Dynamics*, p. 483–492, Southampton.
- [22] Vanhoenacker, K., Schoukens, J., Swevers, J., and Vaes, D. (2002). Summary and comparing overview of techniques for the detection of non-linear distortions. In *Proceedings of ISMA 2002*, volume 3, p. 1241–1255, Leuven.
- [23] Yang, X. and Deb, S. (2009). Cuckoo search via Lévy flights. In *Proceeding of the World Congress on Nature Biologically Inspired Computing*, p. 210–214.
- [24] Yao, L. and Sethares, W. (1994). Nonlinear parameter estimation via the genetic algorithm. *IEEE Transactions on Signal Processing*, 42(4):927–935.
- [25] Zhu, Q. (2005). An implicit least squares algorithm for nonlinear rational model parameter estimation. *Applied Mathematical Modelling*, 29(7):673–689.

# Capítulo 5

## *UM ALGORITMO HÍBRIDO BASEADO EM OTIMIZAÇÃO POR NUVEM DE PARTÍCULAS E BUSCA TABU PARA O PROBLEMA DA DIVERSIDADE MÁXIMA*

*Edison Luiz Bonotto*

*Lucídio dos Anjos Formiga Cabral*

**Resumo:** Este artigo descreve um algoritmo híbrido (PSO TS) que utiliza Otimização por Nuvem de Partículas (PSO) e Busca Tabu (TS) para a resolução do Problema da Diversidade Máxima (MDP). O MDP é um problema da área de Otimização Combinatória que tem por objetivo selecionar um número pré-estabelecido de elementos de um dado conjunto de maneira tal que a soma das diversidades entre os elementos selecionados seja a maior possível. A abordagem PSO simula o comportamento de um bando de pássaros em voo com seu movimento localmente aleatório, mas globalmente determinado, para encontrar máximos locais. A utilização do PSO TS alcança a grande maioria dos melhores resultados encontrados para as instâncias testes presentes na literatura e melhorou a solução conhecida em seis casos, demonstrando assim ser competitivo com os melhores algoritmos em termos de qualidade das soluções.

**Palavras Chave.** Problema da Diversidade Máxima (MDP), Otimização por Nuvem de Partículas (PSO), Busca Tabu (TS).

## 1 INTRODUÇÃO

Este artigo descreve uma aplicação baseada em um algoritmo de Otimização por Nuvem de Partículas com refinamento das soluções através de um procedimento de busca local baseado em Busca Tabu para tratar o Problema da Diversidade Máxima.

O Problema da Diversidade Máxima encontra aplicações nos mais variados campos da ciência, como na indústria farmacêutica [Agrafiotis, 1997], na biologia, onde através de cruzamentos seletivos pode-se extrair características desejáveis em animais, plantas e outros seres vivos [Takane, 2011], e nas empresas, onde pode ser usado para selecionar equipes de trabalho heterogêneas ou para descobrir qual a combinação de produtos que gera o maior lucro possível.

Há na literatura diversas abordagens para resolver o MDP. Entre eles podemos citar:

greedy randomized adaptive search procedure (GRASP) Ghosh [1996]; Silva et al. [2004, 2007], tabu search (TS) Aringhieri e Cordone [2011]; Aringhieri et al. [2008]; Duarte e Martí [2007]; Palubeckis [2007]; Wang et al. [2012], iterated local search (ILS) Duarte et al. [2008], iterated greedy (IG) Lozano et al. [2011], scatter search (SS) Gallego et al. [2009], variable neighborhood search (VNS) Aringhieri e Cordone [2011], Brimberg et al. [2009], simulated annealing (SA) Agrafiotis [1997] e memetic algorithm (MA) Wang et al. [2014]. Martí et al. [2013] faz uma comparação detalhada entre várias heurísticas e metaheurísticas para o problema. A Tabela 1 apresenta um resumo das principais abordagens encontradas na literatura para solucionar o Problema da Máxima Diversidade.

Tabela 1: Principais abordagens encontradas na literatura para o MDP

| Artigo / ano             | Abordagem   |
|--------------------------|---|
| Kuo et al. [1993]        | Mathematical Formulation                                  |
| Ghosh [1996]             | Greedy Randomized Adaptive Search Procedure               |
| Agrafiotis [1997]        | Simulated Annealing                                       |
| Silva et al. [2004]      | Greedy Randomized Adaptive Search Procedure               |
| Duarte e Martí [2007]    | Tabu Search + Greedy Randomized Adaptive Search Procedure |
| Palubeckis [2007]        | Iterated Tabu Search                                      |
| Aringhieri et al. [2008] | Tabu Search   |
| Duarte et al. [2008]     | Iterated Local Search                                     |
| Brimberg et al. [2009]   | Variable Neighborhood Search                              |
| Gallego et al. [2009]    | Scatter Search  |
| Martí et al. [2010]      | Exact Methods   |
| Lozano et al. [2011]     | Iterated Greedy   |
| Takane [2011]            | Mathematical Formulation + Exact Methods                  |
| Wu e Hao [2013]          | Memetic Algorithm + Tabu Search                           |
| Wang et al. [2014]       | Memetic Algorithm + Tabu Search                           |

### 1.1 O PROBLEMA DA DIVERSIDADE MÁXIMA (MAXIMUM DIVERSITY PROBLEM MDP)

O Problema da Diversidade Máxima é um problema de Otimização Combinatória que tem por objetivo selecionar um número pré-estabelecido de elementos de um dado conjunto de tal maneira que a soma das diversidades entre os elementos selecionados seja a maior possível, ou seja, consiste em selecionar um subconjunto  $M$  com  $m$  elementos de um conjunto  $N$ , onde  $N = \{1, \dots, n\}$ , de tal forma que a diversidade entre os seus elementos selecionados seja máxima.

O MDP é um problema NP-Difícil, isto é, não existe algoritmo conhecido que o resolva de forma exata em tempo polinomial e, à medida

que o tamanho das instâncias aumenta, tornase proibitivo o tempo necessário para que estas sejam resolvidas. Segundo Wu e Hao [2013], o problema também é conhecido por maximum dispersion, MAX-AVG dispersion, edge-weighted clique, remote-clique, maximum edge-weighted subgraph e dense  $k$ -subgraph.

Para calcular a diversidade como uma distância é comum usar uma métrica normalizada entre os atributos dos elementos de  $N$ . Uma métrica de distância normalmente utilizada é a norma Euclidiana. Usamos a Equação (1) para definir a distância euclidiana entre os elementos  $i$  e  $j$  denotada por  $d_{ij}$ , onde  $p = 2$ ,  $k$  pertence ao conjunto  $R$ ,

$R$  é o conjunto dos atributos (Tabela 2),  $a_{ik}$  denota o  $k$ -ésimo atributo do elemento  $i$  e  $a_{jk}$

denota o  $k$ -ésimo atributo do elemento  $j$ .

$$d_{ij} = \sqrt[p]{\sum_{k \in R} (|a_{ik} - a_{jk}|)^p} \quad (1)$$

## 1.2 EXEMPLO

Nesta seção, mostraremos um exemplo de aplicação do MDP a fim de melhorar a compreensão do problema. Suponha que uma seção de recursos humanos de uma determinada empresa queira selecionar três funcionários dentre cinco pré-selecionados

para compor uma equipe de trabalho e deseja que esta equipe tenha os mais diversos atributos possíveis. Inicialmente, são definidos quais atributos são considerados importantes para que esta equipe seja a mais heterogênea possível. Em seguida, define-se um peso para cada atributo, de acordo com sua relevância.

Tabela 2: Mapeamento de atributos pessoais

| Classificação | Atributos |       |                  |                            |                    |
|---------------|-----------|-------|------------------|----------------------------|--------------------|
|               | Sexo      | Idade | Região de Origem | Áreas de Conhecimento      | Experiência (anos) |
| 1             | Masculino | 16-19 | Norte            | Ciências Exatas            | 0-1                |
| 2             | Feminino  | 20-23 | Centro-Oeste     | Ciências Humanas           | 2-3                |
| 3             |           | 24-27 | Nordeste         | Ciências Sociais Aplicadas | 4-5                |
| 4             |           | 28-31 | Sul              | Ciências da Saúde          | 6-7                |
| 5             |           | 31-34 | Sudeste          | Ciências Agrárias          | 8-10               |

A Tabela 2 mostra o mapeamento dos atributos e a Tabela 3 mostra o mapeamento dos atributos dos candidatos, onde cada linha

representa um candidato com seus respectivos atributos.

Tabela 3: Atributos mapeados dos candidatos

| Funcionário | Atributos |       |                  |                       |                    |
|-------------|-----------|-------|------------------|-----------------------|--------------------|
|             | Sexo      | Idade | Região de Origem | Áreas de Conhecimento | Experiência (anos) |
| A           | 1         | 2     | 2                | 2                     | 1                  |
| B           | 2         | 3     | 3                | 2                     | 1                  |
| C           | 2         | 1     | 1                | 3                     | 2                  |
| D           | 1         | 4     | 5                | 5                     | 3                  |
| E           | 2         | 4     | 5                | 4                     | 4                  |

Com base na Tabela 3, calculamos a distância entre cada par de candidatos. A Tabela 4 mostra os valores resultantes do

cálculo da distância Euclidiana usando a Equação (1).

Tabela 4: Matriz de distâncias entre cada par de candidatos

|   | A    | B    | C    | D    | E    |
|---|------|------|------|------|------|
| A | 0,00 | 1,73 | 2,24 | 5,10 | 5,20 |
| B | 1,73 | 0,00 | 3,16 | 4,36 | 4,24 |
| C | 2,24 | 3,16 | 0,00 | 5,57 | 5,48 |
| D | 5,10 | 4,36 | 5,57 | 0,00 | 1,73 |
| E | 5,20 | 4,24 | 5,48 | 1,73 | 0,00 |

A solução ótima para o problema é calculada analisando todas as combinações possíveis dos cinco candidatos agrupados de três em

três e escolhendo os que apresentarem a maior diversidade entre eles. Neste caso teríamos 10 combinações possíveis.

$$\binom{5}{3} = \frac{5!}{3!2!} = \frac{5 \times 4 \times 3}{3 \times 2 \times 1} = \frac{60}{6} = 10$$

A Tabela 5 apresenta os grupos juntamente com a diversidade total entre eles. A solução do problema (solução ótima) seria a escolha

dos candidatos B, C e D, pois apresentam a maior diversidade entre eles.

Tabela 5: Diversidade máxima com todas as soluções possíveis

| Combinações |              | Diversidade Total  |       |
|-------------|--------------|--------------------|-------|
| ABC         | AB + AC + BC | 1,73 + 2,24 + 3,16 | 7,13  |
| ABD         | AB + AD + BD | 1,73 + 5,10 + 4,36 | 11,19 |
| ABE         | AB + AE + BE | 1,73 + 5,20 + 4,24 | 11,17 |
| ACD         | AC + AD + CD | 2,24 + 5,10 + 5,57 | 12,91 |
| ACE         | AC + AE + CE | 2,24 + 5,10 + 5,48 | 12,82 |
| ADE         | AD + AE + DE | 5,10 + 5,20 + 1,73 | 12,03 |
| BCD         | BC + BD + CD | 3,16 + 4,36 + 5,57 | 13,09 |
| BCE         | BC + BE + CE | 3,16 + 4,24 + 5,48 | 12,88 |
| BDE         | BD + CE + DE | 4,36 + 5,48 + 1,73 | 11,57 |
| CDE         | CD + CE + DE | 5,57 + 5,48 + 1,73 | 12,78 |

## 2 MÉTODO PROPOSTO

O algoritmo proposto utiliza uma versão da meta-heurística Otimização por Nuvem de

Partículas combinada com um algoritmo de Busca Tabu para refinar as soluções. As partículas iniciais a serem utilizadas pelo PSO serão geradas por duas heurísticas de construção, sendo uma aleatória e outra gulosa. O método foi denominado PSO TS.

### 2.1 ALGORITMO PSO TS

O algoritmo PSO TS inicia com a geração de  $q$  partículas de tamanho  $m$ , onde  $Nr$  PG%

delas sendo no mínimo uma partícula é (são) criada(s) através de um algoritmo guloso simples e, nas partículas restantes, os  $m$  elementos são escolhidos aleatoriamente obedecendo aos limites da instância. A quantidade  $q$  de partículas é definida dinamicamente no início da execução do algoritmo através da Equação (2), e varia de acordo com o número de elementos  $m$  que compõem a solução, tendo como um limite mínimo de 25 partículas. Durante a criação das partículas, GBest e P Best vão sendo atualizados.

$$q = \max(25, q \times m) \quad (2)$$

Após a criação das partículas, elas são ordenadas de acordo com sua diversidade máxima e as  $N_r$  Ch T S In melhores partículas servirão como soluções de partida para um algoritmo de Busca Tabu. Caso haja melhora na diversidade máxima, a solução retornada da TS substituiu a partícula e  $GBest$  e  $PBest$  são atualizados.

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (3)$$

Onde:

$X_{id}^k$  é a posição do elemento  $i$  na dimensão  $d$  da partícula no espaço multidimensional na  $k$ -ésima iteração, e;

$$V_{id}^{k+1} = W * V_{id}^k + c_1 * Cr * (P_{id}^k - X_{id}^k) + c_2 * (1 - Cr) * (P_{gd}^k - X_{id}^k) \quad (4)$$

Onde:

$W$  é uma variável que define o peso inercial que será aplicado à velocidade da partícula e o valor é escolhido através da equação  $W = 0.9 - ((0.9 - 0.1)/k) * q$ . Esta equação, baseada em Coelho Filho [2013], faz com que nas  $k$  iterações iniciais o valor de  $W$  entre  $[0.1, \dots, 0.9]$  seja próximo do máximo e nas iterações finais fique próximo do mínimo. Isto faz com que a partícula se desloque mais rapidamente no início do processamento e consiga "escapar" de ótimos locais.

$Cr$  é escolhido através da equação  $Cr = y * Cr * (1 - Cr)$ , descrita por Chuang et al. [2011], e gera sequências de números caóticos em substituição a  $r_1$  e  $r_2$  da equação original do PSO. Inicialmente  $Cr$  é escolhido aleatoriamente entre  $[0, \dots, 1]$ . Foi usado  $y = 4$ , o que faz com que a equação tenha um comportamento caótico, mas gere a maioria dos números nas proximidades de 0 ou 1;

$c_1$  é uma constante que define a velocidade em direção à  $PBest$  e expressa o fator cognitivo da partícula;

$c_2$  é uma constante que define a velocidade em direção à  $GBest$  e expressa o fator social da partícula. Os valores das constantes  $c_1$  e  $c_2$  são mostrados na Tabela 6 da Seção 2.4;

$X_{id}^k$  é a posição atual da partícula na  $k$ -ésima iteração;

A heurística gulosa é descrita na Seção 2.2.2 e o procedimento TS é descrito com detalhes na Seção 2.3. Os valores das constantes  $q$ ,  $N_r$  P G e  $N_r$  Ch T S In são mostrados na Tabela 6 da Seção 2.4.

A cada iteração  $k$ , a partícula desloca-se pelo espaço multidimensional conforme a equação (3):

$V_{id}^k$  é a velocidade de deslocamento do elemento  $i$  na dimensão  $d$  da partícula na  $k$ -ésima iteração.

A velocidade de deslocamento da partícula é calculada pela Equação (4):

$P_{id}^k$  ( $PBest$ ) a melhor posição já visitada pela partícula até a  $k$ -ésima iteração, e;

$P_{gd}^k$  ( $GBest$ ) a melhor posição já visitada por qualquer uma das partículas até a  $k$ -ésima iteração.

Em algumas situações, a nova posição do elemento  $i$  da partícula ocorre fora do espaço vetorial. Neste caso, o elemento não é deslocado nesta iteração.

Também acontecem casos em que a nova posição do elemento coincide com a posição de outro elemento da mesma partícula. Sendo assim, este elemento também não é deslocado na iteração corrente.

A cada iteração as partículas se movem para a nova posição e a diversidade máxima é calculada. Caso haja melhora na diversidade localmente (maior diversidade máxima já atingida pela partícula),  $PBest$  é atualizado. Igualmente, caso haja melhora em  $GBest$ , ele é atualizado.

Ao final de cada iteração, as partículas são ordenadas em ordem decrescente de acordo com a sua diversidade e um algoritmo de Busca Tabu (TS) é iniciada  $N_r$  Ch T S vezes. A TS tem como soluções de partida as  $N_r$  Ch T S partículas com maior diversidade máxima na iteração corrente.

Ao término da Busca Tabu, as partículas que serviram de solução inicial para a TS são

substituídas pelas partículas retornadas da Busca Tabu e  $GBest$  e  $PBest$  são atualizados.

O algoritmo termina quando atingir um número  $Kmax$  de iterações ou atingir o tempo  $t$  máximo de processamento.

$$Kmax = \text{máximo}(100, Kmax * m) \quad (5)$$

Os valores das constantes  $Kmax$  e  $Nr\_Ch\_TS$  são mostrados na Tabela 6 da Seção 2.4; O último  $GBest$  encontrado é a solução do

O número máximo de iterações do PSO é definido dinamicamente no início da execução do algoritmo através da Equação (5) e depende do número de elementos  $m$  da instância e do valor parâmetro  $Kmax$ .

problema. O pseudocódigo do algoritmo PSO TS é mostrado no Algoritmo 1.

**Algoritmo 1:** Pseudocódigo do Algoritmo PSO\_TS

```

input : Matriz  $M_{n \times n} (d_{ij})$  com cardinalidade  $m \leq n$ 
output: Melhor solução encontrada  $Gb^*$ 
1  $P \leftarrow \{x^1, \dots, x^{\text{max}(1, \text{int}(m \cdot Nr\_PG/100))}\} \leftarrow \text{Gera\_particulas\_gulosas}()$  /* Seção 2.2.1 */;
2  $P \leftarrow \{x^{\text{max}(1, \text{int}(m \cdot Nr\_PG/100))+1}, \dots, x^q\} \leftarrow \text{Gera\_particulas\_aleatorias}()$ ;
3  $Pb^* \leftarrow \{x^1, \dots, x^q\} \leftarrow P = \{x^1, \dots, x^q\}$ ;
4  $Gb^* \leftarrow \text{argmax}\{f(x^i) | i = \{1, \dots, q\}\}$ ;
5  $k \leftarrow 0$ ;
6 while  $k < Kmax$  do
7    $Nr\_Ch = Nr\_Ch\_TS$ ;
8   if  $k = 0$  then
9      $Nr\_Ch = Nr\_Ch\_TS \cdot n \times m / 100$ ;
10  end
11   $Ordena\ P = \{x^1, \dots, x^q\}$  (Descrescente pela diversidade máxima);
12  for  $i \leftarrow 0$  to  $Nr\_Ch$  do
13     $P(s) = TS(P(s))$  /* Seção 2.3 */;
14     $Atualiza\ Gb^*$  e  $Pb^*(s)$ ;
15    if  $Tempo\_processamento > t$  then
16      return  $Gb^*$ ;
17    end
18  end
19  for  $i \leftarrow 0$  to  $q - 1$  do
20     $Move\ P[i]$  /* Eq. (3) */;
21     $Atualiza\ Gb^*$  e  $Pb^*(i)$ ;
22  end
23 end
24 return  $Gb^*$ .

```

## 2.2 ALGORITMOS DE CONSTRUÇÃO

### 2.2.1 HEURÍSTICA DE CONSTRUÇÃO GULOSA

A heurística gulosa cria uma partícula (solução válida) conforme os passos a seguir:

(1) escolhe aleatoriamente um elemento da instância  $N$  para fazer parte da solução  $U$ ;

$$S_i = \sum_{j \in U} d_{ij} \quad (6)$$

(4) adiciona à solução parcial  $U$  o primeiro elemento encontrado com a maior contribuição no vetor

$S$  e exclui este elemento do vetor;

(5) O vetor  $S$  é atualizado conforme a Equação (6), assim como a diversidade máxima da solução parcial, e;

(2) cria um vetor  $S$  contendo os elementos que não fazem ainda parte de  $U$ ;

(3) armazena em  $S$  o valor que cada elemento  $N/U$  contribuiria ao fazer parte da solução parcial, conforme Equação (6);

(6) retorna ao passo (4) até que uma solução válida com  $m$  elementos seja gerada.

A escolha do primeiro elemento sendo aleatória torna o algoritmo não determinístico e permite a geração de mais de uma partícula com elementos diferentes. Durante a criação das partículas é calculada a diversidade máxima de cada uma delas a fim de que seja

escolhido o  $GBest$  (partícula com a maior diversidade máxima entre todas as

posição da partícula encontrada até o momento). O pseudocódigo da Heurística

**Algoritmo 2: Pseudocódigo da Heurística de Construção Gulosa**

```

1 for  $i \leftarrow 0$  to  $\max(1, \text{int}(m * 0.4)) - 1$  do
2    $Diversidade[i] \leftarrow 0$ ;
3    $S \leftarrow$  valor que cada elemento ( $N/M$ ) contribuiria ao fazer parte da
   solução parcial /* Equação (6) */;
4   for  $j \leftarrow 1$  to  $m - 1$  do
5      $P[i][j] \leftarrow$  elemento de  $S$  que apresenta maior contribuição;
6     for  $x \leftarrow 0$  to  $j - 1$  do
7        $Diversidade[i] \leftarrow Diversidade[i] + diversidade[P[i][j], x]$ ;
8     end
9     Atualiza  $S$ ;
10  end
11  Atualiza  $Gb^*$ ;
12  Atualiza  $Pb^*(i)$ ;
13 end
14 return  $P$ .
```

partículas). A própria partícula, na sua criação, é definida como P Best (melhor

Gulosa é mostrado no Algoritmo 2.

## 2.2.2 HEURÍSTICA DE CONSTRUÇÃO ALEATÓRIA

A Heurística de Construção Aleatória escolhe aleatoriamente um elemento, que não faça

parte da solução, após o outro até que seja gerada uma solução válida com  $m$  elementos. O pseudocódigo da Heurística Aleatória é mostrado no Algoritmo 3.

**Algoritmo 3: Pseudocódigo da Heurística de Construção Aleatória**

```

1 for  $i \leftarrow \max(1, \text{int}(m * 0.4))$  to  $q - 1$  do
2    $Diversidade[i] \leftarrow 0$ ;
3    $P[i][0] \leftarrow \text{rand}() \% (n)$ ;
4   for  $j \leftarrow 1$  to  $m - 1$  do
5      $P[i][j] \leftarrow \text{rand}() \% (n/m)$ ;
6     for  $x \leftarrow 0$  to  $j - 1$  do
7        $Diversidade[i] \leftarrow Diversidade[i] + diversidade[P[i][j], x]$ ;
8     end
9   end
10  Atualiza  $Gb^*$ ;
11  Atualiza  $Pb^*(i)$ ;
12 end
13 return  $P$ .
```

## 2.3 BUSCA TABU

A meta-heurística Busca Tabu é um procedimento adaptativo auxiliar que utiliza uma estrutura de memória para armazenar as soluções geradas ou características destas e guiar um algoritmo de busca local na exploração contínua do espaço de busca. A partir de uma solução inicial, move-se de uma solução para uma outra solução vizinha, até que se satisfaça um determinado critério de parada.

A Busca Tabu empregada foi descrita por Wang et al. [2014] e se mostrou eficiente na melhoria das soluções. Consiste em aplicar um operador de troca restrita para trocar uma

variável que tem valor 1 (elemento que faz parte da solução) para outra com valor 0. Formalmente, dada uma solução  $x = \{x_1, \dots, x_n\}$  onde  $U$  e  $Z$  representam respectivamente as variáveis com valor 1 e 0 em  $x$ . Os vizinhos  $N(x)$  são todas as soluções obtidas pela substituição de duas variáveis, onde  $x_i \in U$  e  $x_j \in Z$ . A TS utiliza um mecanismo que determina rapidamente o ganho do movimento de troca, similar ao utilizado por Aringhieri et al. [2008].

Inicialmente, é empregado um vetor  $\Delta$  para armazenar a variação da função objetivo decorrente do movimento de  $x_i$  do subconjunto atual de  $U/Z$  para o outro



subconjunto  $Z/U$ . O vetor é inicializado

conforme a Equação (7) que segue:

$$\Delta_i = \begin{cases} \sum_{j \in U} -d_{ij}(x_i \in U) \\ \sum_{j \in U} d_{ij}(x_i \in Z) \end{cases} \quad (7)$$

O ganho do movimento entre duas variáveis  $x_i \in U$  e  $x_j \in Z$  pode ser calculado usando a seguinte Fórmula (8):

$$\delta_{ij} = \Delta_i + \Delta_j - d_{ij} \quad (8)$$

Após realizado o movimento de troca entre as variáveis  $x_i$  e  $x_j$  o vetor  $\Delta$  deve ser atualizado utilizando a seguinte Equação (9):

$$\Delta_k = \begin{cases} -\Delta_i + d_{ij}(k=i) \\ -\Delta_j + d_{ij}(k=j) \\ \Delta_k + d_{ik} + d_{jk}(k \notin \{i,j\}, x_k \in U) \\ \Delta_k - d_{ik} + d_{jk}(k \notin \{i,j\}, x_k \in Z) \end{cases} \quad (9)$$

O tamanho da vizinhança de troca é igual a  $m(n - m)$  e pode ser computacionalmente custoso para identificar o melhor movimento a cada iteração da TS. Para superar esse obstáculo, foi empregada a estratégia de filtro aplicada a lista de candidatos sucessivo de Glover e Laguna [1997] que restringe a análise para os movimentos que produzem resultados de alta qualidade para cada operação específica.

Para o movimento de troca, é necessário mover a variável  $x_i$  de  $U$  para  $Z$  e mover a variável  $x_j$  de  $Z$  para  $U$ . A diferença na função objetivo resultante de cada operação pode ser facilmente obtida a partir do vetor  $\Delta$ . Para cada operação são escolhidas as principais  $c/s$  variáveis ( $c/s$  é um parâmetro que define tamanho da lista de candidatos) e são gravadas em ordem decrescente em relação aos valores de  $\Delta$  para construir a lista de candidatos  $UCL$  e  $ZCL$ . Finalmente, os movimentos de troca são restritos as variáveis de  $UCL$  e  $ZCL$ . O tamanho da lista tabu foi definido como  $c/s = \min(\sqrt{m}, \sqrt{n - m})$ .

O tamanho da  $c/s$  impacta diretamente no procedimento da TS. Um valor muito grande

pode incluir alguns movimentos pouco atraentes na exploração da vizinhança, enquanto um valor muito baixo poderia excluir alguns movimentos atraentes.

Para garantir que soluções visitadas dentro de um determinado espaço de iterações não sejam revisitadas, a Busca Tabu incorpora uma memória de curto prazo, conhecida como Lista Tabu Glover e Laguna [1997]. Cada vez que duas variáveis  $x_i$  e  $x_j$  são trocadas, dois inteiros aleatórios são gerados a partir de um intervalo de  $tt = [a, b]$  para evitar qualquer movimento envolvendo  $x_i$  ou  $x_j$  ser selecionado por um determinado número de iterações. Os números inteiros que definem a gama de  $tt$ ,  $a = 15$  e  $b = 25$ , são os mesmos empregados por Wang et al. [2014]. Quando  $x_i$  e  $x_j$  são trocados, vamos atribuir a soma de um número inteiro aleatório de  $tt$  e o número atual da iteração  $Iter$  ao elemento  $T[i]$  de  $T$  e a soma de outro número inteiro aleatório de  $tt$  e  $Iter$  para  $T[j]$ . Posteriormente, para qualquer iteração  $Iter$ , uma variável  $x_k$  é proibida de participar de um movimento de troca se  $T[k] > Iter$ . O número de iterações  $\alpha$  foi definido como  $6m$ . O Algoritmo 4 mostra o pseudocódigo da Busca Tabu.

**Algoritmo 4:** Pseudocódigo da Busca Tabu (Baseado em Wang et al. [2014])

```

input : Solução PSO
output: Solução melhorada

1 Iter ← 0;
2 SemMelhora ← 0;
3  $T[x_i, \dots, x_n] \leftarrow 0$  (Lista Tabu);
4 Criar e inicializar o vetor  $\Delta$  de acordo com a Eq. (7);
5 while SemMelhora <  $\alpha$  do
6   Construir a lista de candidatos UCL e ZCL, gravadas em ordem
   decrecente em relação aos valores de  $\Delta$ ;
7   Executa o movimento de troca de  $x_i$  para  $x_j$  que proporcione a melhor
   troca de acordo com a equação  $\delta_{ij} = \Delta_i + \Delta_j - d_{ij}$  e desde que
    $T[x_i] \leq \textit{Iter}$  e  $T[x_j] \leq \textit{Iter}$ ;
8   Atualiza o vetor  $\Delta$  de acordo com a Eq. (9);
9   Atualiza  $T[x_i]$  e  $T[x_j]$ ;
10  Iter ++;
11  if Melhorou Solução then
12    | SemMelhora ← 0;
13  else
14    | SemMelhora ++;
15  end
16 end
17 return Solução melhorada.

```

## 2.4 PARAMETROS

Para a execução do algoritmo, algumas constantes devem ser previamente definidas. Foram testados valores com  $q$

= .15, .20, .25, .5 e 1,  $Nr_{PG} = 1\%, 10\%, 40\%, 70\%, 80\%$  e  $90\%$ ,  $c1 = .8, 1$  e  $1.2$ ,  $c2 = 1.2, 1.4$  e  $1.8$ ,  $Nr_{Ch_{TS}} = 4, 5, 6, 7, 8$  e  $9$  e  $Nr_{Ch_{TS}In} = 0\%, 20\%, 40\%$  e  $60\%$ .

Tabela 6: Valores definidos para as constantes

| Instances | $q$ | $Nr_{PG}$ | $Kmax$ | $c1$ | $c2$ | $Nr_{Ch_{TS}}$ | $Nr_{Ch_{TS}In}$ |
|-----------|-----|-----------|--------|------|------|----------------|------------------|
| SOM-b     | 1   | 80 %      | 1      | 1.00 | 1.40 | 7              | 0                |
| MDG-a     | .20 | 80 %      | 1      | 1.00 | 1.40 | 8              | 0                |
| MDG-c     | .20 | 60 %      | 4      | 1.00 | 1.40 | 6              | 20               |

Os parâmetros foram escolhidos de forma empírica, com base nos resultados dos testes realizados, analisando as melhores soluções e soluções médias encontradas. A Tabela 6. mostra os valores para as constantes, onde  $m$  é o número de elementos da solução,  $q \times m =$  número de partículas iniciais,  $Nr_{PG} =$

porcentagem de partículas criadas de forma gulosa,  $Kmax \times m =$  número máximo de iterações do *PSO*,  $c1$  and  $c2 =$  constantes de aceleração positivas,  $Nr_{Ch_{TS}} =$  número de chamadas a *TS* a cada iteração do *PSO* e  $Nr_{Ch_{TS}In} =$  porcentagem de partículas iniciais que são submetidas a Busca Tabu.

## 3 RESULTADOS

Para verificar a eficiência do algoritmo proposto foram empregadas as instâncias de médio porte SOM-b, criadas por Silva et al. [2004], que contém uma população de tamanho  $n = 100, n = 200, n = 300, n = 400$  e  $n = 500$ , e testadas soluções de tamanho  $m = 10\%n, 20\%n, 30\%n$  e  $40\%n$ . A meta-heurística *PSO TS* foi executada 10 vezes pelo tempo máximo de 1s. As melhores soluções conhecidas da literatura (BKS) foram retiradas

de Duarte et al. [2008]. Para as instâncias de grande porte foram usadas 20 instâncias MDG-a criadas por Duarte e Martí [2007] e foram executadas 30 vezes por um tempo máximo de 16s e 20 instâncias MDG-c, similares as usadas por Palubeckis [2007], que foram executadas por 10 vezes por um tempo máximo de 480s. Para a diversidade máxima, foram considerados o resultado com maior diversidade (#Best) e a diversidade média (#Avg).

Tabela 7: Resultados das instâncias SOM-b (Tempo &lt; 1s)

| Instâncias | n   | m   | BKS   | PSO TS           |      |      |      |
|------------|-----|-----|-------|------------------|------|------|------|
|            |     |     |       | Tempo (segundos) |      |      |      |
|            |     |     |       | Média            | Mín  | Méd  | Máx  |
| SOM-b 1    | 100 | 10  | 333   | 333              | 0.00 | 0.00 | 0.00 |
| SOM-b 2    | 100 | 20  | 1195  | 1195             | 0.00 | 0.00 | 0.00 |
| SOM-b 3    | 100 | 30  | 2457  | 2457             | 0.00 | 0.00 | 0.00 |
| SOM-b 4    | 100 | 40  | 4142  | 4142             | 0.00 | 0.00 | 0.01 |
| SOM-b 5    | 200 | 20  | 1247  | 1247             | 0.00 | 0.02 | 0.08 |
| SOM-b 6    | 200 | 40  | 4450  | 4450             | 0.00 | 0.02 | 0.05 |
| SOM-b 7    | 200 | 60  | 9437  | 9437             | 0.00 | 0.01 | 0.01 |
| SOM-b 8    | 200 | 80  | 16225 | 16225            | 0.01 | 0.01 | 0.02 |
| SOM-b 9    | 300 | 30  | 2694  | 2694             | 0.00 | 0.04 | 0.12 |
| SOM-b 10   | 300 | 60  | 9689  | 9689             | 0.01 | 0.10 | 0.29 |
| SOM-b 11   | 300 | 90  | 20743 | 20743            | 0.01 | 0.06 | 0.19 |
| SOM-b 12   | 300 | 120 | 35881 | 35881            | 0.04 | 0.09 | 0.15 |
| SOM-b 13   | 400 | 40  | 4658  | 4658             | 0.02 | 0.09 | 0.15 |
| SOM-b 14   | 400 | 80  | 16956 | 16956            | 0.02 | 0.14 | 0.37 |
| SOM-b 15   | 400 | 120 | 36317 | 36317            | 0.05 | 0.10 | 0.23 |
| SOM-b 16   | 400 | 160 | 62487 | 62484            | 0.09 | 0.23 | 0.31 |
| SOM-b 17   | 500 | 50  | 7141  | 7141             | 0.04 | 0.26 | 0.52 |
| SOM-b 18   | 500 | 100 | 26258 | 26258            | 0.03 | 0.12 | 0.31 |
| SOM-b 19   | 500 | 150 | 56572 | 56572            | 0.06 | 0.08 | 0.16 |
| SOM-b 20   | 500 | 200 | 97344 | 97344            | 0.11 | 0.26 | 0.52 |
| Média      |     |     |       |                  | 0.02 | 0.08 | 0.17 |

O algoritmo foi implementado utilizando a linguagem de programação C++, compilador gcc compilado com a opção -O2 e SO Linux Ubuntu V. 4.2.0-35. Para os testes, foi utilizado um computador com processador Intel i7 950 @ 3.07GHz e 4 cores, 8 GB Memória e HD 2 TB. Apenas 1 core do processador foi utilizado.

A Tabela 7 mostra os resultados médios obtidos para as instâncias de médio porte SOM-b. Nesta tabela, a coluna 1 indica a instância utilizada; a coluna 2, a dimensão n desta instância; a coluna 3, a quantidade m de elementos da partícula (solução); a coluna 4, a melhor solução (diversidade máxima) conhecida na literatura (BKS); a coluna 5, a solução média encontrada nos 10 testes; as colunas 6 e 7 e 8, respectivamente, o tempo mínimo, médio e máximo de processamento até encontrar a solução.

Conforme podemos observar, nas instâncias de médio porte, em todas as instâncias e em

100% dos 10 testes executados em cada uma das delas, o algoritmo PSO TS encontrou a melhor solução conhecida literatura, ou seja, um desvio padrão nulo. O tempo máximo para encontrar a melhor solução variou de 0 a 0.52 segundos, dependendo do valor de n e m da instância e o tempo médio para encontrar as soluções foi de 0,08 segundos.

Comparamos os resultados com algoritmos state-of-the-art que utilizaram as instancias SOM-b para os testes. Várias Meta-heurística foram analisadas no trabalho de Martí et al. [2013], que realizaram os testes por um tempo de processamento de 10s e 20s em um processador Intel Core2 Quad Q8300 2.5GHz. Não foi encontrado na literatura testes com tempo de processamento inferior a 10 segundos. O PSO TS sempre encontrou resultados melhores ou iguais, mesmo sendo executado com um limite de tempo de apenas 1 segundo.

Tabela 8: Resultados das instâncias MDG-a (Tempo  $\leq 16$ s)

| Instância | BKS    | PSO_TS  |          |       |       | CV      |
|-----------|--------|---------|----------|-------|-------|---------|
|           |        | Solução |          | GAP   |       |         |
|           |        | #Best   | #Avg.    | #Best | #Avg. |         |
| MDG-a_21  | 114271 | 114271  | 114237.6 | 0     | 33    | 0.00018 |
| MDG-a_22  | 114327 | 114327  | 114298.8 | 0     | 28    | 0.00047 |
| MDG-a_23  | 114195 | 114195  | 114157.4 | 0     | 38    | 0.00029 |
| MDG-a_24  | 114093 | 114091  | 114042.7 | 2     | 50    | 0.00022 |
| MDG-a_25  | 114196 | 114196  | 114108.1 | 0     | 88    | 0.00035 |
| MDG-a_26  | 114265 | 114265  | 114232.0 | 0     | 33    | 0.00016 |
| MDG-a_27  | 114361 | 114361  | 114351.9 | 0     | 9     | 0.00016 |
| MDG-a_28  | 114327 | 114327  | 114280.0 | 0     | 47    | 0.00054 |
| MDG-a_29  | 114199 | 114199  | 114169.6 | 0     | 29    | 0.00018 |
| MDG-a_30  | 114229 | 114229  | 114193.6 | 0     | 35    | 0.00024 |
| MDG-a_31  | 114214 | 114214  | 114160.4 | 0     | 54    | 0.00025 |
| MDG-a_32  | 114214 | 114214  | 114151.8 | 0     | 62    | 0.00034 |
| MDG-a_33  | 114233 | 114233  | 114189.9 | 0     | 43    | 0.00028 |
| MDG-a_34  | 114216 | 114216  | 114181.9 | 0     | 34    | 0.00033 |
| MDG-a_35  | 114240 | 114240  | 114229.1 | 0     | 11    | 0.00013 |
| MDG-a_36  | 114335 | 114335  | 114313.3 | 0     | 22    | 0.00025 |
| MDG-a_37  | 114255 | 114255  | 114218.5 | 0     | 37    | 0.00029 |
| MDG-a_38  | 114408 | 114408  | 114398.6 | 0     | 9     | 0.00015 |
| MDG-a_39  | 114201 | 114201  | 114183.7 | 0     | 17    | 0.00031 |
| MDG-a_40  | 114349 | 114349  | 114297.5 | 0     | 52    | 0.00038 |
| Média     |        |         |          | 0.1   | 36.6  | 0.00027 |

Na Tabela 8 nós temos os resultados encontrados para as instâncias grandes MDG-a, onde as colunas indicam, respectivamente, a instância usada; a melhor solução conhecida na literatura; a melhor solução encontrada nos testes; o GAP (BKS-#Best); o GAP (BKS-#Avg.) e o coeficiente de variação, obtido pela razão entre desvio padrão e a média entre as 30 execuções.

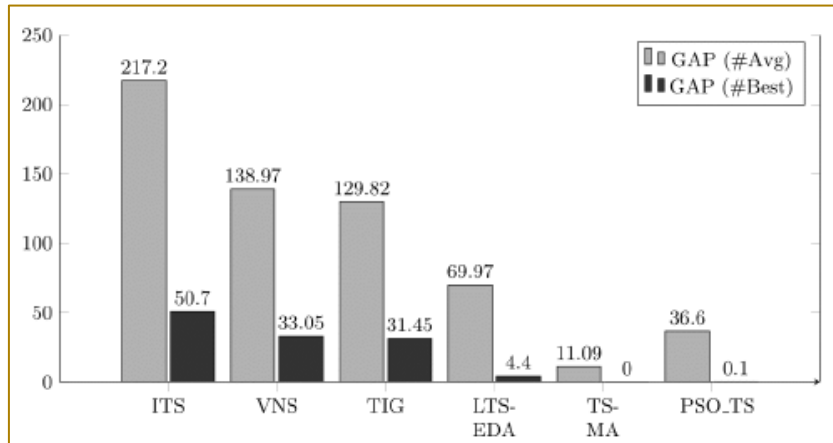
Nós podemos observar que em apenas uma instância o algoritmo não conseguiu encontrar a melhor solução conhecida na literatura, sendo que o GAP foi de apenas 2 unidades e o GAP médio foi de 36.6 unidades. O coeficiente de variação médio foi de 0.00027, o que demonstra a robustez do algoritmo.

Os resultados obtidos nas instâncias MDG-a foram comparados com algoritmos state-of-the-art, que foram relatados em Wang et al. [2014] e foram executados por um tempo de 17 segundos em um computador Xeon E5440 com 2.83 GHz CPU e 8 GB RAM ou máquinas equivalentes.

Para determinar o tempo de parada do nosso algoritmo, nós usamos o CPU Benchmarks

Single Thread Performance 1 que indicou que nosso computador tem um fator 1.036 mais rápido que aqueles usados nos testes por Wang et al. [2014]. Então, nós determinamos o tempo de parada como 16 segundos para as instâncias MDG-a.

Figura 1: Comparação entre PSO TS e outros algoritmos state-of-the-art (instâncias MDG-a)



O gráfico da Figura 1 mostra o GAP médio da melhor solução encontrada pelo algoritmo PSO TS em cada uma das 20 instâncias MDG-a (média da coluna #Best da Tabela 8), além do GAP médio da média das soluções encontradas pelo algoritmo PSO TS nas execuções realizadas em cada uma das 20 instâncias MDG-a (média da coluna #Avg da Tabela 8). Então, é feita uma comparação com os resultados de diferentes meta-heurísticas encontradas na literatura.

Podemos observar que obtemos resultados melhores que os algoritmos ITS Palubeckis [2007], VNS Brimberg et al. [2009], TIG Lozano et al. [2011] e LTS EDA Wang et al. [2012]. Na comparação com o algoritmo TS-MA Wang et al. [2014] obtivemos resultados ligeiramente inferiores em relação à qualidade das soluções. Acreditamos que isso se deva às características da meta-heurística PSO, que requer um tempo maior de processamento para convergir para as soluções ótimas.

Para a análise das instâncias MDG-c, também consideradas instâncias de grandes porte, foram realizadas 10 execuções e utilizado um tempo máximo de processamento 480 segundos.

Na Tabela 9 nós temos os resultados encontrados para estas instâncias, onde as colunas indicam, respectivamente, a instância usada; a melhor solução conhecida na literatura; a melhor solução encontrada nos testes; a solução média encontrada; o GAP da melhor solução e o GAP da solução média em relação a melhor solução conhecida.

O GAP foi calculado de acordo com a Equação  $GAP = \left( \frac{\#Best}{\#Soluc} - 1 \right) \times 100$ , onde

#Best é a melhor solução encontrada na literatura e Soluc é solução analisada. GAPs negativos representam melhores resultados. As melhores soluções conhecidas na literatura foram retiradas de Martí et al. [2013].

Como podemos ver, o PSO TS encontrou a melhor solução conhecida em 04 instâncias e, principalmente, melhorou os resultados conhecidos para seis casos. Em dois casos, até mesmo a sua qualidade média foi melhor do que os resultados conhecidos anteriormente relatados na literatura.

Os resultados obtidos nas instâncias MDG-c foram comparados com algoritmos state-of-the-art, que foram relatados em Martí et al. [2013] e que foram executados por um tempo de processamento de 600 segundos em um computador Intel Core 2 Quad CPU Q 8300 with 6 GB de RAM.

Para determinar o tempo de parada do nosso algoritmo, nós usamos o CPU Benchmarks

Single Thread Performance que indicou que nosso computador tem um fator 1.246 mais rápido que aqueles usados nos testes por Martí et al. [2013]. Então, nós determinamos o tempo de parada como 480 segundos para as instâncias MDG-c.

O gráfico da Figura 2 mostra a comparação entre o GAP da melhor solução conhecida na literatura e a média das soluções encontrada no PSO TS com os resultados das diferentes metaheurísticas. Nós podemos observar que o algoritmo PSO TS obteve melhores resultados que todos os algoritmos comparados: GRASP D2+I LS Duarte e Martí [2007], ITS Palubeckis [2007], B VNS Brimberg et al. [2009], A VNS Aringhieri e Cordone [2011] e G SS Gallego et al. [2009].

Tabela 9: Resultado das instâncias MDG-c (Tempo  $\leq$  480s)

| Instances | BKS      | PSO_TS   |          |         |         |
|-----------|----------|----------|----------|---------|---------|
|           |          | Solution |          | GAP %   |         |
|           |          | Best     | Avg.     | Best    | Avg.    |
| MDG-c_1   | 24924685 | 24926344 | 24918989 | -0.0067 | 0.0229  |
| MDG-c_2   | 24909199 | 24912546 | 24903365 | -0.0134 | 0.0234  |
| MDG-c_3   | 24900820 | 24896090 | 24890210 | 0.0190  | 0.0426  |
| MDG-c_4   | 24904964 | 24909710 | 24900660 | -0.0191 | 0.0173  |
| MDG-c_5   | 24899703 | 24895704 | 24890870 | 0.0161  | 0.0355  |
| MDG-c_6   | 43465087 | 43444251 | 43436138 | 0.0480  | 0.0666  |
| MDG-c_7   | 43477267 | 43477267 | 43466250 | 0.0000  | 0.0253  |
| MDG-c_8   | 43458007 | 43462452 | 43458214 | -0.0102 | -0.0005 |
| MDG-c_9   | 43448137 | 43448137 | 43444905 | 0.0000  | 0.0074  |
| MDG-c_10  | 43476251 | 43465561 | 43460248 | 0.0246  | 0.0368  |
| MDG-c_11  | 67009114 | 67021132 | 67015085 | -0.0179 | -0.0089 |
| MDG-c_12  | 67021888 | 67014170 | 66999552 | 0.0115  | 0.0333  |
| MDG-c_13  | 67024373 | 67016057 | 67011065 | 0.0124  | 0.0199  |
| MDG-c_14  | 67024804 | 67027878 | 67022870 | -0.0046 | 0.0029  |
| MDG-c_15  | 67056334 | 67056334 | 67050797 | 0.0000  | 0.0083  |
| MDG-c_16  | 95637733 | 95637733 | 95633768 | 0.0000  | 0.0041  |
| MDG-c_17  | 95645826 | 95569412 | 95556782 | 0.0800  | 0.0932  |
| MDG-c_18  | 95629207 | 95532141 | 95518634 | 0.1016  | 0.1158  |
| MDG-c_19  | 95633549 | 95594864 | 95590270 | 0.0405  | 0.0453  |
| MDG-c_20  | 95643586 | 95580853 | 95573081 | 0.0656  | 0.0738  |
| Average   |          |          |          | 0.017   | 0.033   |

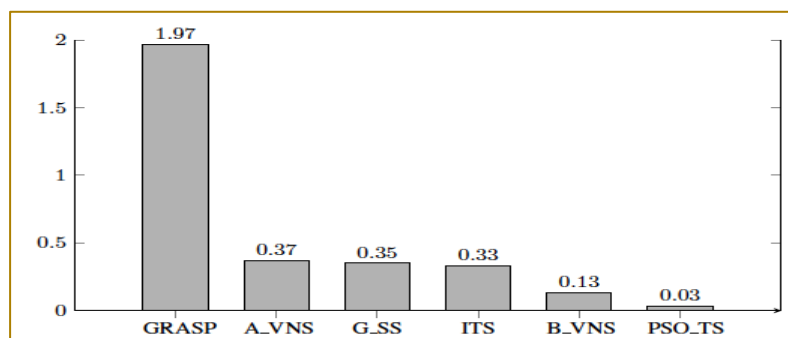
#### 4 CONCLUSÕES

Neste trabalho foi proposta uma nova abordagem para a resolução do Problema da Diversidade Máxima baseada na meta-heurística Otimização por Nuvem de Partículas juntamente com Busca Tabu. Para a geração das partículas iniciais foram implementadas uma heurística de construção gulosa e outra aleatória.

O método, denominado PSO TS, mostra-se simples na sua implementação e os

resultados demonstram que o algoritmo consegue excelente desempenho, obtendo, em todas as instâncias de médio porte testadas, as melhores soluções encontradas na literatura. Nas instâncias de grande porte MDG-a apenas em uma instância este resultado não foi alcançado. Nas instâncias MDG-c, o algoritmo melhorou a solução conhecida em seis casos e encontrou a melhor solução em outros quatro casos.

Figura 2: Comparação das instâncias MDG-c entre PSO TS e outros algoritmos state-of-the-art



O PSO TS apresentou-se bastante robusto em relação ao comportamento médio da qualidade das soluções obtidas, com um coeficiente de variação nulo para as instâncias de médio porte SOM-b. Nas instâncias de grande porte MDG-a, o coeficiente de variação foi de 0,00027 e de apenas 0,00012 para as instâncias MDG-c. Em resumo, uma variabilidade muito baixa.

Como passos futuros, visando melhorar a qualidade média das soluções e,

principalmente, reduzir o tempo de processamento, têm-se como principais sugestões o estudo do ajuste de parâmetros usados no algoritmo buscando melhorar o desempenho, explorar novas estruturas de vizinhança para as partículas, onde o PBest é escolhido a partir do melhor vizinho e desenvolver abordagens paralelas para o PSO TS para resolver as instâncias grandes.

## REFERÊNCIAS

- [1]. Agrafiotis, D. (1997). Stochastic algorithms for maximizing molecular diversity. *Journal of Chemical Information and Modeling*, 37:841–851.
- [2]. Aringhieri, R. e Cordone, R. (2011). Comparing local search metaheuristics for the maximum diversity problem. *Journal of the Operational Research Society*, 62:266–280.
- [3]. Aringhieri, R., Cordone, R., e Melzani, Y. (2008). Tabu search versus GRASP for the maximum diversity problem. *4OR*, 6(1):45–60. ISSN 16194500.
- [4]. Brimberg, J., Mladenovic, N., Uros̃evic, D., e Ngai, E. (2009). Variable neighborhood search for the heaviest k-subgraph. *Computers and Operations Research*, 36(11):2885–2891.
- [5]. Chuang, L.-Y., Hsiao, C.-J., e Yang, C.-H. (2011). Chaotic particle swarm optimization for data clustering. *Expert Systems with Applications*, 38(12):14555–14563. ISSN 09574174.
- [6]. Coelho Filho, O. P. (2013). Uma nova abordagem híbrida do algoritmo de otimização por enxame de partículas com busca local iterada para o problema de clusterização de dados. *Journal of Chemical Information and Modeling*, 53(9):1689–1699. ISSN 1098-6596.
- [7]. Duarte, A. e Martí, R. (2007). Tabu search and GRASP for the maximum diversity problem. *European Journal of Operational Research*, 178(1):71–84. ISSN 03772217.
- [8]. Duarte, I. L., da Silva, G. C., e Costa, T. A. (2008). Algoritmos heurísticos para o problema da diversidade máxima. *Anais do XL SBPO*. João Pessoa, PB, p. 1126–1137.
- [9]. Gallego, M., Duarte, A., Laguna, M., e Martí, R. (2009). Hybrid heuristics for the maximum diversity problem. *Computational Optimization and Applications*, 44(3):411–426. ISSN 09266003.
- [10]. Ghosh, J. B. (1996). Computational aspects of the maximum diversity problem. *Operations Research Letters*, 19(4):175–181. ISSN 01676377.
- [11]. Glover, F. e Laguna, M. (1997). *Tabu search*. Kluwer Academic Publishers.
- [12]. Kuo, C.-C., Glover, F., e Dhir, K. S. (1993). Analyzing and Modeling the Maximum Diversity Problem by Zero-One Programming. *Decision Sciences*, 24(6):1171–1185. ISSN 0011-7315.
- [13]. Lozano, M., Molina, D., e García-Martínez, C. (2011). Iterated greedy for the maximum diversity problem. *European Journal of Operational Research*, 214(1):31–38. ISSN 03772217.
- [14]. Martí, R., Gallego, M., e Duarte, A. (2010). A branch and bound algorithm for the maximum diversity problem. *European Journal of Operational Research*, 200(1):36–44. ISSN 03772217.
- [15]. Martí, R., Gallego, M., Duarte, A., e Pardo, E. G. (2013). Heuristics and metaheuristics for the maximum diversity problem. *Journal of Heuristics*, 19(4):591–615. ISSN 13811231.
- [16]. Palubeckis, G. (2007). Iterated tabu search for the maximum diversity problem. *Applied Mathematics and Computation*, 189(1):371–383. ISSN 00963003.
- [17]. Silva, G., Ochi, L., e Martins, S. (2004). Experimental comparison of greedy randomized adaptive search procedures for the maximum diversity problem. *Lecture Notes on Computer Science*, p. 498–512.

- [18]. Silva, G. C., De Andrade, M. R. Q., Ochi, L. S., Martins, S. L., e Plastino, A. (2007). New heuristics for the maximum diversity problem. *Journal of Heuristics*, 13(4):315–336. ISSN 13811231.
- [19]. Takane, B. (2011). Um algoritmo exato para o problema da diversidade máxima. Dissertação de m.sc., UFMG, Belo Horizonte, MG, Brasil.
- [20]. Wang, J., Zhou, Y., Cai, Y., e Yin, J. (2012). Learnable tabu search guided by estimation of distribution for maximum diversity problems. *Soft Computing*, 16(4):711–728. ISSN 14327643.
- [21]. Wang, Y., Hao, J.-K., Glover, F., e Lu", Z. (2014). A tabu search based memetic algorithm for the maximum diversity problem. *Engineering Applications of Artificial Intelligence*, 27:103–114.
- [22]. Wu, Q. e Hao, J. K. (2013). A hybrid metaheuristic method for the Maximum Diversity Problem.
- [23]. *European Journal of Operational Research*, 231(2):452–464. ISSN 03772217.



# Capítulo 6

## AUXÍLIO MULTICRITÉRIO NA PRIORIZAÇÃO DE RISCOS DE PROJETOS

*Ricardo Luiz Fernandes Bella*

**Resumo:** O objetivo deste artigo é apresentar uma aplicação multicritério em uma questão central na análise de riscos: a priorização. Para isso, utilizou-se como metodologia a modelagem e simulação da aplicação combinada de 2 ferramentas multicritério nos dados de um projeto real. As ferramentas multicritério utilizadas foram: *Elimination Et Choix Traduisant la Réalité* (ELECTRE-Tri) para a classificação; e, *Analytic Hierarchy Process* (AHP) para a priorização. Nesse estudo, o resultado da aplicação das ferramentas multicritério foi comparado com o resultado real do projeto obtido pela aplicação da metodologia *Project Management Institute* (PMI). Como conclusão, observa-se que a abordagem multicritério teve um resultado similar ao PMI. Assim, sugere-se que a priorização de riscos do projeto, que é concebida sob uma lógica qualitativa pelo PMI, possa ser aparamentada de métodos multicritérios que apoiem a tomada de decisão. Nesse sentido, o artigo traz, como principal contribuição, uma proposta de modelo para a priorização de riscos de projetos.

**Palavras Chave:** Multicritério, Riscos, Projetos. Tópico (ADM – Apoio à Decisão Multicritério)

## 1. INTRODUÇÃO

O gerenciamento de riscos em projetos tem por objetivo maximizar ganhos e diminuir perdas através de processos que identificam, analisam e monitoram os riscos (PMI, 2013). Junior & Carvalho (2013) ressaltam a importância desse tópico para a gestão de projetos relacionando o gerenciamento de riscos com o sucesso do projeto.

Nesse sentido, propõe-se o tema “auxílio multicritério na priorização de riscos de projetos” com o objetivo de ilustrar a aplicação de métodos quantitativos que possam robustecer o processo de análise de riscos de projetos. Para isso, o artigo faz um recorte sobre o gerenciamento de riscos de projetos delimitando a pesquisa sobre dois processos de análise de riscos: a classificação e priorização de riscos.

Os métodos quantitativos utilizados foram as ferramentas multicritério: *Elimination Et Choix Traduisant la Réalité* (ELECTRE-Tri) para a classificação de riscos e *Analytic Hierarchy Process* (AHP) para a priorização dos

mesmos. Além disso, o estudo lançou mão de uma abordagem tradicional pautada nas boas práticas de gerenciamento de projetos do *Project Management Institute* (PMI) como instrumento de comparação.

Sendo assim, para atingir o objetivo do artigo, o conteúdo foi organizado nas seguintes seções: (1) introdução; (2) referencial teórico; (3) metodologia; (4) modelagem; (5) resultado; e (6) conclusão.

## 2. REFERENCIAL TEÓRICO

### 2.1 GESTÃO DE RISCOS DE PROJETOS

O gerenciamento de risco de projetos envolve processos que ocorrem durante todas as fases de um projeto. Segundo o PMI (2013), estes processos de gerenciamento de riscos têm o objetivo de controlar ameaças e fomentar oportunidades. Na tabela 1, os processos que compõem o gerenciamento de riscos de projetos segundo as boas práticas do PMI e seus respectivos objetivos podem ser observados.

Tabela 1: Processos de gerenciamento de riscos de projetos segundo o PMI

| Processo                           | Objetivo                                     |
|------------------------------------|--|
| Planejar o Gerenciamento de Riscos | Definir a condução do gerenciamento de risco |
| Identificar os Riscos              | Registrar riscos e suas informações          |
| Realizar a Análise Qualitativa     | Avaliar e classificar os riscos              |
| Realizar a Análise Quantitativa    | Quantificar os impactos prováveis            |
| Planejar as Respostas aos Riscos   | Desenvolver opções e ações de resposta       |
| Controlar os Riscos                | Monitorar riscos e implementar planos        |

Fonte: PMI (2013)

Conforme a tabela 1, a abordagem tradicional segmenta o processo de análise de riscos em dois grupos: a análise qualitativa orientada para a classificação dos riscos; e a análise quantitativa que é voltada para os impactos financeiros dos riscos.

Desse modo, a ênfase do estudo recai sobre a análise qualitativa que possui como seu principal objetivo a classificação de riscos. Na tabela 2, o processo de análise qualitativa de riscos pode ser mais bem entendido a partir de três conjuntos: entradas, ferramentas e saídas.

Tabela 2: Processo de análise qualitativa de riscos segundo o PMI

| Entradas                            | Ferramentas                                     | Saídas                                |
|-------------------------------------|---|---------------------------------------|
| Plano de Gerenciamento de Risco     | Avaliação de Probabilidade e Impacto dos Riscos | Atualização nos Documentos do Projeto |
| Linha de Base do Escopo             | Matriz de Probabilidade e Impacto               |                                       |
| Registro dos Riscos                 | Avaliação de Qualidade dos Dados sobre Risco    |                                       |
| Fatores Ambientais da Empresa       | Categorização de Riscos                         |                                       |
| Ativos de Processos Organizacionais | Avaliação da Urgência dos Riscos                |                                       |
|                                     | Opinião Especializada                           |                                       |

Fonte: PMI (2013)

Conforme se percebe na tabela 2, a maioria dos elementos que compõe esse processo de análise (isto é, as entradas, ferramentas e saídas) tem caráter qualitativo. Dessa maneira, ao se realizar a análise qualitativa dos riscos, as principais informações geradas são: a classificação dos riscos segundo as variáveis de probabilidade e impacto; e a priorização dos riscos segundo o grau de urgência, isto é, a ordem de prioridade para o tratamento dos riscos.

## 2.2 ANÁLISE MULTICRITÉRIO

O campo de estudo de análise multicritério ou auxílio multicritério é abordado pela literatura internacional como *Multicriteria Decision Aid* (MCDA) e está baseado na aplicação de métodos multicritério na solução de questões tais como: ranqueamento, classificação e

seleção.

Nesse contexto, existem famílias e escolas de métodos multicritérios que agrupam os métodos segundo as suas similaridades (COSTA et al, 14), sendo as principais a escola americana e a francesa (SILVA et al, 07). Nesse artigo foram utilizadas as ferramentas mais populares de cada uma dessas escolas: a AHP da escola americana e o ELECTRE da escola francesa.

## 2.3 ELECTRE-TRI

Na descrição de Gomes (2007) os métodos ELECTRE incluem seis versões, sendo todas refinamentos e adaptações do primeiro método multicritério proposto por Bernard Roy. Na tabela 3, pode-se observar as versões do método, o ano de desenvolvimento e a sua aplicação.

Tabela 3: Origem E Versões Do Método Electre

| Versão | Autor           | Ano  | Aplicação     |
|--------|-----------------|------|---------------|
| I      | Roy             | 1968 | Seleção       |
| II     | Roy e bertier   | 1973 | Ordenação     |
| III    | Roy             | 1978 | Ordenação     |
| IV     | Roy e hugonnard | 1982 | Ordenação     |
| IS     | Roy e skalka    | 1985 | Seleção       |
| Tri    | Yu wei          | 1992 | Classificação |

Fonte: Gomes (2007)

Conforme a tabela 3 nota-se que a versão de maior aderência ao objetivo de classificação

de riscos é a versão Tri do ELECTRE proposta por Yu Wei. Esse método foi projetado para

classificar alternativas avaliadas segundo vários critérios (FREITAS et al, 1998). Além disso, o ELECTRE-Tri é um dos mais populares métodos na literatura de classificação (BOUYSSOU et al, 2014), por isso foi escolhido para auxiliar na classificação dos riscos.

## 2.4 AHP

O método AHP é um dos métodos multicritério mais populares na literatura de MCDA. Isso pode ser verificado em uma busca rápida por artigos na base Scopus que contenham "AH" em seus títulos, resumos ou palavras-chave. Em abril de 2015, o resultado dessa busca foi de 9.161 artigos publicados em jornais. Como régua de comparação, o mesmo método de busca para a palavra "ELECT E" retornou 549 artigos. Por essa razão, escolheu-se utilizar o método AHP.

Entretanto, o AHP sofre uma crítica forte relacionada a sua aplicação em problemas de ordenação chamada de "reversão de". A reversão de ordem consiste na mudança da ordem das alternativas avaliadas pelo método caso este seja reutilizado após uma redução no conjunto de alternativas, isto é, excluir do modelo uma ou mais alternativas pode gerar uma mudança drástica no resultado (GAUDARD et al, 2012).

No caso desse artigo, o problema de reversão de ordem foi mitigado através de uma modelagem orientada pela classificação prévia dos riscos, isso significa que as alternativas foram avaliadas separadamente segundo a sua classe de risco. Desse modo, mesmo que ocorra uma reversão de ordem, o risco continuará na sua faixa de classificação.

## 3. METODOLOGIA

Este estudo foi realizado com informações de um projeto do programa social de uma empresa de grande porte. Essa empresa visualizou a oportunidade de capacitar pessoas com potencial para se tornarem atletas de ponta através de projetos desportivos com ênfase nos esportes olímpicos. Para isso, esse programa buscou a integração de outras empresas como patrocinadoras para alavancar os projetos, sendo estes projetos de longo prazo (início em 2013 e término em 2020).

A escolha da empresa estudada foi por critério de conveniência dado que um dos

pesquisadores trabalha na consultoria que apoiou as equipes de projeto deste programa. A escolha do projeto também foi por critério de conveniência dado que este foi escolhido devido o seu maior grau de andamento. O projeto selecionado foi o projeto do esporte boxe, denominado projeto boxe.

Para preservar a empresa e seu programa social, o nome da empresa não será revelado e os dados utilizados neste trabalho foram adaptados, entretanto buscou-se manter a correspondência semântica e a ordem lógica dos dados.

A pesquisa delimitou-se em três etapas metodológicas básicas: (1) modelagem segundo as duas abordagens a serem comparadas, i.e, a tradicional (PMI) e a multicritério (ELECTRE-Tri / AHP);

(2) simulação dos modelos com base nos dados do projeto boxe; e (3) apresentação e comparação dos resultados.

### 3.1 ABORDAGEM TRADICIONAL (PMI)

A abordagem tradicional, conforme sugere o PMI, geralmente segue 6 etapas:

- - identificação dos riscos do projeto;
- - avaliação da qualidade dos dados sobre os riscos;
- - predefinição das escalas de avaliação de probabilidade e impacto;
- - avaliação de probabilidade e impacto dos riscos;
- - classificação dos riscos segundo a matriz de probabilidade e impacto;
- - priorização dos riscos segundo a avaliação de urgência dos mesmos.

### 3.2 ABORDAGEM MULTICRITÉRIO (ELECTRE-TRI / AHP)

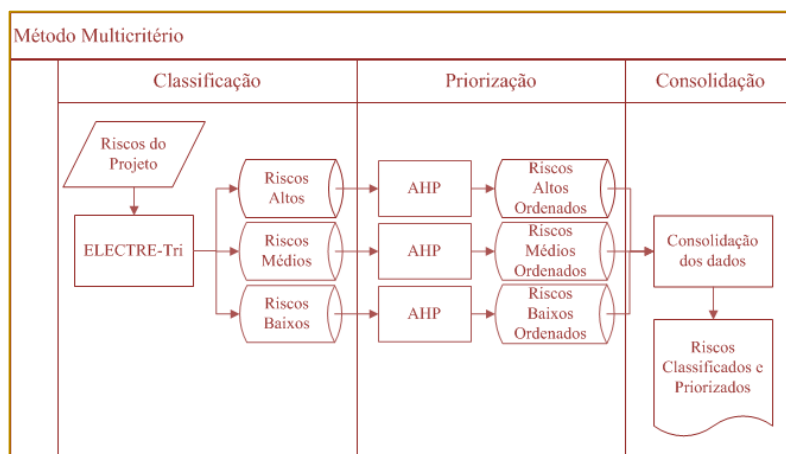
A abordagem multicritério foi modelada segundo os mesmos passos da abordagem tradicional, se diferenciando apenas pela aplicação de ferramentas multicritério aos dois últimos passos da abordagem tradicional, isto é, classificar e priorizar os riscos.

Sendo assim, ressaltam-se os passos necessários para a aplicação das ferramentas multicritério. Na figura 2, observa-se a aplicação das ferramentas multicritério em 3

passos: (1) classificação dos riscos do projeto com o auxílio da ferramenta ELECTRE-Tri; (2) priorização dos riscos previamente

classificados com o auxílio da ferramenta AHP; e (3) consolidação dos riscos classificados e priorizados.

Figura 2: Fluxograma do método multicritério para a classificação e priorização dos riscos



Fonte: Metodologia

Conforme ilustrado na figura 2, o processo de priorização de riscos proposto se dá pela combinação das duas ferramentas multicritério escolhidas. Sendo assim, é necessário destacar a modelagem de cada uma dessas ferramentas.

#### 4. MODELAGEM

##### 4.1 MODELAGEM DA ABORDAGEM TRADICIONAL (PMI)

A utilização do método tradicional requer as seguintes definições (PMI, 2013):

Os riscos do projeto;

A escala de avaliação de probabilidade;

A escala de avaliação de impacto;

A escala de classificação dos riscos.

Dessa maneira, a modelagem da abordagem tradicional consistiu nas seguintes definições:

a) Os riscos do projeto. Foram identificados 7 riscos para o projeto, são eles:

Atraso na liberação de recursos financeiros;

Falta de apoio político das entidades esportivas ligadas ao esporte;

Falta de patrocínio;

Falta de pessoal da equipe técnica e de apoio;

Indisponibilidade de local para realizar

Treinamentos;

Lesão em competição e/ou treinamento;

Ocorrência de acidente grave ou fatal.

b) A escala de avaliação de probabilidade.

0,75 para probabilidade maior que 65% de ocorrer o risco;

0,55 para probabilidade entre 45 e 65% de ocorrer o risco;

0,35 para probabilidade entre 25 e 45% de ocorrer o risco;

0,15 para probabilidade entre 5 e 25% de ocorrer o risco;

0,05 para probabilidade menor que 5% de ocorrer o risco.

c) A escala de avaliação de impacto.

0,8 para impacto muito alto sobre o objetivo do projeto em questão;

0,4 para impacto alto sobre o objetivo do projeto em questão;

0,2 para impacto moderado sobre o objetivo do projeto em questão;

0,1 para impacto baixo sobre o objetivo do projeto em questão;

0,05 para impacto muito baixo sobre o objetivo do projeto em questão.

d) A escala de classificação dos riscos. Essa escala toma como referência o grau de

importância dos riscos calculado pela multiplicação das avaliações de probabilidade e de impacto dos riscos. Nesse modelo foi utilizada a seguinte escala:

Alto para uma importância do risco entre 0,14 e 1;

Médio para uma importância do risco entre 0,04 e 0,14 (inclusive);

Baixo para uma importância do risco entre 0 e 0,04 (inclusive).

## 4.2 MODELAGEM DA ABORDAGEM MULTICRITÉRIO (ELECTRE-TRI / AHP)

### ELECTRE-TRI PARA A CLASSIFICAÇÃO DOS RISCOS

A utilização do método ELECTRE-Tri requer as seguintes definições (COSTA et al, 2007):

O conjunto "A" das alternativas viáveis:  $A=\{a_1, a_2, a_3, \dots, a_n\}$ ;

O conjunto "F" dos critérios:  $F=\{g_1, g_2, g_3, \dots, g_n\}$ ;

O vetor "W" dos pesos dos critérios:  $W=\{w_1, w_2, w_3, \dots, w_n\}$ ;

Os Limites de preferência;

Os Limites de indiferença;

Os Limites de veto;

O conjunto "B" dos valores que definem as "n+1" categorias:  $B=\{b_1, b_2, \dots, b_n\}$ .

Dessa maneira, a modelagem da abordagem multicritério para a classificação de riscos consistiu nas seguintes definições:

a) O conjunto "A" das alternativas viáveis:  $A=\{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ , onde:

a1= Atraso na liberação de recursos financeiros;

a2= Falta de apoio político das entidades esportivas ligadas ao esporte;

a3= Falta de patrocínio;

a4= Falta de pessoal da equipe técnica e de apoio;

a5= Indisponibilidade de local para realizar Treinamentos;

a6= Lesão em competição e/ou treinamento;

a7= Ocorrência de acidente grave ou fatal.

b) O conjunto "F" dos critérios:  $F=\{g_1, g_2, g_3,$

$g_4, g_5\}$ , onde:

g1= Probabilidade de ocorrência do risco do projeto;

g2= Impacto do risco sobre o custo do projeto;

g3= Impacto do risco sobre o cronograma do projeto;

g4= Impacto do risco sobre o escopo do projeto;

g5= Impacto do risco sobre a qualidade do projeto.

c) O vetor "W" dos pesos dos critérios:  $W=\{w_1, w_2, w_3, w_4, w_5\}$ , onde:

$w_1= w_2= w_3= w_4= w_5= 0,2$  (isto é, os pesos foram tomados como iguais sendo uniformemente distribuídos)

d) Os Limites de preferência são 0,2 para o critério g1 e 0,1 para os demais critérios.

e) Os Limites de indiferença são 0,1 para o critério g1 e 0,05 para os demais critérios.

f) Os Limites de veto são 0,3 para o critério g1 e 0,15 para os demais critérios.

g) O conjunto "B" dos valores que definem as "n+1" categorias:  $B=\{b_1, b_2\}$ , onde:

$b_1= 0,7$  para o critério g1 e 0,4 para os demais critérios. Sendo estes os limites entre as categorias de risco alto e médio;

$b_2= 0,3$  para o critério g1 e 0,2 para os demais critérios. Sendo estes os limites entre as categorias de risco médio e baixo.

### AHP PARA A PRIORIZAÇÃO DE RISCOS

A utilização do método AHP requer as seguintes definições (BELLA et al, 2014):

O foco principal da análise;

O conjunto "C" de critérios:  $C=\{c_1, c_2, c_3, \dots, c_n\}$ ;

O conjunto "S" de subcritérios:  $S=\{s_1, s_2, s_3, \dots, s_n\}$ ;

O conjunto "A" de alternativas:  $A=\{a_1, a_2, a_3, \dots, a_n\}$ ;

A estrutura hierárquica do processo.

Dessa maneira, a modelagem da abordagem multicritério para a classificação de riscos consistiu nas seguintes definições:

a) O foco principal da análise consistiu na ordenação dos riscos previamente classificados pelo método ELECTRE-Tri

b) O conjunto “C” de critérios:  $C=\{c1, c2, c3, c4\}$ , onde:

c1= Influência sobre o custo do projeto;

c2= Influência sobre o cronograma do projeto;

c3= Influência sobre o escopo do projeto;

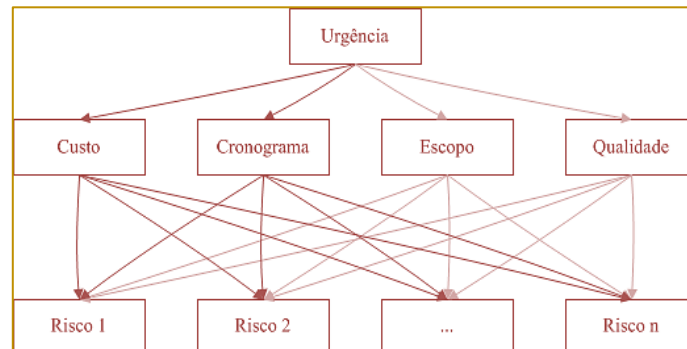
c4= Influência sobre a qualidade do projeto.

c) O conjunto “S” de subcritérios. O conjunto “S” para esse modelo é considerado vazio devido a complexidade nesse contexto de se escolher bons subcritérios sem tendenciar o modelo.

d) O conjunto “A” de alternativas consistiu nos grupos de riscos previamente classificados pelo ELECTRE-Tri. Dessa maneira, o método AHP foi executado para três grupos de alternativas, são eles: o grupo de riscos altos; o grupo de riscos médios; e o grupo de riscos baixos.

e) A estrutura hierárquica do processo dependente da definição prévia dos conjuntos de critérios, subcritérios e alternativas. Desse modo, organizou-se uma estrutura genérica aplicável aos três grupos de alternativas geradas pelo processo de classificação. A estrutura hierárquica do modelo está apresentada na figura 1.

Figura 1: Estrutura hierárquica do processo de priorização de riscos



Fonte: Modelo de priorização (seção 4.3)

## 5.RESULTADOS

### 5.1 ABORDAGEM TRADICIONAL

A abordagem tradicional foi realizada pela equipe de projeto da empresa estudada, sendo o resultado apresentado nesta seção uma adaptação dos documentos do projeto boxe (vide seção 3). Na tabela 4, observa-se a classificação e a prioridade dos riscos do projeto boxe segundo a abordagem tradicional realizada com ênfase em um

processo de decisão consensual.

Conforme pode-se observar na tabela 4, a classificação dos riscos ficou determinada da seguinte maneira: três riscos foram classificados como altos (alternativas a1, a2 e a3); dois foram classificados como médios (alternativas a4 e a6); e, dois como baixos (alternativas a5 e a7). Sendo a ordem de prioridade encontrada conforme a sequência de alternativas: a3, a2, a1, a6, a4, a7, a5.

Tabela 4: Classificação e prioridade dos riscos do projeto boxe segundo a abordagem tradicional

| Classificação | Prioridade | Risco (alternativa de referência)  |
|---------------|------------|--|
| Alto          | 1          | Falta de patrocínio (a3)   |
|               | 2          | Falta de apoio político das entidades esportivas ligadas ao esporte (a2) |
|               | 3          | Atraso na liberação de recursos financeiros (a1)                         |
| Médio         | 4          | Lesão em competição e/ou treinamento (a6)                                |
|               | 5          | Falta de pessoal da equipe técnica e de apoio (a4)                       |
| Baixo         | 6          | Ocorrência de acidente grave ou fatal (a7)                               |
|               | 7          | Indisponibilidade de local para realizar Treinamentos (a5)               |

Fonte: Documentos do projeto estudado (projeto boxe)

## 5.2 ABORDAGEM MULTICRITÉRIO

A abordagem multicritério foi realizada considerando os julgamentos de valor de um especialista em gestão de projetos que participou do projeto como consultor. Vale ressaltar que o ideal nesse ponto, seria considerar que os julgamentos de valor fossem pesados segundo o consenso da mesma equipe de projeto que realizou a abordagem tradicional. Entretanto, não houve disponibilidade para tal reunião, sendo assim, optou-se por considerar a percepção de um

especialista envolvido no projeto.

Na tabela 5, observa-se a classificação e a prioridade dos riscos do projeto boxe segundo a abordagem multicritério. Conforme a tabela 5, a classificação dos riscos ficou determinada da seguinte maneira: cinco riscos foram classificados como médios (alternativas a1, a2, a3, a6 e a7); dois foram classificados como baixos (alternativas a4 e a5); e nenhum como alto. Sendo a ordem de prioridade encontrada conforme a sequência de alternativas: a1, a3, a2, a6, a7, a4, a5.

Tabela 5: Classificação e prioridade dos riscos do projeto boxe segundo a abordagem multicritério

| Classificação | Prioridade | Risco (alternativa de referência)  |
|---------------|------------|--|
| Médio         | 1          | Atraso na liberação de recursos financeiros (a1)                         |
|               | 2          | Falta de patrocínio (a3)   |
|               | 3          | Falta de apoio político das entidades esportivas ligadas ao esporte (a2) |
|               | 4          | Lesão em competição e/ou treinamento (a6)                                |
|               | 5          | Ocorrência de acidente grave ou fatal (a7)                               |
| Baixo         | 6          | Falta de pessoal da equipe técnica e de apoio (a4)                       |
|               | 7          | Indisponibilidade de local para realizar Treinamentos (a5)               |

Fonte: Modelagem e Aplicação dos métodos multicritério (seções 4 e 5)

## COMPARAÇÃO ENTRE ABORDAGENS

Para facilitar a comparação entre as abordagens tradicional e multicritério foi utilizado um código para representar a classificação e a prioridade do risco. O

código é formado por uma letra maiúscula oriunda da inicial de sua classe e um número referente a sua ordem no *ranking* de prioridade.

Na tabela 6, observa-se as mudanças de



classe e na ordem de prioridade entre as abordagens. Conforme a tabela 6, na comparação da abordagem tradicional para a multicritério: dois riscos mantiveram a mesma classificação, quatro foram rebaixados em uma classe e um promovido em uma classe.

Em relação a ordenação, tem-se que: dois riscos mantiveram a posição no *ranking* de prioridade, três foram rebaixados em uma posição, e dois foram promovidos sendo um em uma posição e o outro em duas.

Tabela 6: Comparativo entre as abordagens tradicional e multicritério para a classificação e priorização de riscos

| Risco (alternativa de referência)  | Tradicional | Multicritério | Mudança de classe | Mudança de ordem |
|--|-------------|---------------|-------------------|------------------|
| Atraso na liberação de recursos financeiros (a1)                         | A3          | M1            | Desceu 1          | Subiu 2          |
| Falta de apoio político das entidades esportivas ligadas ao esporte (a2) | A2          | M3            | Desceu 1          | Desceu 1         |
| Falta de patrocínio (a3)   | A1          | M2            | Desceu 1          | Desceu 1         |
| Falta de pessoal da equipe técnica e de apoio (a4)                       | M5          | B6            | Desceu 1          | Desceu 1         |
| Indisponibilidade de local para realizar Treinamentos (a5)               | B7          | B7            | Manteve           | Manteve          |
| Lesão em competição e/ou treinamento (a6)                                | M4          | M4            | Manteve           | Manteve          |
| Ocorrência de acidente grave ou fatal (a7)                               | B6          | M5            | Subiu 1           | Subiu 1          |

Fonte: Abordagens tradicional e multicritério (tabelas 4 e 5)

## 6. CONCLUSÃO

A comparação entre as abordagens mostrou que a abordagem multicritério vai ao encontro da abordagem tradicional. Tal fato sugere que, a abordagem multicritério pode ser utilizada para robustecer a análise qualitativa de riscos sem que seja necessário abrir mão do caráter qualitativo do processo. Isso por que, as ferramentas multicritério trabalham no sentido de racionalizar a subjetividade inerente ao processo, permitindo que os julgamentos de valor possam ser verificados, analisados e questionados.

Nesse sentido, como principal contribuição o artigo traz uma proposta de modelo multicritério para a classificação e priorização de riscos. Dessa maneira, a metodologia proposta pode ser aplicada na prática para outros projetos. Sendo assim, é importante ressaltar as principais limitações presente no artigo, são elas: a utilização de um grupo pequeno de riscos; a utilização de grupos diferentes na aplicação das abordagens; e a utilização dos julgamentos de valor de um único especialista na abordagem multicritério.

## REFERÊNCIAS

[1] Bella, RLF & Mendes, AM & Ferraz, EM. Análise multicritério na seleção de projetos. Anais do 34 Encontro Nacional de Engenharia de Produção (Xxxiv Enegep), 2014.

[2] Bouyssou, D & Marchant, T. On the relations between Electre TRI-B and Electre TRI-C and on a new variant of Electre TRI-B. European Journal of Operational Research, Volume 242, Issue 1, 2015.

[3] Costa, HG & Santafé Junior, Hpg & Haddad, AN. Uma contribuição do método ELECTRE TRI à obtenção da classificação de riscos industriais. Investigação Operacional, Vol. 27, 2007.

[4] Costa, HG & Vilas Boas, Gar & Freitas, Alp & Gomes, CFS. Modelo multicritério para avaliação e classificação da gestão organizacional: proposta e caso de uso. Production, v. 24, n. 3, p. 521-535, 2014.

- [5] Freitas, ALP & Costa, HG. Avaliação e classificação da qualidade de serviços utilizando uma abordagem multicritério. *Gestão & Produção*, v.5, n.3, p. 272-283, 1998.
- [6] Gaudard, RM & Ribeiro, ALD & Oliveira, EC & Souza, JCF & Monteiro, SBS. A
- [7] avaliação dos riscos da copa do mundo de futebol de 2014: uma ferramenta usando análise de multicritério. *Anais do 32 Encontro Nacional de Engenharia de Produção (Xxxii Enegep)*, 2012.
- [8] Gomes, LFAM. Teoria da decisão. São Paulo: Thomson Learning, 2007.
- [9] Junior, RR & Carvalho, MM. Relacionamento entre gerenciamento de risco e sucesso de projetos. *Produção*, v. 23, n. 3, p. 570-581, jul./set. 2013.
- [10] Project Management Institute (PMI). Um guia do conhecimento em gerenciamento de projetos (Guia PMBoK). 5 edição, editora Saraiva, 2013.
- [11] Silva, ACS & Nascimento, LPAS & Belderrain, MCN. Método de apoio multicritério à decisão na seleção e priorização de portfólio de projetos. *Anais do 13 Encontro de Iniciação Científica e PósGraduação do ITA (Xiii Encita)*, 2007.

## ANEXO

Tabela 7: Julgamentos de valor utilizados para a classificação dos riscos pelo método ELECTRE-Tri

|    | a1   | a2   | a3   | a4   | a5   | a6   | a7   |
|----|------|------|------|------|------|------|------|
| g1 | 0,35 | 0,55 | 0,55 | 0,15 | 0,15 | 0,35 | 0,15 |
| g2 | 0,8  | 0,2  | 0,4  | 0,1  | 0,1  | 0,4  | 0,4  |
| g3 | 0,2  | 0,2  | 0,2  | 0,4  | 0,1  | 0,4  | 0,4  |
| g4 | 0,1  | 0,1  | 0,1  | 0,1  | 0,1  | 0,1  | 0,1  |
| g5 | 0,2  | 0,2  | 0,4  | 0,8  | 0,4  | 0,4  | 0,8  |

Fonte: Especialistas em gestão do projeto (consultores do projeto boxe)

Tabela 8: Julgamentos de valor utilizados para a priorização dos riscos pelo método AHP no nível de comparação entre critérios

|    | c1 | c2 | c3 | c4 |
|----|----|----|----|----|
| c1 | 1  | 1  | 1  | 1  |
| c2 | 1  | 1  | 1  | 1  |
| c3 | 1  | 1  | 1  | 1  |
| c4 | 1  | 1  | 1  | 1  |

Fonte: Especialistas em gestão do projeto (consultores do projeto boxe)

Tabela 9: Julgamentos de valor utilizados para a priorização dos riscos pelo método AHP no nível de comparação entre alternativas (grupo de riscos médios)

|    | a1  | a2  | a3  | a6  | a7 |
|----|-----|-----|-----|-----|----|
| a1 | 1   | 4   | 2   | 6   | 6  |
| a2 | 1/2 | 1   | 1/3 | 2   | 3  |
| a3 | 1/4 | 3   | 1   | 4   | 5  |
| a6 | 1/6 | 1/2 | 1/4 | 1   | 2  |
| a7 | 1/6 | 1/3 | 1/5 | 1/2 | 1  |

Fonte: Especialistas em gestão do projeto (consultores do projeto boxe)

Tabela 10: Julgamentos de valor utilizados para a priorização dos riscos pelo método AHP no nível de comparação entre alternativas (grupo de riscos baixos)

|    | a4  | a5 |
|----|-----|----|
| a4 | 1   | 4  |
| a5 | 1/4 | 1  |

Fonte: Especialistas em gestão do projeto (consultores do projeto boxe)

# Capítulo 7

## O PROBLEMA DO CAIXEIRO VIAJANTE COM PASSAGEIROS E LOTAÇÃO

*Ranmsés E. M. Bastos*

*Marco C. Goldberg*

*Elizabeth F. G. Goldberg*

**Resumo:** O Problema do Caixeiro Viajante com Passageiros e Lotação é uma versão do Problema do Caixeiro Viajante clássico onde o caixeiro é o motorista de um veículo que compartilha os custos de viagem com passageiros. Além de dividir os custos do percurso, o caixeiro pode se valer, também, dos descontos das *high-occupancy vehicle lanes*, que são faixas de trânsito que isentam veículos lotados do pagamento de pedágio. Este artigo apresenta um modelo matemático resumido para este problema e um algoritmo baseado nas meta-heurísticas *Simulated Annealing* e Busca em Vizinhança Variável. Os resultados dos algoritmos heurísticos são comparados às soluções ótimas obtidas por um algoritmo exato.

**Palavras Chave:** Ridesharing. Carpool. Simulated annealing.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2016)

## 1. INTRODUÇÃO

Este trabalho apresenta um novo problema de roteamento intitulado Problema do Caixeiro Viajante com Passageiros e Lotação (PCV-PL). O modelo é uma variante do Problema do Caixeiro Viajante (PCV) e incorpora elementos das classes *Pickup and Delivery* (P&D) e *Ridesharing*. O problema em questão inclui, no PCV tradicional, a consideração do valor econômico relativo ao aproveitamento do potencial de transporte solidário do veículo do caixeiro para reduzir as despesas do motorista e proporcionar um meio de transporte mais eficiente para os passageiros e de menor impacto sobre o meio ambiente. O problema está correlacionado com o Problema do Caixeiro Viajante com Passageiros (PCV-Pa) apresentado por Calheiros [2015], sendo dele uma variante. No PCV-Pa, um motorista deve percorrer um ciclo hamiltoniano sobre um conjunto de localidades e pode transportar passageiros com o objetivo de, com eles, ratear custos de viagem. No PCV-PL o caixeiro pode beneficiar-se adicionalmente pela isenção de pedágio em vias que possuam tal desconto. A exigência para a isenção é de que o veículo tenha uma lotação de passageiros maior ou igual a um certo valor mínimo. Essas vias são denominadas usualmente de *high-occupancy vehicle lanes* ou faixas *hov*. Em outras palavras, a parcela de custos referente ao pedágio é nula caso o veículo esteja suficientemente ocupado ao trafegar pelas conexões *hov*. Estas faixas especiais de trânsito estão presentes em várias partes do mundo e fazem parte de esforços governamentais para redução nominal do tráfego em rodovias. Além do incentivo financeiro da isenção de pedágio, é comum que as faixas sejam exclusivas e apresentem menor incidência de congestionamentos, bem como permitam evitar as filas ou afunilamentos nas praças de pedágio.

O PCV-PL está relacionado a outros problemas clássicos de roteamento. O PCV é um caso particular do PCV-PL quando não são oferecidos assentos para passageiros. Ao permitir o compartilhamento de assentos, o problema também pode ser incluído genericamente na classe dos problemas de *ridesharing* a qual é definida sob diferentes pontos de vista conforme Amey et al. [2011]. De uma forma geral, um problema da classe *ridesharing* corresponde à reunião de duas ou mais pessoas para partilhar uma única viagem em um veículo, sem levar em conta

um acordo prévio ou um histórico de cooperação [Dailey et al. 1999]. Uma visão estruturada do problema pode ser obtida em [Furuhata et al. 2013].

Como um problema que embarca e desembarca passageiros ao longo da rota, o PCV-PL compartilha elementos com os problemas de roteamento com P&D. Parragh et al. [2008a, 2008b] revisam a literatura do P&D. O PCV-PL distingue-se do P&D tanto pela função objetivo que associa o ganho do problema à razão entre o custo de trafegar na via e o carregamento do veículo, quanto em virtude da programação de embarque e desembarque atender exclusivamente ao interesse do caixeiro. No PCV-PL não existe demanda de transporte a ser atendida. Uma das variantes do P&D que se aproxima do PCV-PL é o *Single Vehicle Routing Problem with Deliveries and Selective Pickups* [Gribkovskaia et al. 2008], que permite não satisfazer as coletas demandadas.

O PCV-PL ainda compartilha elementos com os problemas de roteamento com coleta, como *Travelling Salesman with profits* [Feillet et al. 2005], *Prize Collecting Travelling Salesman* [Balas 2004], *Attractive Travelling Salesman* [Erdoğan, et al. 2010], *Selective Travelling Salesman* [Gendreau et al. 1998] e *Orienteering Problems* [Vansteenwegen et al. 2011], contudo o ganho associado ao embarque do passageiro não é fixo ou limitado, sendo variável em função da rota adotada.

Poucos trabalhos que abordam a otimização de rotas em função de ganhos obtidos com ligações do tipo *hov* são relatados na literatura. Wang et al. [2013] abordam uma variante P&D com ligações *hov*, janelas de tempo e *ridesharing*.

Neste trabalho é apresentado um algoritmo exato e algoritmos baseados nas meta-heurísticas *Simulated Annealing* e *Busca em Vizinhança Variável para o PCV-PL*. São reportados os resultados de um experimento computacional onde foram comparadas as abordagens propostas.

As próximas seções estão organizadas da seguinte forma: a seção 2 apresenta o PCV-PL; a seção 3 apresenta os algoritmos desenvolvidos; a seção 4 detalha os experimentos computacionais e a seção 5 apresenta a conclusão do trabalho.

## 2. PROBLEMA DO CAIXEIRO VIAJANTE COM PASSAGEIROS E LOTAÇÃO

Mesclando a classe dos problemas de *ridesharing* com a classe dos problemas do caixeiro viajante, o problema pode ser entendido como uma variante do Caixeiro Viajante Capacitado em que o caixeiro pode reduzir seus custos através de compartilhamento de despesas de viagem com eventuais passageiros embarcados e, adicionalmente, beneficia-se de incentivos sociais pelo transporte solidário. No PCV-PL considera-se que a disposição das localidades e ligações rodoviárias associadas ao problema são representadas por um grafo ponderado  $G = (N, M)$  onde  $N = \{1, \dots, n\}$  modela o conjunto de localidades e  $M = \{1, \dots, m\}$  o conjunto de ligações rodoviárias existentes entre as localidades. O ciclo do caixeiro é iniciado na localidade de número 1.

Os passageiros desejam, de uma forma geral, embarcar e desembarcar em localidades que pertencem ao ciclo do caixeiro. Consequentemente, os passageiros podem compartilhar assentos no veículo do caixeiro em parte de sua rota, desde que a cidade de embarque do passageiro venha antes da cidade de desembarque. O caixeiro não cobra tarifa fixa pelos deslocamentos e sim rateia as despesas do deslocamento de cada trecho com os passageiros embarcados no trecho. Neste trabalho considera-se que os passageiros não são sensíveis ao caminho percorrido entre a cidade de embarque e a cidade destino, ou seja, ao número de cidades intermediárias entre o embarque e o desembarque. No caso examinado neste trabalho considera-se que existe apenas 1 pessoa solicitando transporte em cada cidade, i.e., o número total de solicitações é  $n$ . Cada passageiro possui um limite de orçamento, ou seja, um valor máximo que o mesmo está disposto a pagar por suas despesas de deslocamento. A despesa total do deslocamento não deve ultrapassar  $t_l$  unidades monetárias,  $l = 1, \dots, n$ . O  $l$ -ésimo passageiro deseja iniciar seu caminho na cidade  $P_l$  e o terminá-lo na cidade  $Q_l$ ,  $Q_l \neq P_l$ ,  $l = 1, \dots, n$ . No caso particular examinado,  $Q_1 \neq 1$ .

Ao embarcar na cidade  $P_l$ , o passageiro  $l$  passa a dividir em rateio uniforme, com os demais ocupantes do veículo, outros

passageiros e o motorista, as despesas dos trechos em que  $l$  seguir embarcado. A parte da despesa que cabe a cada ocupante, passageiros e motorista, para trafegar em cada trecho  $(i, j)$ ,  $(i, j) \in M$ , é obtida pela divisão do custo do deslocamento entre as cidades  $i$  e  $j$  pelo número de pessoas que ocupam o veículo na aresta  $(i, j)$ .

Quando o veículo do caixeiro atravessar uma aresta que está habilitada ao benefício do desconto associado ao transporte solidário, chamada de aresta *hov*, e o veículo possuir a lotação mínima exigida para habilitar o incentivo, o valor da aresta é decrescido do valor do desconto. Consequentemente, o valor de rateio da aresta *hov* sofre um decréscimo em seu custo operacional caso o veículo por ela circule atendendo a lotação socialmente incentivada. De uma forma geral, trafegar com a lotação exigida em arestas *hov* sugere a possibilidade de reduções adicionais e que podem ser somadas às reduções obtidas pelo rateio.

A solução do PCV-PL consiste na determinação de um ciclo Hamiltoniano no grafo  $G$  e da lista de pessoas, passageiros, que terão suas solicitações de transporte atendidas. As restrições devem garantir que as condições de embarque e desembarque, os limites de orçamento dos passageiros e a capacidade do veículo sejam atendidas. As arestas do tipo *hov* devem ser diferenciadas em valor das arestas normais do problema e é necessário que restrições controlem a ativação do desconto de ocupação máxima. O conjunto *Hov* contém as arestas do tipo *hov*.

O objetivo é minimizar a função (1) onde  $d_{ij}$  denota o custo da aresta  $(i, j)$ ;  $f_{ij}$  é uma variável Booleana igual a 1 se a aresta  $(i, j)$  pertence ao ciclo Hamiltoniano definido para o caixeiro e 0 caso contrário;  $c_{ij}$  é a redução prevista para o custo da aresta  $(i, j)$  caso o carro circule em  $(i, j)$  com lotação que garanta o desconto na via;  $w_{ij}$  é uma variável Booleana igual a 1 se o caixeiro atravessa a aresta  $(i, j)$  com lotação que garanta o desconto na via e 0 caso contrário;  $L$  denota o conjunto de pessoas demandando transporte e  $v_{ij}^l$  é uma variável Booleana que possui valor 1 se o carona  $l$  foi transportado na aresta  $(i, j)$  e 0 em caso contrário.

$$\sum_{i,j \in N} \frac{d_{ij}f_{ij} - c_{ij}w_{ij}}{\sum_{l \in L} v_{ij}^l + 1} \quad (1)$$

A função (1) é de natureza não linear. Um modelo matemático deve forçar que o produto  $c_{ij}w_{ij}$  seja igual a zero se  $(i, j) \notin \text{Hov}$ . *Todavia, para que seja diferente de zero, é necessário que  $(i, j) \in \text{Hov}$  e a lotação do veículo expressa por  $\sum_{l \in L} v_{ij}^l$  seja maior ou igual a  $h_{ij}$ , a lotação mínima do veículo*

exigida na aresta  $(i, j)$  para que o desconto seja alcançado. Assim, a variável  $w_{ij}$  deve receber valor 1 quando  $\sum_{l \in L} v_{ij}^l \geq h_{ij}$ .

O total de passageiros transportados em uma aresta  $(i, j)$  não deve ultrapassar a capacidade  $R$  do veículo. Esta restrição está expressa em (2).

$$Rf_{ij} - \sum_{l \in L} v_{ij}^l \geq 0 \quad \forall (i, j) \in M \quad (2)$$

A condição de limite de orçamento de cada passageiro pode ser expressa como em (3).

$$\sum_{i \in N} \sum_{j \in N} \frac{d_{ij}v_{ij}^l}{\sum_{k \in L} v_{ij}^k + 1} - t_l \leq 0 \quad \forall l \in L \quad (3)$$

Como uma generalização do PCV, a solução do PCV-PL é pelo menos tão difícil quanto a solução do PCV, um conhecido problema NP-Árduo [Garey e Johnson 1979]. Todavia, o número de possíveis decisões no PCV-PL é maior do que no PCV, uma vez que o custo do ciclo não pode ser obtido *a priori* através da soma do custo das arestas por depender do carregamento do veículo. Por sua vez, o carregamento depende da sequência de visitas, da decisão de embarque e da capacidade do veículo. O esquema de embarques e desembarques ótimo calculado sobre a rota ótima do caixeiro viajante clássico pode não ser a solução ótima para o PCV-PL, uma vez que o carregamento ótimo do veículo dependerá também do destino desejado pelos passageiros. Há que se determinar uma rota de baixo custo e que permita um carregamento excelente do veículo. A dificuldade da solução desse problema é sugerida na própria modelagem simplificada, pois se evidencia a necessidade de alcançar a solução de um problema de programação inteira com função objetivo não linear e pelo menos uma restrição não linear de tarifa, além de um conjunto de restrições lógicas que garantam o comportamento em função do desconto da lotação máxima.

### 3. ALGORITMOS PARA O PCV-PL

Foi desenvolvido um algoritmo exato do tipo *backtracking* de forma a estabelecer um referencial de comparação de desempenho para os algoritmos heurísticos para as instâncias onde foi possível encontrar soluções ótimas. Uma solução  $s$  para o problema é composta de duas partes: uma permutação de  $n$  vértices que define a rota do caixeiro e uma lista de atribuição de passageiros. A lista com a permutação dos  $n$  vértices será notada pelo próprio  $s$  e a lista de atribuição de passageiros será notada por  $L$ .

#### 3.1 ALGORITMO EXATO

O Algoritmo 1 descreve, em pseudocódigo, a estratégia utilizada para determinação da solução exata. No passo 1 são gerados todos os ciclos Hamiltonianos com início e fim no vértice de origem. O melhor custo de uma solução é iniciado no passo 2. No laço dos passos 3-6 do algoritmo 1 é chamado o procedimento *Embarque* que define quais passageiros serão embarcados. No passo 3, é feita uma verificação para evitar que soluções não promissoras sejam desenvolvidas. O teste é feito verificando se a

melhor solução gerada até o momento tem custo maior que a solução parcial sendo gerada. O limite inferior é definido pela divisão entre o somatório dos custos do ciclo supondo que todos os descontos de pedágios são ativados ( $s.comprimento$ ) e a capacidade  $R$  do veículo. Uma árvore de busca pelo melhor esquema de carregamento de  $s$  é iniciada no passo 4. O custo inicial da solução é dado pelo custo da rota do caixeiro em  $s$ . No procedimento *Embarque* é construída a árvore de busca para alocação de passageiros em  $s$ . O custo do nó da árvore é atualizado sempre que uma solução é

criada ou modificada (passos 1, 6 e 9 do procedimento *Embarque*). Para cada cidade  $s[i]$  é verificado se o candidato a embarque naquela cidade,  $p$ , pode ser embarcado. Caso seja possível, é criado um nó na árvore de busca considerando a inclusão deste passageiro, ou seja,  $L[s[i]]$  recebe valor 1 e o custo com a inclusão de  $p$  é calculado (passos 4 e 5 do procedimento *Embarque*). Também, em toda iteração, é criado um nó na árvore de busca onde o passageiro  $p$  não é embarcado (passos 7 e 8 do procedimento *Embarque*).

### Algoritmo 1 – Backtracking

| Algoritmo 1 – BT                 |  |
|----------------------------------|--|
| 1                                | <b>Gere todos os ciclos hamiltonianos e coloque no conjunto <math>CH</math></b>                                      |
| 2                                | $melhor.custo \leftarrow infinito$   |
| 3                                | <b>Para cada <math>s \in CH</math>, se <math>melhor.custo &gt; (s.comprimento)/R</math></b>                          |
| 4                                | <b>Inicie uma árvore de busca <math>\beta</math> tal que <math>\beta.raiz \leftarrow s</math></b>                    |
| 5                                | $\beta.raiz.custo \leftarrow calcula\_custo(s)$  |
| 6                                | <b>Execute <math>Embarque(\beta.raiz, L)</math></b>  |
| Procedimento $Embarque(node, L)$ |  |
| 1                                | <b>Se <math>node.custo \leq melhor.custo</math>, faça <math>melhor.custo \leftarrow node.custo</math></b>            |
| 2                                | <b>Para <math>i \leftarrow 1</math> até <math>n</math></b>   |
| 3                                | <b>Se o embarque do passageiro em <math>s[i]</math> é viável faça</b>  |
| 4                                | $L[s[i]] \leftarrow 1$ ; $node.right.custo \leftarrow calcula\_custo(s)$   |
| 5                                | $Embarque(node.right, L)$  |
| 6                                | <b>Se <math>node.right.custo \leq melhor.custo</math> faça <math>melhor.custo \leftarrow node.right.custo</math></b> |
| 7                                | $L[s[i]] \leftarrow 0$   |
| 8                                | $Embarque(node.left, L)$   |
| 9                                | <b>Se <math>node.left.custo \leq melhor.custo</math> faça <math>melhor.custo \leftarrow node.left.custo</math></b>   |

### 3.2 ALGORITMOS HEURÍSTICOS

Os algoritmos heurísticos desenvolvidos foram inspirados nas meta-heurísticas *Simulated Annealing* (SA) e *Busca em Vizinhaça Variável* (VNS). Tal escolha levou em conta os resultados promissores relatados para o SA em problemas de diversas áreas, bem como a flexibilidade que a estratégia VNS permite para buscas locais. A composição das duas abordagens supostamente tem potencial para uma evolução progressiva da busca e fuga de ótimos locais através do esquema *annealing*, conhecido como critério de Metropolis, em referência ao trabalho tido como precursor do desenvolvimento do método [Metropolis et al.

1953]. Por outro lado, herda a capacidade de exploração sistemática e aprofundada de boas vizinhanças da proposta VNS.

Basicamente, o SA pode partir de qualquer solução inicial  $s$ . Seu procedimento iterativo consiste na geração de uma solução vizinha  $v \in \varphi(s)$ , onde  $\varphi(s)$  representa a vizinhança de  $s$ , e no teste de aceitação para a transição de  $s$  para  $v$ . A aceitação da transição é realizada pela avaliação da variação da energia  $\Delta = \xi(v) - \xi(s)$  entre as duas soluções. A probabilidade da aceitação da transição  $P_i(s \rightarrow v)$  é calculada pelo conjunto de equações (4). O parâmetro  $T_i$ , associado à probabilidade  $P_i$  no passo  $i$ , permite modular a probabilidade da aceitação da transição



tanto em função da diferença de energia entre

as soluções quanto do estágio  $i$  da busca.

$$P_i(s \rightarrow v) = \begin{cases} e^{-\Delta/T_i}, & \text{se } \Delta \geq 0 \\ 1, & \text{se } \Delta < 0 \end{cases} \quad (4)$$

O termo  $e^{-\Delta/T_i}$  representa a probabilidade da busca evoluir para um estado pior (mais energético) que o estado  $i$ , o que diminui a chance de estagnação da busca em ótimos locais. A cada rodada o sistema é lentamente resfriado e o *loop* descrito se prolonga até que seja alcançada uma temperatura mínima  $T_{min}$  ou até que sejam analisados  $lt$  estados (equivalente à quantidade de iterações). Para

cada implementação é preciso definir os seguintes parâmetros:  $T_{min}$  e  $lt$ , já descritos;  $T_0$  - temperatura inicial;  $T_i$  - mecanismo de decréscimo da temperatura;  $\varphi$  - função geradora da vizinhança; e  $\xi$  - função energia. No presente trabalho foi adotado o mecanismo homogêneo para  $T_i$ , o qual encontra-se descrito na equação (5).

$$T_{i+1} = \alpha \cdot T_i \quad 0 < \alpha < 1 \quad (5)$$

A variação da vizinhança foi inserida no SA como um esquema de geração da solução  $v \in \varphi(s)$ . A função vizinhança  $\varphi(s)$  passa a ser uma interface para uma escolha equiprovável de  $k$  estruturas distintas de vizinhança  $\varphi_1(s), \varphi_2(s), \dots, \varphi_k(s)$ . Foram adotadas  $k = 4$  estruturas de vizinhança para variação da rota, cada uma delas baseada nos operadores: *troca*, *inserção*, *inversão* e *vizinho mais próximo*. No operador *troca*, troca-se a posição de duas cidades no ciclo. Por exemplo, na figura 1, são trocadas de

posição as cidades 4 e 2. No operador *inserção* uma cidade é removida de sua posição e inserida após outra. Por exemplo, na figura 1 a cidade 4 é removida e inserida após a 9. No operador *inversão*, uma sequência de cidades é invertida. Por exemplo, a sequência entre as cidades 4 e 2 é invertida. No operador *vizinho mais próximo*, é escolhida uma cidade e é colocada ao lado da primeira sua cidade mais próxima. Por exemplo, na Figura 1 a cidade 2 é colocada ao lado da cidade 4.

Figura 2 – Operadores para variação da rota



Para atribuir passageiros à rota, foi adotada a estratégia construtiva descrita no Algoritmo 2, designado por *Heurística de Carregamento (HC)*. A heurística *HC* recebe como parâmetros de entrada uma solução  $s$  ainda sem passageiros atribuídos e a lista *Tarmax* que contém o máximo que cada passageiro está disposto a pagar. A saída de *HC* é a rota

e a atribuição de passageiros na lista  $L$ . No passo 1 do Algoritmo 2, a variável  $\theta$  recebe valor falso. Esta variável receberá valor verdadeiro quando for atribuído um carregamento de passageiros válido para a rota da solução  $s$ . Inicialmente, todos os candidatos a embarque estão desmarcados e somente o motorista está no carro (passo 2 do

Algoritmo 2). O laço entre as linhas 3 e 16 faz a alocação de passageiros. No passo 5 são verificadas as primeiras condições de embarque do passageiro. Tais condições são:

- a) o ponto de embarque do passageiro encontra-se antes do seu ponto de desembarque na rota e
- b) existe possibilidade de a tarifa final do passageiro ser menor que sua tarifa limite, ou seja, se a tarifa limite do passageiro é maior que a divisão entre o somatório dos custos dos trechos entre o embarque e o desembarque, supondo que todos os descontos de pedágios são ativados, e a capacidade  $R$ .

O embarque do passageiro somente é aceito se o passageiro não tiver sido marcado em

uma iteração anterior e se ainda existir assento disponível, conforme verificado no passo 6 do Algoritmo 2. O custo da solução (rota e embarques) é calculado no passo 7. As linhas 8-14 verificam se existe algum passageiro excedendo a tarifa máxima admitida para ser marcado. No caso de existir mais de um, é marcado o passageiro que tiver a maior diferença entre o seu limite e a tarifa atual. O passageiro marcado não é considerado em futuras iterações do algoritmo. Se o esquema de embarque possui algum passageiro na condição descrita, então o esquema é descartado e o algoritmo reinicia a construção da alocação de passageiros (linhas 15-16). Se o esquema é válido, o algoritmo é finalizado.

### Algoritmo 2 – Heurística de Carregamento

| Algoritmo 2 – $HC(s, T_{max})$ |  |
|--------------------------------|--|
| 1                              | $\theta \leftarrow FALSE$  |
| 2                              | <b>Desmarque todos os passageiros; Para <math>i \leftarrow 1</math> até <math>n</math> faça <math>L[i] \leftarrow 0</math></b> |
| 3                              | <b>Enquanto <math>\theta = FALSE</math></b>  |
| 4                              | <b>Para <math>i \leftarrow 1</math> até <math>n</math></b>   |
| 5                              | <b>Se o embarque do passageiro <math>p</math> no vértice <math>s[i]</math> é viável faça</b>                                   |
| 6                              | <b>Se <math>p</math> é desmarcado e existem assentos disponíveis faça <math>L[s[i]] \leftarrow 1</math></b>                    |
| 7                              | $s.custo \leftarrow calcula\_custo(s)$   |
| 8                              | $maxtar \leftarrow 0; j \leftarrow 0$  |
| 9                              | <b>Para <math>i \leftarrow 1</math> até <math>n</math></b>   |
| 10                             | <b>Se <math>L[i] = 1</math> faça <math>tarifa[i] \leftarrow calcula\_tarifa(i)</math></b>                                      |
| 11                             | <b>Se <math>maxtar &lt; tarifa[i] - T_{max}[i]</math> faça</b>   |
| 12                             | $maxtar \leftarrow tarifa[i] - T_{max}[i]; j \leftarrow i$   |
| 13                             | <b>Se <math>j \neq 0</math> faça</b>   |
| 14                             | <b>Marque o passageiro <math>j</math></b>  |
| 15                             | <b>Para <math>i \leftarrow 1</math> até <math>n</math> faça <math>L[i] \leftarrow 0</math></b>                                 |
| 16                             | <b>Caso contrário, <math>\theta \leftarrow TRUE</math></b>   |

Foram desenvolvidos dois algoritmos heurísticos que possuem o mesmo arcabouço: *Simulated Annealing com operadores (SAop)* e *Simulated Annealing com Busca Local (SABL)*. Os dois distinguem-se pela aplicação ou não de buscas locais na geração de vizinhos. O método para gerar a solução inicial é o mesmo em ambos os casos e consiste em gerar uma rota aleatória, aplicar a heurística de Lin e Kernighan [1973], *LKH*, e atribuir os embarques conforme o Algoritmo 2.

Partindo da solução  $s$ , o Algoritmo 3 gera um vizinho para  $s$  por meio da aplicação única de um operador selecionado em sorteio aleatório equiprovável (passo 6), de modo que a cada iteração é avaliada apenas uma nova solução. O carregamento de passageiros na nova rota é realizado com o Algoritmo 2 (passo 7 do Algoritmo 3). A nova solução,  $v$ , é avaliada. O valor de avaliação associado a uma solução é notado por  $\xi()$ . A avaliação é feita pela expressão (1).

### Algoritmo 2 – Simulated Annealing com operadores

| Algoritmo 3 – $SAop(It, T, Tarmax)$ |  |
|-------------------------------------|--|
| 1                                   | <b>Crie</b> a rota da solução inicial ( $s$ )  |
| 2                                   | $s \leftarrow HC(s, Tarmax)$ ; $T \leftarrow T_0$ ; $melhor \leftarrow s$                                |
| 3                                   | <b>Enquanto</b> ( $T \geq T_{min}$ )   |
| 4                                   | <b>Para</b> $i \leftarrow 1$ até $It$  |
| 5                                   | <b>Sorteie</b> um operador $\gamma$  |
| 6                                   | <b>Aplique</b> $\gamma$ à rota de $s$ gerando $s'$   |
| 7                                   | $v \leftarrow HC(s', Tarmax)$  |
| 8                                   | $\Delta \leftarrow \xi(v) - \xi(s)$  |
| 9                                   | <b>Se</b> $\Delta < 0$ <b>faça</b>   |
| 10                                  | $s \leftarrow v$   |
| 11                                  | <b>Se</b> $\xi(melhor) > \xi(s)$ <b>faça</b> $melhor \leftarrow s$                                       |
| 12                                  | <b>Se</b> $\Delta \geq 0$ <b>faça</b> $v$ é aceita para substituir $s$ com probabilidade $e^{-\Delta/T}$ |
| 13                                  | <b>Atualize</b> $T$  |

### Algoritmo 3 – Simulated Annealing com Buscas Locais

| Algoritmo 4 – $SABL(It, T, Tarmax)$ |  |
|-------------------------------------|--|
| 1                                   | <b>Crie</b> a rota da solução inicial ( $s$ )  |
| 2                                   | $s \leftarrow HC(s, Tarmax)$ ; <b>Aplique</b> <i>Busca Local para Carregamento</i> em $s$                |
| 3                                   | $T \leftarrow T_0$ ; $melhor \leftarrow s$   |
| 4                                   | <b>Enquanto</b> ( $T \geq T_{min}$ )   |
| 5                                   | <b>Para</b> $i \leftarrow 1$ até $It$  |
| 6                                   | <b>Sorteie</b> um operador $\gamma$  |
| 7                                   | <b>Execute</b> <i>Busca Local para Rota</i> em $s$ utilizando $\gamma$ gerando $v$                       |
| 8                                   | $\Delta \leftarrow \xi(v) - \xi(s)$  |
| 9                                   | <b>Se</b> $\Delta < 0$ <b>faça</b>   |
| 10                                  | $s \leftarrow v$   |
| 11                                  | <b>Se</b> $\xi(melhor) > \xi(s)$ <b>faça</b> $melhor \leftarrow s$                                       |
| 12                                  | <b>Se</b> $\Delta \geq 0$ <b>faça</b> $v$ é aceita para substituir $s$ com probabilidade $e^{-\Delta/T}$ |
| 13                                  | <b>Se</b> $v$ substituiu $s$ <b>faça</b> /* Neste caso $s \leftarrow v$ em passo anterior */             |
| 14                                  | <i>Busca Local para Carregamento</i> em $s$ gerando $v'$   |
| 15                                  | <b>Se</b> $\xi(v') < \xi(s)$ <b>faça</b> $s \leftarrow v'$   |
| 16                                  | <b>Se</b> $\xi(melhor) > \xi(s)$ <b>faça</b> $melhor \leftarrow s$                                       |
| 17                                  | <b>Atualize</b> $T$  |

O Algoritmo 4 faz uso de buscas locais para rota e carregamento. A inserção de tal artifício objetivou explorar o espaço de soluções de maneira mais completa que o Algoritmo 3, abrangendo regiões do espaço que potencialmente não foram por este visitadas.

A *Busca Local para Carregamento* (passos 2 e 14 do Algoritmo 4) mantém inalterada a rota e objetiva encontrar um esquema de

carregamento que resulte em uma solução de menor custo. São avaliados todos os passageiros embarcáveis que não embarcaram na solução original. O passageiro que possui o trecho mais longo é inserido no esquema de embarque. A nova solução é avaliada. Se ela for viável, é tomada como resultado uma vez que é, necessariamente, melhor que a solução

prévia. Caso contrário, é iniciado o procedimento de busca por uma solução viável que obrigatoriamente inclua o passageiro com o maior trecho e exclua o embarque de, no máximo, 2 passageiros que estão na solução corrente. Seguindo este critério, todas as combinações de 1 ou 2 são então avaliadas e a saída do procedimento é a melhor solução dentre as avaliadas na busca e a própria solução inicial.

A *Busca Local para Rota* (passo 7 do Algoritmo 4) altera a rota e consequentemente o esquema de carregamento. O processo utiliza um dos operadores de modificação de rota, selecionado em sorteio equiprovável. O operador é aplicado sistematicamente a cada um dos vértices da rota, gerando um conjunto de novas rotas com cardinalidade  $O(n)$ . Cada rota do conjunto é então carregada por meio do Algoritmo 2 e as novas soluções são avaliadas. A saída é a melhor solução dentre aquelas pertencentes ao conjunto gerado, que é obrigatoriamente diferente da solução original.

#### 4. EXPERIMENTO COMPUTACIONAL

Os algoritmos foram implementados na linguagem *C++*. As execuções dos algoritmos *SAop* e *SABL* se deram num computador portátil com memória RAM de 8 GB, processador *Intel Core i7* modelo *3615QM* com 2,3 GHz e sistema operacional *OSX 10.11.4*. A determinação das soluções exatas pelo algoritmo BT foi realizada numa estação de trabalho com memória RAM de 96 GB, 2 processadores *Intel Xeon* modelo *X5690* com 3,47 GHz e sistema operacional *CentOS 6.7*.

Para avaliar o desempenho dos algoritmos, foi elaborado um conjunto de casos de teste composto por grafos simétricos com custos das arestas variando aleatoriamente entre 50 e 500. Em cada vértice existe um passageiro cujo destino é selecionado em sorteio aleatório com distribuição uniforme. A tarifa limite para os passageiros é calculada em função de um percentual da árvore geradora mínima do grafo, variando entre 25% e 50%, conforme o caso teste. A designação das arestas *hov* foi realizada em sorteio aleatório equiprovável sobre as arestas da instância, com base em um percentual do total de arestas previamente fixado  $S$ ,  $S \in \{10\%; 20\%; 25\%\}$ . Os descontos que formam a variável  $c_{ij}$  foram designados em sorteio aleatório equiprovável sobre as arestas *hov* em um percentual previamente fixado  $F$ ,  $F \in \{20\%; 25\%; 30\%; 40\%\}$ . O desconto é ativado quando, ao percorrer uma aresta *hov*, o veículo encontra-se pleno em sua capacidade  $R \in \{3; 4\}$ . Foram criadas 12 classes de instâncias, cada classe possuindo 25 (vinte e cinco) problemas e sendo caracterizada por um par dos valores permitidos para  $S$  e  $F$ . Cada instância é unicamente identificada pelo par  $C/$ - classe - e  $ln$  - dígito identificador de instância relacionado ao valor  $N$ , o número de vértices da instância. Para a realização do experimento, foram considerados 5 problemas de cada classe dos casos teste com  $N = 10$ , totalizando 60 instâncias. As instâncias estão disponíveis em <http://www.dimap.ufrn.br/lae/projetos/CaRS.php>. O experimento consistiu na execução de 100 execuções de cada algoritmo heurístico para cada instância não utilizada no procedimento de ajustes de parâmetro.

Tabela 3 – Parâmetros dos algoritmos *SAop* e *SABL*.

| Parâmetro | Valores possíveis              | <i>SAop</i> | <i>SABL</i> |
|-----------|--------------------------------|-------------|-------------|
| $T_0$     | {1; 20; 40; 60; 80; 100}       | 1           | 60          |
| $T_{min}$ | {0,01; 0,001; 0,0001; 0,00001} | 0,00001     | 0,001       |
| $\alpha$  | [0,75; 0,99]                   | 0,84        | 0,78        |
| $lt$      | {100; 500; 1000; 1500; 2000}   | 2000        | 2000        |

O ajuste de parâmetros foi realizado por meio da ferramenta *IRACE - Iterated Racing for Automatic Algorithm Configuration* [López-Ibáñez et al. 2011], que executa um processo automatizado de afinamento. Foram

selecionadas de forma aleatória, do banco de 60 instâncias, 10 para o conjunto de treinamento. As configurações selecionadas estão na Tabela 1.

Tabela 4 – Resultados do experimento computacional

| Inst. | <i>BT</i> |         | <i>SAop</i> |        |          |            | <i>SABL</i> |        |        |          |            |        |
|-------|-----------|---------|-------------|--------|----------|------------|-------------|--------|--------|----------|------------|--------|
|       | Exata     | T (s)   | Mín.        | Méd.   | $\sigma$ | $\Delta\%$ | Tm (s)      | Mín.   | Méd.   | $\sigma$ | $\Delta\%$ | Tm (s) |
| 01-01 | 590,83    | 127,07  | 590,83      | 620,56 | 27,49    | 0,000      | 0,15        | 590,83 | 591,93 | 6,29     | 0,000      | 0,57   |
| 01-02 | 661,33    | 119,03  | 661,33      | 667,66 | 21,50    | 0,000      | 0,15        | 661,33 | 661,33 | 0,00     | 0,000      | 0,58   |
| 01-03 | 567,93    | 115,68  | 567,93      | 640,38 | 79,72    | 0,000      | 0,15        | 567,93 | 569,27 | 13,34    | 0,000      | 0,60   |
| 01-04 | 837,67    | 120,67  | 837,67      | 851,58 | 19,49    | 0,000      | 0,15        | 837,67 | 837,67 | 0,00     | 0,000      | 0,58   |
| 02-01 | 582,17    | 153,40  | 582,17      | 626,77 | 50,66    | 0,000      | 0,15        | 582,17 | 582,89 | 7,20     | 0,000      | 0,61   |
| 02-02 | 868,50    | 217,99  | 868,50      | 915,87 | 46,31    | 0,000      | 0,15        | 868,50 | 869,59 | 7,37     | 0,000      | 0,64   |
| 02-03 | 653,83    | 109,52  | 653,83      | 694,06 | 38,64    | 0,000      | 0,15        | 653,83 | 653,83 | 0,00     | 0,000      | 0,59   |
| 02-04 | 883,67    | 89,69   | 883,67      | 897,19 | 28,34    | 0,000      | 0,16        | 883,67 | 883,67 | 0,00     | 0,000      | 0,71   |
| 03-01 | 394,73    | 251,09  | 394,73      | 404,32 | 21,28    | 0,000      | 0,16        | 394,73 | 394,73 | 0,00     | 0,000      | 0,74   |
| 03-02 | 548,42    | 199,37  | 548,42      | 659,23 | 79,18    | 0,000      | 0,16        | 548,42 | 555,54 | 28,35    | 0,000      | 0,67   |
| 03-03 | 469,00    | 189,82  | 469,00      | 487,96 | 24,67    | 0,000      | 0,17        | 469,00 | 469,00 | 0,00     | 0,000      | 0,76   |
| 03-04 | 608,05    | 261,16  | 608,47      | 629,47 | 48,81    | 0,069      | 0,17        | 608,05 | 608,05 | 0,00     | 0,000      | 0,81   |
| 03-05 | 579,17    | 176,76  | 579,17      | 580,43 | 8,58     | 0,000      | 0,16        | 579,17 | 579,17 | 0,00     | 0,000      | 0,66   |
| 04-01 | 443,42    | 122,51  | 443,42      | 461,04 | 27,47    | 0,000      | 0,15        | 443,42 | 443,42 | 0,00     | 0,000      | 0,62   |
| 04-02 | 386,13    | 149,79  | 386,13      | 488,27 | 78,06    | 0,000      | 0,16        | 386,13 | 401,79 | 44,77    | 0,000      | 0,68   |
| 04-03 | 389,75    | 175,35  | 389,75      | 465,20 | 64,33    | 0,000      | 0,17        | 389,75 | 395,57 | 17,61    | 0,000      | 0,78   |
| 04-04 | 333,25    | 150,15  | 354,75      | 401,17 | 85,90    | 6,452      | 0,16        | 333,25 | 333,25 | 0,00     | 0,000      | 0,70   |
| 04-05 | 692,40    | 140,47  | 692,40      | 769,95 | 91,67    | 0,000      | 0,16        | 692,40 | 692,40 | 0,00     | 0,000      | 0,71   |
| 05-01 | 454,67    | 128,24  | 454,67      | 457,53 | 16,43    | 0,000      | 0,15        | 454,67 | 454,67 | 0,00     | 0,000      | 0,57   |
| 05-03 | 695,42    | 221,45  | 695,42      | 718,37 | 39,17    | 0,000      | 0,15        | 695,42 | 695,81 | 1,47     | 0,000      | 0,57   |
| 05-04 | 521,92    | 270,04  | 557,08      | 572,41 | 20,62    | 6,738      | 0,15        | 521,92 | 521,92 | 0,00     | 0,000      | 0,60   |
| 05-05 | 585,67    | 232,58  | 585,67      | 586,90 | 7,04     | 0,000      | 0,15        | 585,67 | 585,67 | 0,00     | 0,000      | 0,62   |
| 06-02 | 636,75    | 215,24  | 641,92      | 669,32 | 26,33    | 0,811      | 0,15        | 641,92 | 642,69 | 4,40     | 0,811      | 0,60   |
| 06-04 | 462,67    | 153,14  | 462,67      | 478,55 | 28,01    | 0,000      | 0,16        | 462,67 | 462,67 | 0,00     | 0,000      | 0,67   |
| 06-05 | 441,58    | 162,53  | 453,33      | 488,03 | 42,79    | 2,661      | 0,16        | 441,58 | 443,23 | 4,10     | 0,000      | 0,63   |
| 07-01 | 309,33    | 277,01  | 309,33      | 332,14 | 28,42    | 0,000      | 0,15        | 309,33 | 309,33 | 0,00     | 0,000      | 0,66   |
| 07-04 | 510,39    | 363,35  | 510,39      | 514,85 | 13,95    | 0,000      | 0,15        | 510,39 | 510,39 | 0,00     | 0,000      | 0,64   |
| 07-05 | 565,88    | 548,27  | 567,63      | 575,93 | 11,69    | 0,311      | 0,14        | 565,88 | 565,88 | 0,00     | 0,000      | 0,62   |
| 08-01 | 465,60    | 591,39  | 465,60      | 482,70 | 22,89    | 0,000      | 0,15        | 465,60 | 465,72 | 1,16     | 0,000      | 0,72   |
| 08-02 | 314,96    | 315,67  | 314,96      | 324,22 | 13,63    | 0,000      | 0,17        | 314,96 | 315,87 | 3,00     | 0,000      | 0,80   |
| 08-03 | 452,86    | 392,94  | 452,86      | 460,75 | 13,94    | 0,000      | 0,15        | 452,86 | 452,99 | 0,65     | 0,000      | 0,65   |
| 08-05 | 386,37    | 351,80  | 386,37      | 390,24 | 8,40     | 0,000      | 0,15        | 386,37 | 386,37 | 0,00     | 0,000      | 0,70   |
| 09-01 | 445,03    | 362,05  | 445,03      | 459,79 | 14,81    | 0,000      | 0,16        | 445,03 | 445,03 | 0,00     | 0,000      | 0,67   |
| 09-02 | 451,40    | 319,41  | 472,73      | 495,07 | 19,47    | 4,726      | 0,15        | 451,40 | 455,45 | 8,41     | 0,000      | 0,64   |
| 09-03 | 420,20    | 452,86  | 420,20      | 431,79 | 14,45    | 0,000      | 0,15        | 420,20 | 420,32 | 0,91     | 0,000      | 0,61   |
| 09-04 | 489,93    | 378,41  | 489,93      | 500,42 | 15,04    | 0,000      | 0,14        | 489,93 | 489,93 | 0,00     | 0,000      | 0,56   |
| 09-05 | 350,17    | 327,36  | 350,17      | 352,83 | 9,41     | 0,000      | 0,15        | 350,17 | 350,17 | 0,00     | 0,000      | 0,59   |
| 10-01 | 504,43    | 401,92  | 508,60      | 525,16 | 18,32    | 0,826      | 0,14        | 504,43 | 504,43 | 0,00     | 0,000      | 0,55   |
| 10-02 | 429,23    | 317,58  | 429,23      | 472,15 | 33,94    | 0,000      | 0,14        | 429,23 | 429,66 | 2,99     | 0,000      | 0,53   |
| 10-03 | 608,17    | 418,62  | 608,17      | 620,00 | 16,53    | 0,000      | 0,14        | 608,17 | 608,17 | 0,00     | 0,000      | 0,50   |
| 10-05 | 427,30    | 389,76  | 427,30      | 430,51 | 8,32     | 0,000      | 0,15        | 427,30 | 427,30 | 0,00     | 0,000      | 0,57   |
| 11-01 | 466,42    | 643,42  | 466,42      | 485,76 | 21,22    | 0,000      | 0,15        | 466,42 | 466,45 | 0,36     | 0,000      | 0,61   |
| 11-03 | 417,22    | 1084,24 | 417,22      | 434,21 | 13,06    | 0,000      | 0,15        | 417,22 | 417,27 | 0,28     | 0,000      | 0,65   |
| 11-04 | 456,70    | 1311,22 | 456,70      | 473,97 | 20,11    | 0,000      | 0,15        | 456,70 | 456,70 | 0,00     | 0,000      | 0,66   |
| 11-05 | 458,28    | 749,36  | 458,28      | 469,21 | 11,36    | 0,000      | 0,14        | 458,28 | 458,28 | 0,00     | 0,000      | 0,62   |
| 12-01 | 428,07    | 919,52  | 428,07      | 444,28 | 15,43    | 0,000      | 0,15        | 428,07 | 428,07 | 0,00     | 0,000      | 0,64   |
| 12-02 | 437,07    | 722,92  | 437,07      | 447,57 | 15,12    | 0,000      | 0,15        | 437,07 | 437,07 | 0,00     | 0,000      | 0,61   |
| 12-03 | 504,10    | 1028,79 | 504,10      | 511,91 | 8,46     | 0,000      | 0,14        | 504,10 | 504,10 | 0,00     | 0,000      | 0,60   |
| 12-04 | 414,78    | 643,84  | 414,78      | 422,63 | 9,75     | 0,000      | 0,14        | 414,78 | 414,78 | 0,00     | 0,000      | 0,61   |
| 12-05 | 434,60    | 754,31  | 441,52      | 446,70 | 13,31    | 1,592      | 0,14        | 434,60 | 434,60 | 0,00     | 0,000      | 0,62   |

A Tabela 2 apresenta os resultados do algoritmo exato *BT* e dos algoritmos heurísticos *SAop* e *SABL* para as instâncias do banco excetuando-se aquelas empregadas no passo de calibração de parâmetros, totalizando 50 casos de teste. Cada linha mostra os resultados da instância identificada na primeira coluna. São mostrados os valores da solução ótima (Exata) e do tempo execução do algoritmo *BT* em segundos (T). As colunas relativas aos algoritmos heurísticos mostram o valor da menor solução encontrada (Mín), a média das melhores soluções (Méd), o desvio padrão ( $\sigma$ ), o afastamento percentual da melhor solução para a solução ótima ( $\Delta\%$ ) e o tempo médio de execução em segundos (Tm).

Os resultados exibidos na Tabela 2 mostram que, de modo geral, o algoritmo *SABL* apresentou um desempenho superior. Ele encontrou a solução ótima de 49 das 50 instâncias, enquanto o algoritmo *SAop* conseguiu alcançar a solução ótima em 41 instâncias. O tempo médio de execução do *SAop* foi 0,15s e 0,64s para o *SABL*. O *SABL* também apresentou os menores valores médios e a menor variância dos resultados.

## REFERÊNCIAS

- [1]. Amey, A., Attanucci, J. & Mishalani, R. (2011). Real-Time Ridesharing – The Opportunities and Challenges of Utilizing Mobile Phone Technology to Improve Rideshare Services. TRB Annual Meeting, 1-17.
- [2]. Balas, E., (2004). The prize collecting traveling salesman problem and its applications, in: G. Gutin e A. Punnen, editors, The Traveling Salesman Problem and Its Variations, Combinatorial Optimization, 12, Springer US, 663-695.
- [3]. Calheiros, Z. S. A. (2015). Algoritmos Evolucionários na Solução do Caixeiro Viajante com Caronas, Monografia de Graduação, Universidade Federal do Rio Grande do Norte.
- [4]. Dailey, D. J., Lose, D. & Meyers, D. (1999). Seattle smart traveler: dynamic ridematching on the world wide web. Transportation Research Part C 7(1):17-32.
- [5]. Erdogan, G., Cordeau, J-F. & Laporte, G. (2010). The Attractive Traveling Salesman Problem, European Journal of Operational Research 203:59-69

## 5. CONCLUSÃO

Este te sugere para o modelo maior complexidade que o PCV com um maior número de decisões livresrabalho apresentou um novo problema de roteamento, chamado Problema do Caixeiro Viajante com Passageiros e Lotação (PCV-PL). Embora tenha como base o PCV, o PCV-PL possui características qu. A dicotomia entre priorizar a rota menos custosa ou buscar a máxima coleta de passageiros confere ao problema uma interessante característica de *trade off*. Observe-se que a coleta de passageiros é associada não só à sequência de localidades da rota quanto a restrições de tarifa dos próprios passageiros. Por outro lado, a capacidade do carro interfere na coleta e, conseqüentemente, também na rota ótima.

Como atividades futuras será realizada uma modelagem matemática, solução da mesma para diversas instâncias e desenvolvimento de novos algoritmos heurísticos.

## AGRADECIMENTOS

Este trabalho foi financiado parcialmente pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), projetos 301845/2013-1 e 308062/2014-0.

- [6]. Feillet, D., Dejax, P. & Gendreau, M. (2005). Travelling salesman problems with profits, Transportation Science, 39:188-205.
- [7]. Furuhata, M., Dessouky, M., Ordonez, F., Brunet, M., Wang, X. & Koenig, S. (2013). Ridesharing: The state-of-the-art and future directions. Transportation Research Part B 57:28-46.
- [8]. Garey, M. R. e Jonhson, D. S. (1979). Computers and intractability: a guide to the theory of NP-completeness, W.H. Freeman and Co.
- [9]. Gendreau, M., Laporte, G. & Semet, F. (1998). A tabu search heuristic for the undirected selective travelling salesman problem, European Journal of Operational Research 106:539-545.
- [10]. Gribkovskaia, I., Laporte, G. & Shyshou A. (2008). The single vehicle routing problem with deliveries and selective pickups, Computers & Operations Research 35(9):2908-2924.
- [11]. Lin, S., Kernighan, B. W. (1973). An Effective Heuristic Algorithm for the Traveling-Salesman Problem, Operations Research, 21.
- [12]. López-Ibáñez, M., Dubois-Lacoste, J., Stützle, T. & Birattari, M. (2011). The irace package,

Iterated Race for Automatic Algorithm Configuration, Technical Report TR/IRIDIA/2011-004, IRIDIA, Université Libre de Bruxelles, Belgium.

[13]. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21(6):1087-1092.

[14]. Parragh, S. N., Doerner, K. F. & Hartl, R. F. (2008a). A survey on pickup and delivery problems - Part I: Transportation between customers and depot, *Journal für Betriebswirtschaft* 58(1): 21-51.

[15]. Parragh, S. N., Doerner, K. F. & Hartl, R. F. (2008b). A survey on pickup and delivery problems

[18].

Part II: Transportation between pickup and delivery locations, *Journal für Betriebswirtschaft* 58(2): 81-117.

[16]. Vansteenwegen, P., Souffriau, W. & Van Oudheusden, D. (2011). The orienteering problem: A survey, *European Journal of Operational Research* 209:1-10.

[17]. Wang, X., Dessouky, M. & Ordóñez, F. (2013). A Pickup and Delivery Problem for Ridesharing Considering Congestion. *Transportation Letters: The International Journal of Transportation Research*.

# Capítulo 8

## ALGORITMO GENÉTICO PARA O PROBLEMA DE ROTEAMENTO EM ARCOS CAPACITADO E ABERTO

*Rafael Kendy Arakaki*

*Fábio Luiz Usberti*

**Resumo:** Problemas de roteamento em arcos têm por objetivo determinar rotas de custo mínimo que visitam um subconjunto de arcos de um grafo, com uma ou mais restrições adicionais. A solução desses problemas remete à diminuição de custos logísticos, melhorando a competitividade das empresas. O Problema de Roteamento em Arcos Capacitado e Aberto (OCARP Open Capacitated Arc Routing Problem) é um problema de otimização combinatória NP-difícil com aplicações práticas, como o problema de roteamento de leituristas. Este trabalho propõe uma nova metodologia de algoritmos genéticos para tratar o OCARP. São apresentados experimentos computacionais com instâncias da literatura.

**Palavras chave.** roteamento em arcos, roteamento de leituristas, algoritmo genético, otimização combinatória, programação linear inteira.



## 1 INTRODUÇÃO

Os problemas de roteamento podem ser subdivididos em duas grandes áreas: os problemas de roteamento em nós e os problemas de roteamento em arcos (ARPs Arc Routing Problems). Os problemas de roteamento em nós, como o clássico Problema do Caixeiro Viajante (TSP Travelling Salesman Problem), objetivam a criação de rota(s) que visita(m) um subconjunto de nós. Por outro lado, os ARPs objetivam a criação de rota(s) que visita(m) um subconjunto de arestas. Durante a última década, os problemas de roteamento em arcos têm sido uma área de pesquisa bastante ativa [Wøhlk, 2008].

O problema de roteamento em arcos capacitado e aberto (OCARP Open Capacitated Arc Routing Problem) é um problema de otimização combinatória definido em um grafo conexo não-direcionado com custos e demandas não-negativas nas arestas. É disponibilizada uma frota de veículos com capacidade limitada para atender as demandas das arestas. O objetivo do OCARP é encontrar um conjunto de rotas que atenda às demandas das arestas e possua custo mínimo [Usberti et al., 2011].

O OCARP é um problema NP-difícil, portanto se  $P \neq NP$  não existe um algoritmo de tempo polinomial para resolvê-lo [Usberti, 2012]. Trata-se de um problema de otimização combinatória que possui diversas aplicações práticas.

Uma das aplicações práticas do OCARP é um problema real de roteamento de leituristas. Empresas distribuidoras de energia elétrica, água e gás possuem colaboradores que registram o consumo de clientes dispersos geograficamente. O roteamento desses funcionários apresenta-se como um problema de difícil solução, o qual consiste em determinar as rotas que os leituristas devem percorrer para realizar as leituras de consumo dos clientes [Usberti et al., 2008]. Outra aplicação do OCARP é o problema de determinação do caminho de corte, o qual consiste em guiar o corte de chapas metálicas em peças de formas poligonais utilizadas na indústria metal-mecânica [da Silva, 2016].

O presente trabalho propõe uma metodologia de solução heurística baseada em algoritmos genéticos para o OCARP. O método desenvolvido melhorou

substancialmente as soluções conhecidas para várias instâncias da literatura.

Este trabalho está dividido em seis seções. A Seção 2 apresenta a definição e a formulação para o problema de roteamento em arcos capacitado e aberto. Adicionalmente, são revisados os principais resultados da literatura relacionada ao OCARP. A Seção 3 descreve a metaheurística de algoritmos genéticos proposta. A Seção 4 apresenta os resultados de experimentos computacionais da heurística. Por fim, a Seção 5 apresenta as conclusões.

## 2 PROBLEMA DE ROTEAMENTO EM ARCOS CAPACITADO E ABERTO

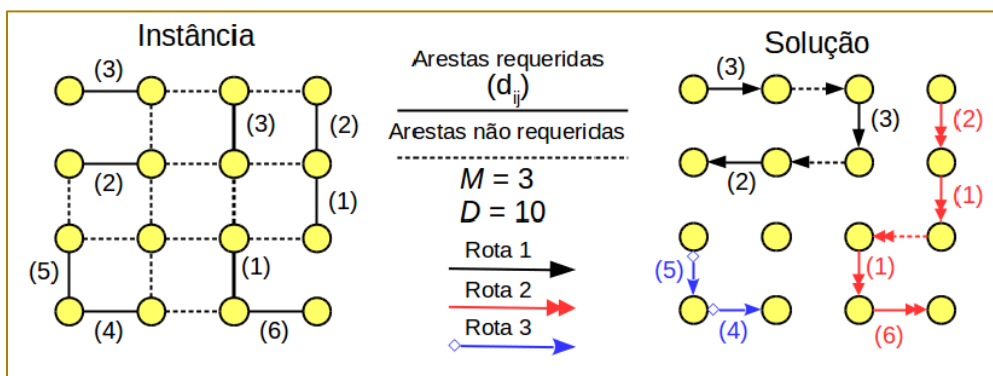
O Problema de Roteamento em Arcos Capacitado e Aberto (OCARP), proposto por Usberti et al. [2011], é definido a seguir. Seja  $G(V, E)$  um grafo conexo não-direcionado com custos não-negativos  $c_{ij}$  e demandas não-negativas  $d_{ij}$  associados a cada aresta. Todas as arestas que possuem demandas positivas ( $d_{ij} > 0$ ) formam o conjunto de arestas requeridas  $ER \subseteq E$ . Cada aresta requerida deve ser atendida apenas uma vez por um único veículo. Entretanto, as arestas podem ser visitadas múltiplas vezes, por um ou mais veículos. Uma frota de  $M$  veículos idênticos com capacidade limitada  $D$  é disponibilizada. Ao percorrer o grafo, um veículo pode escolher entre atender uma aresta ou apenas visitá-la. No caso de atendimento a sua capacidade é reduzida pela demanda da aresta atendida.

O OCARP pode ser considerado uma variante do CARP (Capacitated Arc Routing Problem CARP), porém com a diferença de que o OCARP considera tanto rotas abertas como fechadas. Uma rota aberta possui nós distintos como nó inicial e nó final, enquanto uma rota fechada possui o mesmo nó inicial e final. Uma solução OCARP consiste em um conjunto de  $M$  rotas, uma para cada veículo, para atender todas as arestas requeridas sem exceder a capacidade de nenhum veículo (Figura 1). O objetivo do OCARP é minimizar o custo desse conjunto de rotas.

Do ponto de vista teórico, OCARP e CARP são ambos problemas NP-difíceis, sendo possível a redução em tempo polinomial de um problema para o outro como mostrado por Usberti [2012]. Apesar das semelhanças entre CARP e OCARP, existem diferenças importantes que dificultam a adaptação de algoritmos do CARP para se resolver o OCARP. A maioria das metodologias para o CARP considera o número de veículos  $M$  como uma variável a ser otimizada, entretanto

no OCARP o número de veículos  $M$  é um parâmetro fixo. Essa diferença é importante porque torna a geração de uma solução viável para o OCARP mais difícil, uma vez que se o número de veículos e a capacidade dos veículos forem baixos o suficiente, em relação aos valores das demandas, implica-se em um problema NP-difícil de empacotamento (BPP Bin-Packing Problem) para se obter qualquer solução viável.

Figura 1: OCARP: uma instância e uma solução viável.



## 2.1 MODELO MATEMÁTICO

Uma formulação para o OCARP proposta originalmente por Usberti et al. [2011] é apresentada a seguir.

Apesar do grafo  $G$  ser não-direcionado, é conveniente representar soluções OCARP com arcos direcionados para dar a orientação e ordem em que as arestas são visitadas. Desse modo são definidos quatro conjuntos de variáveis:

$x_{ij}^k = 1$  se a rota  $k$  visita o arco  $(i,j)$ ,  $x_{ij}^k = 0$  do contrário;

$l_{ij}^k = 1$  se a rota  $k$  visita o arco  $(i,j)$ ,  $l_{ij}^k = 0$  do contrário;

$\alpha_{ij}^k = 1$  se a rota  $k$  visita o arco  $(i,j)$ ,  $\alpha_{ij}^k = 0$  do contrário;

$u_s^k$ ,  $v_s^k$  e  $w_s^k$  ( $S \subseteq V$ ) são variáveis auxiliares utilizadas no conjunto de restrições (7) para eliminação de subciclos ilegais.

A função objetivo (1) minimiza o custo total, dado pela soma dos custos das arestas pertencentes a cada rota da solução. Seja  $N$

( $i$ ) o conjunto de vértices vizinhos ao vértice  $i$  em  $G$ , a restrição (2) representa a continuidade das rotas e também impõe  $\alpha_i^k = 1$  se o vértice  $i$  é início da rota  $k$ . A restrição (3) não permite que haja mais de um vértice como início de rota. A restrição

(4) estabelece que arcos atendidos precisam ser visitados. A restrição (5) força que cada aresta requerida (representada por dois arcos) seja atendida uma única vez por uma única rota. A restrição (6) representa o limite de capacidade dos veículos.

O conjunto de restrições (7) não permite subciclos ilegais na solução. Para uma certa rota  $k$ , a rota é considerada ilegal se existe algum subconjunto de nós  $S$  tal que  $v_s^k = u_s^k = w_s^k = 1$ .

A variável  $u_s^k$  identifica se há um ciclo em  $S$ , a variável  $v_s^k$  identifica se não há arestas de corte em  $(S, \bar{S})$  e a variável  $w_s^k$  identifica se não há um vértice inicial em  $S$ . Um exemplo de subciclo ilegal é apresentado na Figura 2. O conjunto de restrições (7) é de ordem exponencial, portanto na prática essas restrições são adicionadas ao modelo à medida em que são necessárias.

Lema 2.1. O conjunto de restrições (7) elimina soluções com rotas ilegais (subciclos ilegais) mas não elimina solução viáveis (rotas legais).

(OCARP)

$$\text{MIN} \sum_{k=1}^M \sum_{(i,j) \in E} c_{ij} x_{ij}^k \quad (1)$$

s.a.

$$\sum_{j \in N(i)} x_{ij}^k - x_{ji}^k \leq \alpha_i^k \quad (i \in V, k \in \{1, \dots, M\}) \quad (2)$$

$$\sum_{i \in V} \alpha_i^k \leq 1 \quad (k \in \{1, \dots, M\}) \quad (3)$$

$$x_{ij}^k \geq l_{ij}^k \quad ((i, j) \in E_R, k \in \{1, \dots, M\}) \quad (4)$$

$$\sum_{k=1}^M (l_{ij}^k + l_{ji}^k) = 1 \quad ((i, j) \in E_R) \quad (5)$$

$$\sum_{(i,j) \in E_R} d_{ij} l_{ij}^k \leq D \quad k \in \{1, \dots, M\}) \quad (6)$$

$$\left. \begin{aligned} \sum_{(i,j) \in (S,S)} x_{ij}^k - |S|^2 u_S^k &\leq |S| - 1 \\ \sum_{(i,j) \in (S,\hat{S})} x_{ij}^k + v_S^k &\geq 1 \\ \sum_{i \in S} \alpha_i^k + w_S^k &\geq 1 \\ u_S^k + v_S^k + w_S^k &\leq 2 \end{aligned} \right\} \quad (S \subseteq V, \hat{S} = V \setminus S, k \in \{1, \dots, M\}) \quad (7)$$

$$x_{ij}^k \in \{0, 1\}, \quad ((i, j) \in E, k \in \{1, \dots, M\}) \quad (8)$$

$$l_{ij}^k \in \{0, 1\}, \quad ((i, j) \in E_R, k \in \{1, \dots, M\}) \quad (9)$$

$$\alpha_i^k \in \{0, 1\}, \quad (i \in V, k \in \{1, \dots, M\}) \quad (10)$$

$$u_S^k, v_S^k, w_S^k \in \{0, 1\}, \quad (k \in \{1, \dots, M\}, S \subseteq V) \quad (11)$$

Demonstração. Primeiro provaremos que o modelo proíbe uma rota ilegal. Suponha uma rota  $k$  ilegal e um grafo  $G'$  induzido pelas arestas que pertencem à rota  $k$ . Se a rota  $k$  é ilegal então o grafo  $G'$  não é conexo. Pela restrição (3) no máximo uma componente de  $G'$  terá um vértice de início de rota ( $\alpha_i^k = 1$ ). Excluindo-se a componente que tem o vértice inicial, há pelo menos mais uma. Essa(s) outra(s) componente(s) de  $G'$  deverão formar subciclo(s) para manter o equilíbrio da

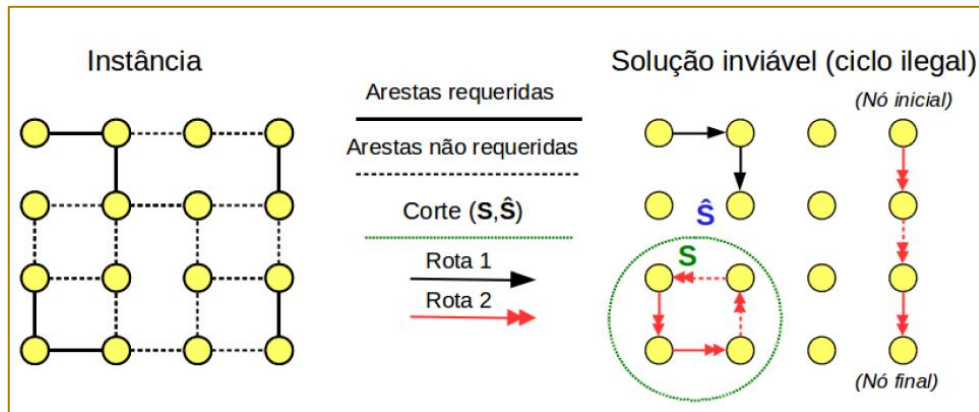
restrição (2). Seja  $S_0$  o conjunto de vértices de uma dessas componentes que não contém o vértice inicial. Então é possível verificar que:

$u_{S'}^k = 1$ , pois  $S'$  corresponde a um subciclo.

$v_{S'}^k = 1$ , porque  $S'$  é uma componente em  $G'$  (não há arestas de corte).

$w_{S'}^k = 1$ , não há nó inicial em  $S'$ .

Figura 2: Uma instância OCARP (à esquerda) e uma solução inviável por causa do subciclo ilegal formado na rota 2 (à direita).



Temos portanto que qualquer rota  $k$  que gere mais de uma componente em um grafo  $G'$  induzido por suas arestas é ilegal no modelo (subciclo ilegal). Para completar a prova, mostraremos que o conjunto de restrições (7) não proíbe rotas legais, ou seja, rotas que gerem somente uma componente em um grafo  $G'$  induzido pelas suas arestas.

Suponha uma rota legal  $k$  e o conjunto  $R$  formado pelos vértices que são visitados por esta rota (vértices de  $G'$ ). Temos então três casos possíveis para um subconjunto de vértices  $S \subseteq V$  em relação a  $R$ :

Se  $R \subseteq S$ , um dos vértices da rota deve ser o inicial (em caso de rota fechada qualquer vértice pode ser inicial), portanto  $w_S^k = 0$ .

Se  $R \not\subseteq S$  e  $R \cap S \neq \emptyset$ , então há pelo menos uma aresta de corte, que liga um elemento de  $R \cap S$  com um elemento de  $R \setminus S$ , portanto  $v_S^k = 0$ .

Se  $R \cap S = \emptyset$ , não há arestas  $(i, j) \in (S, S)$  na rota, portanto  $u_S^k = 0$ .

Como conclusão temos que o conjunto de restrições (7) proíbe exatamente as rotas que não são conexas (rotas ilegais) e em conjunto com as outras restrições do modelo define um conjunto de rotas factíveis para o OCARP.

## 2.2 REVISÃO DA LITERATURA

Os experimentos com a formulação original (Seção 2.1) reportaram soluções ótimas para

várias instâncias do conjunto ogdb (7-27 nós, 11-55 arestas), porém não alcançou bom desempenho para os conjuntos de instâncias maiores: oval (24-50 nós, 34-97 arestas) e oegl (77-140 nós, 98-190 arestas). Segundo os autores a principal deficiência nesse método apresentou-se ser os limitantes duais fracos para as maiores instâncias [Usberti et al., 2011].

Mais tarde outro método de solução exata para o OCARP foi descrito por Usberti et al. [2012]. Trata-se de um algoritmo branch-and-bound que se vale da relação entre o OCARP e outro problema chamado Capacity and Degree Constrained Minimum Spanning Forest (CDCMSF). Esse método aprimorou os limitantes duais conhecidos, porém para o conjunto de instâncias mais difícil oegl ainda foi reportada uma média de desvio de otimalidade substancial (acima de 15%), evidenciando assim a necessidade por métodos mais eficientes.

## 3 ALGORITMO GENÉTICO

Muitos problemas combinatórios NP-difíceis são resolvidos através de métodos heurísticos, uma vez que a solução exata desses problemas requer tempos exponenciais, exceto se  $P = NP$  [Garey e Johnson, 1978]. Uma das metaheurísticas mais utilizadas para a solução de problemas de otimização são algoritmos genéticos [Goldberg, 1988].

Algoritmos genéticos são algoritmos inspirados por elementos da evolução das

espécies como seleção, herança, mutação e cruzamento [Holland, 1991]. Nesta seção é descrita uma metaheurística de algoritmos genéticos proposta para a solução do OCARP.

### 3.1 DESCRIÇÃO DA METAHEURÍSTICA

A heurística apresentada nesta seção trata-se de um algoritmo genético híbrido, além das operações clássicas (cruzamento, seleção, população inicial) considera também os seguintes componentes: (i) rotina de

factibilização, (ii) busca local, (iii) rotina de reinício de população. O Algoritmo 1 mostra o pseudo-código do algoritmo genético.

A rotina de factibilização é utilizada para se obter uma solução factível a partir da codificação de um indivíduo. A rotina de busca local é utilizada para aprimorar as soluções geradas ao longo do processo. Por fim, a rotina de reinício de população justifica-se para evitar a convergência prematura da população, uma vez que não é considerado nenhum operador de mutação.

---

#### Algorithm 1: Algoritmo Genético

---

```

início
  Inicialização da matriz de distâncias.
  Gerar a população inicial.
  enquanto limite de tempo não for excedido faça
    Selecionar pais para cruzamento.
    para um par de pais selecionados (pai1,pai2) faça
       $c = \text{cruzamento}(pai1, pai2)$ 
       $s = \text{factibilização}(c)$ 
       $s' = \text{busca\_local}(s)$ 
      Adiciona a solução  $s'$  à população se factível.
    Selecionar nova população.
    Reinício de população a cada  $k$  gerações.
fim

```

---

### 3.2 INICIALIZAÇÃO DA MATRIZ DE DISTÂNCIAS

Essa etapa computa uma matriz de distâncias que é utilizada em diversas componentes da heurística. Seja  $E_R$  o conjunto de arestas requeridas e seja  $A_R$  o conjunto de arcos requeridos, onde cada aresta requerida  $\{i, j\} \in E_R$  possui dois arcos requeridos correspondentes  $(i, j), (j, i) \in A_R$ .

É gerada uma matriz  $D$  de dimensão  $|A_R| \times |A_R|$ . Cada entrada  $D[e, f]$  dessa matriz é definida

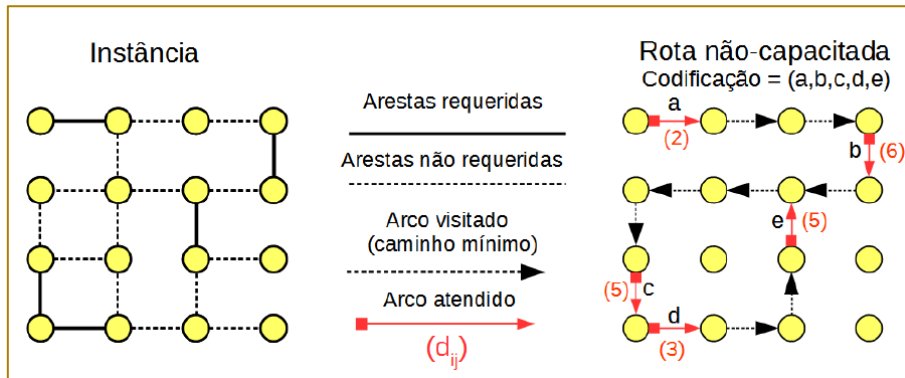
como o custo do caminho mínimo a partir do vértice fim do arco  $e \in A_R$  até o vértice início do arco  $f \in A_R$ . Desse modo a distância entre arcos requeridos pode ser recuperada em tempo  $O(1)$ .

### 3.3 CODIFICAÇÃO

O algoritmo genético utiliza como codificação de uma solução uma sequência de  $|E_R|$  arcos requeridos. Desse modo, mantém-se na codificação a ordem em que os arcos requeridos são atendidos, sem a necessidade de explicitar as arestas não requeridas uma vez que estas estão inseridas implicitamente nos caminhos mínimos entre cada par de arcos requeridos.

A codificação não possui nenhum delimitador entre as rotas, de modo que a codificação pode ser vista como uma única rota não-capacitada, como mostra a Figura 3. A solução OCARP correspondente à codificação somente é obtida após a codificação passar pela etapa de factibilização (Seção 3.8).

Figura 3: Uma instância e uma codificação com sua rota não-capacitada.



### 3.4 FITNESS

A função de fitness de uma solução  $s$  de custo  $c(s)$  é definida como  $1/(c(s) - lb_0)$ , onde  $lb_0$  é um limitante dual trivial para o problema: a soma do custo de todas as arestas requeridas. Se for encontrada uma solução  $s$  onde  $c(s) = lb_0$ , então o algoritmo é encerrado pois foi encontrada uma solução ótima.

### 3.5 GERAÇÃO DA POPULAÇÃO INICIAL

A inicialização da população é dividida em duas etapas. Na primeira etapa resolve-se o problema de roteamento não-capacitado e na segunda etapa procura-se particionar a rota não-capacitada em sub-rotas capacitadas, gerando assim soluções factíveis para o OCARP.

Na primeira etapa a rota não-capacitada é construída por via da heurística Nearest Neighbour, onde a rota é construída de maneira gulosa e iterativa até inserir todos os elementos requeridos. Por fim, a rota é otimizada pelo método de busca local 2-opt [Croes, 1957].

A segunda etapa é composta pela rotina de factibilização (Seção 3.8). Caso o indivíduo gerado seja infactível, outro indivíduo é gerado até que toda a população seja preenchida com indivíduos factíveis.

### 3.6 SELEÇÃO

São realizados  $P - 1$  cruzamentos a cada geração, tal que  $P$  é o tamanho da população

e cada cruzamento requer a seleção de dois indivíduos.

A seleção dos indivíduos é realizada através da técnica Stochastic Universal Selection de Baker [1986]. Nessa técnica, os indivíduos são selecionados em quantidade proporcional à razão de seu fitness sobre a soma de fitness de toda a população.

A nova geração é formada pelo melhor indivíduo e pelos  $P - 1$  indivíduos filhos da geração corrente. Se houver  $C$  cruzamentos que geraram indivíduos filhos infactíveis, então são escolhidos aleatoriamente  $C$  indivíduos da geração atual para serem inseridos na nova geração.

### 3.7 CRUZAMENTOS

Foi adotado o operador de cruzamento  $C1$  de Reeves [1996], onde um gene  $g$  é escolhido aleatoriamente e copia-se para o filho os genes do início até  $g$  do primeiro pai. Em seguida são copiados para o filho os genes que faltam na ordem em que aparecem no segundo pai (Figura 4).

### 3.8 FACTIBILIZAÇÃO

O objetivo da rotina de factibilização é particionar a rota não-capacitada em  $M$  sub-rotas capacitadas (solução OCARP). Para isso é utilizado o algoritmo Split de Ulusoy [1985]. O algoritmo gera  $M$  sub-rotas que não excedem a capacidade dos veículos e que atendem os arcos requeridos na mesma ordem definida pela codificação (Figura 5).

Figura 4: Operador de cruzamento C1.

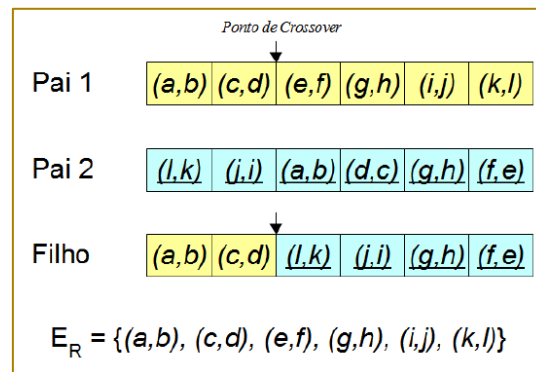
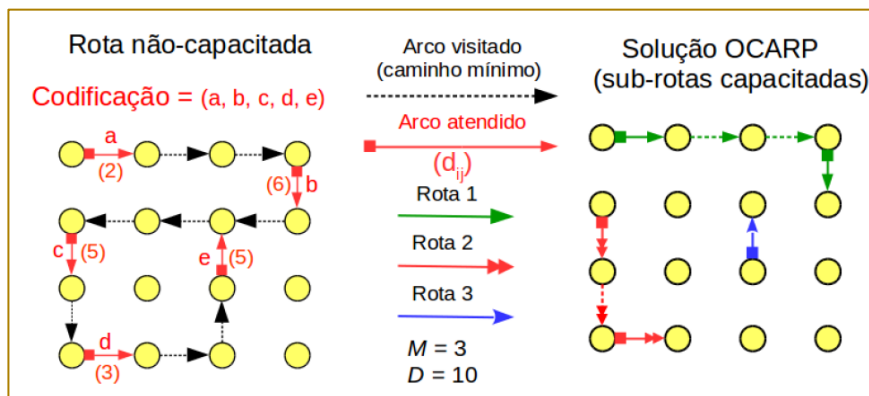


Figura 5: Uma codificação e uma solução correspondente dada pelo algoritmo Split.



Devido ao número fixo de veículos e das capacidades limitadas, não é possível garantir que a solução gerada pelo algoritmo de Ulusoy seja factível. Nesse caso é adotada uma estratégia para tentar tornar a solução factível. Dentre as rotas geradas pelo algoritmo, determinase aquela com a menor capacidade residual adicionando-a a uma nova solução parcial. Em seguida, as arestas requeridas faltantes são adicionadas à solução parcial pelo método construtivo utilizado na inicialização da população (Seção 3.5). Finalmente, chama-se recursivamente a rotina de factibilização com essa nova solução. Se após  $M$  recursões não for encontrada uma solução factível, o indivíduo é declarado inviável.

### 3.9 BUSCA LOCAL

Algoritmos de busca local realizam pequenas modificações em uma solução em busca de soluções similares (vizinhas) de melhor custo.

As soluções geradas no cruzamento possuem uma probabilidade  $p$  de serem submetidas à rotina de busca local.

A busca local utiliza uma estratégia de desconstrução/reconstrução de pares de rotas. A reconstrução utiliza uma heurística construtiva gulosa Nearest Neighbour (Seção 3.5) para gerar uma única rota que atende os arcos do par de rotas desconstruído. Em seguida, essa rota única é particionada em duas rotas pela rotina de factibilização (Seção 3.8). As alterações nas rotas da solução somente são efetivadas se diminuírem o custo da solução.

### 3.10 REINÍCIO DE POPULAÇÃO

A rotina de reinício de população é executada repetidamente ao longo da heurística em um intervalo de  $k$  gerações. Essa rotina tem o objetivo de evitar a convergência prematura da população.

A cada execução da rotina, metade da população atual é substituída por novas soluções. Os indivíduos substituídos são escolhidos aleatoriamente, entretanto a melhor solução nunca é substituída (seleção elitista). As novas soluções são geradas pelo mesmo processo utilizado pela rotina de inicialização da população.

#### 4 EXPERIMENTOS COMPUTACIONAIS

Esta seção apresenta o experimento computacional que tem por objetivo comparar os custos das soluções obtidos pelo algoritmo genético e pela heurística RPS ER apresentada por Usberti et al. [2011].

A máquina utilizada no experimento compõe-se de processador Intel Xeon X3430 2.4 GHz, 8 GB de memória RAM e sistema operacional Linux 64-bit. O algoritmo genético foi executado uma única vez com limitação de uma hora de execução para cada instância.

Os parâmetros definidos para o algoritmo genético são apresentados na Tabela 1.

Os resultados são apresentados nas Tabelas 2 e 3 para os respectivos conjuntos de instâncias: oval (24-50 nós, 34-97 arestas) e oegl (77-140 nós, 98-190 arestas) [Benavent et al., 1991; Li e Eglese, 1995]. Para cada instância, foram considerados três valores distintos para o parâmetro  $M$  (número de veículos):  $M^*$ ,  $M^* + 1$  e  $M^* + 2$ , onde  $M^*$  é o número mínimo de veículos necessários para se obter uma solução factível.

O valor de  $LP_1$  corresponde à melhor solução obtida pela heurística *RPS\_ER* (reportada pelos autores) e  $LP_2$  à melhor solução obtida pelo algoritmo genético. Dados os valores de  $LP_1$  e  $LP_2$  é calculado um valor  $\Delta LP$  (%) que corresponde à diferença proporcional entre os dois limitantes. O tempo de processamento que o algoritmo genético levou para obter o limitante  $LP_2$  é dado por CPU.

Tabela 1: Parâmetros do algoritmo genético

| Parâmetro | Descrição                                     | Valor |
|-----------|---|-------|
| $P$       | Tamanho da população                          | 20    |
| $p$       | Probabilidade de busca local                  | 0.1   |
| $k$       | Intervalo de reinício da população (gerações) | 250   |

A partir das Tabelas 2 e 3 é possível observar que o algoritmo genético obteve resultados expressivamente melhores do que a metodologia da literatura. Pela Tabela 2 observa-se que o algoritmo genético se mostrou eficiente e robusto para encontrar boas soluções com pouco tempo de processamento. A Tabela 3 apresenta os resultados para as instâncias mais difíceis onde pode-se observar que o algoritmo genético obteve resultados expressivamente melhores do que os reportados na literatura, sendo esses, até onde sabemos, atualmente os melhores limitantes conhecidos para as instâncias em questão. Em muitas instâncias a diminuição do custo das soluções superou 15%, com a maior diminuição sendo de

44,45% na instância *egl-s4-C* que é a maior instância considerada nos experimentos.

Um fator que se mostrou de grande importância para a solução de uma instância OCARP é o valor do parâmetro  $M$  que limita o número de veículos disponíveis. Nas Tabelas 2 e 3 é possível verificar que nos casos onde  $M = M^* + 1$  ou  $M = M^* + 2$  a heurística tomou um tempo de processamento consideravelmente menor do que quando  $M = M^*$ .

O bom desempenho do algoritmo genético, em comparação ao algoritmo da literatura, pode ser atribuído à eficiência da rotina de factibilização que permite ao algoritmo encontrar soluções factíveis de bom custo mesmo se o número de veículos  $M$  é limitado.



Tabela 2: Resultados da heurística para as instâncias oval.

| Instância | M  | M*              |                 |                 |       | M*+1            |                 |                 |      | M*+2            |                 |                 |      |
|-----------|----|-----------------|-----------------|-----------------|-------|-----------------|-----------------|-----------------|------|-----------------|-----------------|-----------------|------|
|           |    | LP <sub>1</sub> | LP <sub>2</sub> | $\Delta LP(\%)$ | CPU   | LP <sub>1</sub> | LP <sub>2</sub> | $\Delta LP(\%)$ | CPU  | LP <sub>1</sub> | LP <sub>2</sub> | $\Delta LP(\%)$ | CP   |
| val1A     | 2  | 154             | 154             | 0.00            | 0     | 154             | 151             | 1.95            | 1    | 154             | 149             | 3.25            | 1    |
| val1B     | 3  | 154             | 151             | 1.95            | 1     | 149             | 149             | 0.00            | 1    | 149             | 147             | 1.34            | 1    |
| val1C     | 8  | 173             | 148             | 14.45           | 0     | 149             | 146             | 2.01            | 0    | 146             | 146             | 0.00            | 0    |
| val2A     | 2  | 195             | 195             | 0.00            | 0     | 195             | 192             | 1.54            | 1    | 195             | 189             | 3.08            | 1    |
| val2B     | 3  | 192             | 192             | 0.00            | 0     | 192             | 189             | 1.56            | 1    | 192             | 186             | 3.13            | 0    |
| val2C     | 8  | 197             | 185             | 6.09            | 0     | 186             | 185             | 0.54            | 0    | 185             | 185             | 0.00            | 0    |
| val3A     | 2  | 71              | 71              | 0.00            | 0     | 71              | 69              | 2.82            | 1    | 71              | 67              | 5.63            | 1    |
| val3B     | 3  | 69              | 69              | 0.00            | 0     | 68              | 67              | 1.47            | 1    | 68              | 66              | 2.94            | 1    |
| val3C     | 7  | 68              | 66              | 2.94            | 0     | 65              | 65              | 0.00            | 0    | 65              | 65              | 0.00            | 0    |
| val4A     | 3  | 361             | 358             | 0.83            | 4     | 365             | 354             | 3.01            | 4    | 365             | 350             | 4.11            | 4    |
| val4B     | 4  | 363             | 354             | 2.48            | 6     | 363             | 350             | 3.58            | 3    | 363             | 347             | 4.41            | 2    |
| val4C     | 5  | 364             | 350             | 3.85            | 457   | 355             | 347             | 2.25            | 2    | 355             | 345             | 2.82            | 2    |
| val4D     | 9  | 359             | 343             | 4.46            | 326   | 354             | 343             | 3.11            | 1    | 359             | 343             | 4.46            | 1    |
| val5A     | 3  | 387             | 383             | 1.03            | 3     | 383             | 378             | 1.31            | 4    | 383             | 374             | 2.35            | 3    |
| val5B     | 4  | 385             | 378             | 1.82            | 3     | 386             | 374             | 3.11            | 2    | 386             | 371             | 3.89            | 2    |
| val5C     | 5  | 381             | 374             | 1.84            | 2     | 378             | 371             | 1.85            | 2    | 378             | 368             | 2.65            | 1    |
| val5D     | 9  | 377             | 367             | 2.65            | 1     | 378             | 367             | 2.91            | 1    | 378             | 367             | 2.91            | 0    |
| val6A     | 3  | 196             | 195             | 0.51            | 1     | 196             | 193             | 1.53            | 2    | 196             | 192             | 2.04            | 1    |
| val6B     | 4  | 197             | 194             | 1.52            | 1     | 194             | 192             | 1.03            | 1    | 194             | 191             | 1.55            | 1    |
| val6C     | 10 | 195             | 190             | 2.56            | 0     | 192             | 190             | 1.04            | 0    | 192             | 190             | 1.04            | 0    |
| val7A     | 3  | 267             | 263             | 1.50            | 3     | 267             | 259             | 3.00            | 4    | 267             | 256             | 4.12            | 3    |
| val7B     | 4  | 259             | 259             | 0.00            | 2     | 262             | 256             | 2.29            | 3    | 262             | 253             | 3.44            | 2    |
| val7C     | 9  | 262             | 250             | 4.58            | 2     | 261             | 249             | 4.60            | 1    | 256             | 249             | 2.73            | 0    |
| val8A     | 3  | 366             | 364             | 0.55            | 16    | 366             | 359             | 1.91            | 3    | 366             | 354             | 3.28            | 3    |
| val8B     | 4  | 364             | 359             | 1.37            | 3     | 360             | 354             | 1.67            | 2    | 360             | 351             | 2.50            | 2    |
| val8C     | 9  | 361             | 347             | 3.88            | 11    | 353             | 347             | 1.70            | 0    | 355             | 347             | 2.25            | 0    |
| val9A     | 3  | 303             | 298             | 1.65            | 16    | 303             | 294             | 2.97            | 11   | 303             | 292             | 3.63            | 9    |
| val9B     | 4  | 304             | 294             | 3.29            | 10    | 301             | 292             | 2.99            | 7    | 301             | 290             | 3.65            | 6    |
| val9C     | 5  | 304             | 292             | 3.95            | 6     | 297             | 290             | 2.36            | 5    | 297             | 288             | 3.03            | 4    |
| val9D     | 10 | 300             | 282             | 6.00            | 152   | 294             | 281             | 4.42            | 19   | 294             | 280             | 4.76            | 1    |
| val10A    | 3  | 409             | 402             | 1.71            | 12    | 409             | 396             | 3.18            | 13   | 409             | 391             | 4.40            | 11   |
| val10B    | 4  | 406             | 396             | 2.46            | 11    | 406             | 391             | 3.69            | 9    | 406             | 388             | 4.43            | 7    |
| val10C    | 5  | 406             | 391             | 3.69            | 7     | 404             | 388             | 3.96            | 6    | 404             | 385             | 4.70            | 5    |
| val10D    | 10 | 398             | 379             | 4.77            | 10    | 396             | 378             | 4.55            | 10   | 396             | 377             | 4.80            | 2    |
| média     |    |                 |                 | 2.60            | 31.36 |                 |                 | 2.35            | 3.38 |                 |                 | 3.04            | 2.23 |

LP<sub>1</sub>: limitante primal obtido pela heurística RPS ER de Usberti et al. [2011].

LP<sub>2</sub>: limitante primal obtido pelo algoritmo genético.

$\Delta LP(\%) = 100(LP_1 - LP_2)/LP_1$ : taxa de melhoria de LP<sub>1</sub> para LP<sub>2</sub>. CPU: tempo em segundos para o algoritmo genético obter LP<sub>2</sub>.

Tabela 3: Resultados da heurística para as instâncias oegl.

| Instância | $M^*$  |        |                 |       | $M^{*+1}$ |        |                 |       | $M^{*+2}$ |        |                 |       |      |
|-----------|--------|--------|-----------------|-------|-----------|--------|-----------------|-------|-----------|--------|-----------------|-------|------|
|           | $LP_1$ | $LP_2$ | $\Delta LR(\%)$ | $CPU$ | $LP_1$    | $LP_2$ | $\Delta LR(\%)$ | $CPU$ | $LP_1$    | $LP_2$ | $\Delta LR(\%)$ | $CPU$ |      |
| egl-e1-A  | 5      | 19     | 177             | 9.39  | 13        | 181    | 17              | 5.79  | 11        | 181    | 16              | 8.49  | 29   |
| egl-e1-B  | 7      | 19     | 174             | 10.81 | 5         | 180    | 16              | 9.40  | 6         | 181    | 15              | 12.60 | 568  |
| egl-e1-C  | 1      | 18     | 165             | 9.87  | 65        | 170    | 15              | 7.78  | 60        | 175    | 15              | 11.89 | 1    |
| egl-e2-A  | 7      | 24     | 217             | 11.31 | 343       | 235    | 20              | 11.90 | 56        | 235    | 20              | 13.14 | 5    |
| egl-e2-B  | 1      | 24     | 207             | 14.20 | 118       | 228    | 19              | 12.53 | 1         | 227    | 19              | 13.36 | 2    |
| egl-e2-C  | 1      | 27     | 208             | 22.81 | 255       | 234    | 19              | 14.85 | 1685      | 223    | 19              | 12.24 | 107  |
| egl-e3-A  | 8      | 30     | 252             | 16.96 | 301       | 272    | 24              | 11.40 | 364       | 272    | 23              | 13.01 | 381  |
| egl-e3-B  | 1      | 28     | 240             | 15.80 | 563       | 285    | 23              | 17.68 | 14        | 273    | 23              | 15.26 | 158  |
| egl-e3-C  | 1      | 30     | 235             | 22.59 | 925       | 267    | 22              | 14.48 | 1606      | 272    | 22              | 16.73 | 307  |
| egl-e4-A  | 9      | 32     | 263             | 19.54 | 730       | 311    | 25              | 17.19 | 39        | 298    | 25              | 14.46 | 559  |
| egl-e4-B  | 1      | 32     | 269             | 18.13 | 88        | 311    | 25              | 18.18 | 2986      | 301    | 25              | 16.52 | 328  |
| egl-e4-C  | 1      | 42     | 281             | 33.73 | 3004      | 318    | 25              | 20.99 | 354       | 300    | 24              | 16.85 | 72   |
| egl-s1-A  | 7      | 20     | 179             | 13.30 | 29        | 194    | 16              | 13.51 | 39        | 194    | 16              | 17.57 | 2    |
| egl-s1-B  | 1      | 19     | 174             | 11.11 | 2         | 187    | 16              | 11.61 | 1         | 182    | 15              | 13.62 | 1173 |
| egl-s1-C  | 1      | 22     | 187             | 16.79 | 1686      | 196    | 16              | 16.81 | 2         | 184    | 15              | 18.23 | 1    |
| egl-s2-A  | 1      | 43     | 368             | 15.83 | 68        | 424    | 36              | 14.78 | 242       | 415    | 35              | 14.38 | 463  |
| egl-s2-B  | 2      | 54     | 379             | 30.36 | 755       | 424    | 34              | 17.70 | 3107      | 412    | 34              | 16.98 | 276  |
| egl-s2-C  | 2      | 51     | 373             | 27.94 | 1338      | 431    | 33              | 21.23 | 1486      | 403    | 33              | 17.28 | 1173 |
| egl-s3-A  | 1      | 44     | 380             | 14.00 | 2330      | 430    | 37              | 12.64 | 2653      | 435    | 37              | 14.99 | 3229 |
| egl-s3-B  | 2      | 46     | 367             | 21.45 | 2554      | 435    | 35              | 17.73 | 854       | 417    | 35              | 14.84 | 580  |
| egl-s3-C  | 2      | 48     | 374             | 22.49 | 3233      | 441    | 35              | 19.18 | 2651      | 426    | 34              | 18.12 | 2613 |
| egl-s4-A  | 1      | 52     | 447             | 14.79 | 3377      | 516    | 44              | 14.45 | 193       | 508    | 43              | 13.51 | 184  |
| egl-s4-B  | 2      | 56     | 441             | 21.95 | 3488      | 523    | 43              | 17.29 | 1838      | 498    | 43              | 13.45 | 942  |
| egl-s4-C  | 3      | 90     | 501             | 44.45 | 3100      | 585    | 43              | 25.17 | 1865      | 523    | 42              | 18.20 | 2750 |
| média     |        |        |                 | 19.15 | 1182      |        |                 | 15.17 | 921       |        |                 | 14.82 | 662  |

$LP_1$ : limitante primal obtido pela heurística RPS ER de Usberti et al. [2011].

$LP_2$ : limitante primal obtido pelo algoritmo genético.

$\Delta LR(\%) = 100(LP_1 - LP_2)/LP_1$ : taxa de melhoria de  $LP_1$  para  $LP_2$ .  $CPU$ : tempo em segundos para o algoritmo genético obter  $LP_2$ .

## 5 CONCLUSÃO

Este trabalho apresentou a definição e uma formulação de programação linear inteira para o OCARP. Foi proposto um algoritmo genético para tratar o OCARP. Os principais pontos-chaves do algoritmo apresentado foram:

- Codificação de uma solução ocorre através de uma representação não-capacitada.
- Algoritmo Split para decompor uma rota não-capacitada em sub-rotas capacitadas (solução).
- Rotina de Factibilização com estratégias para guiar o algoritmo Split e obter soluções viáveis.

- Busca Local de Desconstrução e Reconstrução para aprimorar as soluções.

Os resultados obtidos pelo algoritmo genético são favoráveis e mostram que o método conseguiu superar as metodologias existentes na literatura. Entretanto para se provar a otimalidade dos resultados obtidos pelo algoritmo genético ou pelo menos comprovar desvios de otimalidade baixos são necessários modelos matemáticos mais eficientes na prática, portanto esse pode ser um tópico de interesse para pesquisas futuras.

## REFERÊNCIAS

- [1]. Baker, E. J. (1986). Reducing bias and inefficiency in the selection algorithm. Proceedings of the second international conference on genetic algorithms, p. 14–21.
- [2]. Benavent, E., Campos, V., Corberán, A., e Mota, E. (1991). The capacitated arc routing problem: lower bounds. *Networks*, 22:669–690. ISSN 1097-0037.
- [3]. Croes, A. G. (1957). A method for solving traveling-salesman problems. *Operations research*, 6: 791–812. ISSN 0030-364X.
- [4]. da Silva, E. F. (2016). Modelos matemáticos para um problema de caminho de corte. Master's thesis, Universidade de São Paulo.
- [5]. Garey, R. M. e Johnson, S. D. (1978). *Computers and intractability: a guide to the theory of np-completeness*. 1979. San Francisco, LA: Freeman.
- [6]. Goldberg, D. (1988). Genetic algorithms in optimization, search and machine learning. Addison Wesley, 905:205–211.
- [7]. Holland, H. J. (1991). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press. ISBN 0262581116.
- [8]. Li, Y. L. e Eglese, W. R. (1995). An interactive algorithm for vehicle routeing for winter-gritting. *Journal of the Operational Research Society*, p. 217–228. ISSN 0160-5682.
- [9]. Reeves, R. C. (1996). Feature article-genetic algorithms for the operations researcher. *INFORMS journal on computing*, 9:231–250. ISSN 1091-9856.
- [10]. Ulusoy, G. (1985). The fleet size and mix problem for capacitated arc routing. *European Journal of Operational Research*, 22(3):329–337.
- [11]. Usberti, F. L., França, P. M., e França, A. L. M. (2008). Roteamento de Leituristas: Um Problema NP-Difícil. Em: *XL SBPO Brazilian Symposium of Operational Research*. Annals XL SBPO, João Pessoa.
- [12]. Usberti, F. L. (2012). Métodos heurísticos e exatos para o problema de roteamento em arcos capacitado e aberto. PhD thesis, Universidade Estadual de Campinas.
- [13]. Usberti, F. L., França, P. M., e França, A. L. M. (2012). Branch-and-bound algorithm for an arc routing problem. *Annals XLIV SBPO*, Rio de Janeiro.
- [14]. Usberti, F. L., França, P. M., e França, A. L. M. (2011). The open capacitated arc routing problem. *Computers and Operations Research*, 38(11):1543 – 1555. ISSN 0305-0548.
- [15]. Wøhlk, S. (2008). A decade of capacitated arc routing. In *The vehicle routing problem: latest advances and new challenges*, p. 29–48. Springer.

# Capítulo 9

## *META-HEURÍSTICA ILS PARA O PROBLEMA DE POSICIONAMENTO AUTOMÁTICO DE PONTOS DE ACESSO EM AMBIENTES INTERNOS*

*Anderson Rufino*

*Thiago Gouveia da Silva*

*Eduardo Vieira Queiroga*

*Nailson dos Santos Cunha*

*Lucidio dos Anjos F. Cabral*

**Resumo:** As redes Wi-Fi têm se tornado cada vez mais presentes em ambientes internos. O projeto de redes desse tipo, contudo, requer que o posicionamento dos pontos de acesso atenda a diversas restrições relacionadas ao custo do projeto, à área de cobertura, à qualidade do sinal, entre outros critérios. Em virtude da dificuldade do problema quando tratado de forma manual, surgem os métodos que realizam o posicionamento automaticamente. Este trabalho propõe um método baseado na meta-heurística Iterated Local Search (ILS) para o Problema de Posicionamento de Pontos de Acesso (APPP), considerando restrições de capacidade, cobertura interna e externa, e demanda por qualidade. Resultados promissores foram obtidos em experimentos com instâncias sintéticas e reais.

**Palavras-chave:** Projeto de redes. Wireless. ILS-RVND. Área Principal: MH Meta-heurísticas

## 1. INTRODUÇÃO

As redes locais sem fio (WLANs, do inglês Wireless Local Area Networks) têm se tornado cada vez mais importantes para a conectividade de ambientes internos. O crescimento da utilização das redes Wi-Fi (redes WLAN baseadas nos padrões IEEE 802.11) pode ser creditado, principalmente, à diminuição do custo dos equipamentos, ao aumento da confiabilidade e à praticidade da implantação destas redes.

O projeto de redes Wi-Fi de médio e grande porte é uma tarefa complexa que requer, normalmente, que o posicionamento dos roteadores sem fio ou pontos de acesso (APs, do inglês Access Points) siga as diversas restrições relacionadas ao custo do projeto, à área de cobertura, à qualidade do sinal, à minimização da interferência e/ou ao consumo energético, etc. Neste contexto, surgem problemas de otimização combinatória da classe NP-Difícil que podem ser modelados matematicamente e/ou tratados por abordagens exatas e heurísticas.

Neste trabalho, tratamos do Problema de Posicionamento de Pontos de Acesso WLAN (APPP), considerando diferentes tipos de APs, a demanda por qualidade de serviço nos pontos de usuário, a demanda por banda passante e a proibição de cobertura em áreas externas. O APPP é modelado matematicamente como um problema de programação linear inteira baseado no problema de localização de facilidades, provado ser NP-Difícil por Megiddo e Tamir (1982), minimizando os custos dos APs e penalizando a cobertura de pontos proibidos (área externa). Uma meta-heurística baseada no ILS (do inglês, Iterated Local Search) é proposta para resolver instâncias sintéticas e reais de diferentes dimensões. Para avaliar o método proposto, experimentos computacionais envolvendo adaptações de meta-heurísticas da literatura foram realizados.

O trabalho está organizado como segue. Na Seção 2, uma revisão da literatura com trabalhos relacionados ao projeto de redes WLAN em ambientes internos é apresentado. Na Seção 3, o problema é descrito formalmente através de um modelo de programação linear inteira. Na Seção 4, a meta-heurística proposta é apresentada em detalhes. Na Seção 5, o processo de geração de instâncias sintéticas é descrito e justificado. Na Seção 6, os experimentos

computacionais são descritos e os resultados analisados.

## 2. TRABALHOS RELACIONADOS

A literatura apresenta diversos trabalhos que tratam sobre a resolução do APPP. Contudo, a definição do problema possui alta variabilidade, devido à adequação dos métodos para estudos de casos de projetos reais, dificultando a estruturação das pesquisas na área. Apesar de existirem muitos trabalhos propondo métodos exatos para o APPP, por restrições de espaço, o foco desta seção será apenas sobre métodos heurísticos.

Kamenetsky e Unbehauen (2002) propuseram uma combinação de duas abordagens utilizando um algoritmo de remoções sucessivas chamado pruning para obter uma solução inicial, além de refiná-la utilizando a meta-heurística Simulated Annealing (SA). O método é avaliado através de uma função objetivo sobre um espaço discreto, com o intuito de melhorar a área de cobertura e a qualidade do sinal recebido.

Arroyo e Marques (2006) propuseram dois métodos baseados na meta-heurística GRASP para resolver o problema de alocação de antenas de transmissão em uma região. A função objetivo do modelo busca maximizar a cobertura do sinal utilizando um número mínimo de antenas e instalando-as o mais próximo possível dos pontos de demanda, levando em consideração restrições como o alcance do sinal de transmissão e a presença de obstáculos.

O mesmo autor propôs em Marques (2007) outra análise comparativa para o mesmo modelo, desta vez, utilizando algoritmos baseados nas meta-heurísticas GRASP e Algoritmos Genéticos (AG). A função objetivo neste trabalho é definida como a soma dos custos das antenas utilizadas e a soma das distâncias mínimas entre cada ponto de demanda e a antena instalada.

Lu et al. (2006) propuseram uma abordagem baseada na meta-heurística Busca Tabu (TS, do inglês Tabu Search) para a definição da quantidade de APs e o seu posicionamento com respeito à um critério que considera a cobertura, a interferência e a qualidade de serviço (QoS, do inglês Quality of Service). Uma cadeia de Markov é utilizada para avaliar a largura de banda em cada célula da área e

a cobertura é estimada por um simulador de propagação de rádio multiresolução.

Vanhatupa et al. (2007) apresentam um novo algoritmo baseado em AG para a criação de redes WLAN de acordo com os requisitos desejados e faz uma estimativa da qualidade de serviço (cobertura, custos de implantação e capacidade da rede), provendo um feedback para o próprio algoritmo e também para o construtor da rede. O método leva em consideração requisitos como consumo de energia, possíveis locais para posicionamento de APs e quantidade de APs permitidos.

Em Barbosa e Gouvea (2012), uma heurística gulosa e uma abordagem steady-state AG foram propostos para definir o posicionamento de APs, maximizando a cobertura da área e o número de usuários atendido.

Farkas et al. (2013) propuseram um método baseado em SA com o objetivo de encontrar o posicionamento e o número ideal de pontos de acesso Wi-Fi de tal forma que cada ponto na área seja coberto por pelo menos 3 APs.

Por fim, Capdeville e Vianna (2013) propõem duas implementações diferentes do GRASP para resolver o APPP em um ambiente interno de uma instituição federal de ensino. O modelo é baseado no problema de localização de facilidades, levando em consideração a atenuação do sinal por obstáculos. Os autores também propuseram

duas heurísticas de busca local para complementar a meta-heurística.

### 3.FORMULAÇÃO MATEMÁTICA PROPOSTA

Nesta seção, a versão do APPP tratada é formalizada através de um modelo de programação linear inteira. Para discretizar o problema, uma grade de pontos candidatos sobre a área é considerada. Dentre esses pontos,  $P$  é o conjunto de todos os pontos que precisam ser cobertos,  $R$  o conjunto dos pontos onde é possível alocar um AP,  $T$  o conjunto dos diferentes tipos de APs e  $S$  o conjunto dos pontos proibidos.

Temos que  $p_k$  é o custo de utilizar o roteador do tipo  $k \in T$ ;  $q_s$  a penalidade pelo ponto proibido  $s$  estar na área de cobertura de pelo menos um roteador ativo;  $c_k$  a capacidade do roteador do tipo  $k$  e  $d_i$  a demanda por largura de banda do ponto  $i$ . A variável de decisão  $x_{ijk}$  indica se o ponto  $i$  está ligado ao roteador  $j$  do tipo  $k$ . A variável de decisão  $r_{jk}$  assume o valor 1 se o roteador  $j$  do tipo  $k$  está sendo utilizado na solução. A variável de decisão  $s_h$  assume o valor 1 se o ponto proibido  $h$  está na área de cobertura de algum roteador utilizado. Seja  $cob(r_{jk})$  uma função que retorna o conjunto de todos os pontos que estão na área de cobertura do roteador  $r_{jk}$ , e  $cobp(r_{jk})$  uma função análoga para os pontos proibidos. A formulação matemática proposta é apresentada nas expressões (1-6):

$$\begin{aligned}
 & \min \sum_{j \in R} p_k r_{jk} + \sum_{h \in S} q_h s_h & (1) \\
 \text{s.a.} & \sum_{k \in T} x_{ijk} = 1, & \forall i \in P, & (2) \\
 & j \in R, k \in T \\
 & r_{jk} \geq x_{ijk}, & \forall i \in cob(r_{jk}), \forall j \in R, \forall k \in T, & (3) \\
 & \sum_{i \in P} d_i x_{ijk} \leq c_k, & \forall j \in R, \forall k \in T, & (4) \\
 & s_h \geq r_{jk}, & \forall h \in cobp(r_{jk}), \forall j \in R, \forall k \in T, & (5) \\
 & r_{jk}, x_{ijk}, s_h \in \{0, 1\}, & \forall i \in P, \forall j \in R, \forall k \in T, \forall h \in S. & (6)
 \end{aligned}$$

A função objetivo (1) busca minimizar o custo total da utilização dos roteadores, juntamente com a minimização da penalização de cobertura de pontos proibidos. As restrições (2) estabelecem que todos os pontos clientes estejam associados a um roteador. As restrições (3) indicam que cada ponto só pode ser ligado a um roteador ativo.

As restrições (4) definem que a demanda por largura de banda dos pontos conectados ao roteador  $k$  não pode ultrapassar a capacidade do mesmo. A restrição (5) indica se o ponto proibido  $s_h$  está dentro da área de cobertura de algum roteador. Por fim, as restrições (6) definem o domínio das variáveis de decisão.

#### 4. MÉTODO ILS-RVND PROPOSTO

Para a resolução do APPP como definido na Seção 3, propomos um algoritmo baseado na meta-heurística Iterated Local Search (ILS) em conjunto com uma busca local Randomized Variable Neighborhood Descent (RVND). Uma característica importante do método proposto é que ele permite a exploração do espaço de soluções inviáveis do problema. A busca local, por sua vez, tem a responsabilidade de retornar apenas soluções viáveis. Os procedimentos de geração da solução inicial, busca local e perturbação são descritos nas subseções subsequentes.

Segundo Lourenço et al. (2003), a meta-heurística ILS é composta por quatro procedimentos básicos: GeraSoluçãoInicial(), que constrói a primeira solução do problema;

BuscaLocal(), que melhora a solução de entrada fazendo-a convergir para um ótimo local; CritérioDeAceitação(), que determina se a solução corrente será aceita ou não como a nova solução e Perturbação(), que modifica a solução para a próxima iteração do algoritmo. O ILS-RVND proposto é detalhado no Algoritmo 1.

Temos a geração de uma solução inicial na linha 1, sendo  $S$  a solução atual,  $S^*$  a melhor solução da iteração e  $S^{**}$  a melhor solução encontrada pelo método. Na linha 2, o procedimento de iteração do método inicia-se, o qual possui um tempo limite de duração delimitado pela comparação das variáveis tempoExecução e tempoLimite. Na linha 3, são inicializadas as variáveis  $k$ , responsável pela intensidade das perturbações e

Algoritmo 1: ILS-RVND()

```

1  $S^{\square} \leftarrow S^{\square} \leftarrow S$  GeraSoluçãoInicial();
2 enquanto tempoExecução  $\leq$  tempoLimite faça
3    $k \leftarrow 1$ ; iterSemMelhora  $\leftarrow 0$ ;
4   enquanto iterSemMelhora++  $<$  maxIterSemMelhora faça
5      $S \leftarrow$  RVND( $S$ );
6     se  $f_0(S) < f_0(S^{\square})$  então
7        $S^{\square} \leftarrow S$ ;  $k \leftarrow 1$ ; iterSemMelhora  $\leftarrow 0$ ;
8        $S \leftarrow S^{\square}$ ;
9       para cada  $i = 1, 2, \dots, \text{Min}(k, \text{maxK})$  faça
10         $S \leftarrow$  Perturbação( $S$ );
11    $S^{\square\square} \leftarrow$  Best( $S^{\square\square}, S^{\square}$ );
12    $S \leftarrow S^{\square} \leftarrow$  GeraSoluçãoInicial();
13 retorne  $S^{\square\square}$ 

```

iterSemMelhora, que monitora a quantidade de iterações sem melhora da solução. Entre as linhas 4 e 10 temos o principal laço do algoritmo, o qual possui os procedimentos de busca local e perturbação. Este laço encerra-se quando o número de iterações sem melhora da solução atingir um valor máximo, este último representado pela constante maxIterSemMelhora. Neste trabalho utilizamos maxIterSemMelhora=1000.

Na linha 5, temos a obtenção de um ótimo local após a aplicação do método de busca local RVND sobre a solução de entrada. Na linha 6 é verificado se o novo ótimo local será aceito através do valor da função objetivo (novamente, na primeira iteração sempre haverá uma melhora pelo fato da solução de entrada ser inviável) que, em caso positivo, é armazenado e os valores das variáveis  $k$  e

iterSemMelhora são alterados para os valores iniciais.

Na linha 8, o algoritmo garante que o procedimento de perturbação será aplicado ao ótimo local sem comprometer sua integridade. Nas linhas 9 e 10, o procedimento de perturbação com intensidade  $k$  é aplicado ao ótimo local, gerando uma solução (possivelmente inviável) que servirá de entrada para a próxima iteração do laço. Se o número de iterações sem melhora atingir o valor máximo antes do tempo de execução atingir o tempo limite, a melhor solução encontrada até então será armazenada e outra solução será construída.

#### 4.1 PROCEDIMENTOS DE GERAÇÃO DA SOLUÇÃO INICIAL

Na fase de construção, a solução inicial é gerada de forma gulosa. Cada roteador é posicionado de forma a atingir a maior quantidade de pontos clientes cobertos possível, ou seja, dentro do raio de cobertura do roteador. Os empates são quebrados de forma equiprovável. Pelo fato das restrições de capacidade não serem levadas em consideração, soluções inviáveis serão construídas e penalizadas, levando a um alto valor de função objetivo.

#### 4.2 HEURÍSTICAS DE BUSCA LOCAL

Nesta seção descrevemos o procedimento de busca local implementado. Este consiste nas cinco estruturas de vizinhança descritas a seguir, que são executadas conforme o framework RVND padrão.

`allChangeLink()`: Este procedimento verifica, para todos os pontos clientes da solução que estão conectados a algum roteador, se este último está com uma capacidade de banda passante acima do limite. Em caso positivo, será feita uma tentativa de reconectar o ponto corrente para algum outro roteador que ainda não esteja na sua capacidade máxima de banda.

`allDryRouter()`: Procura por roteadores ativos que estejam com uma exigência de banda passante acima do limite. Após identificá-los, um a um, tenta distribuir todos os pontos conectados ao roteador corrente para outros roteadores ativos que ainda não estão na sua capacidade máxima de banda. Se o procedimento for realizado com sucesso, ou seja, se todos os pontos do roteador analisado forem conectados a outros roteadores, o mesmo será retirado da solução, diminuindo o custo da função objetivo.

`bestAddRouter()`: Dados os pontos clientes ainda não conectados a nenhum roteador da solução, procura a posição onde, posicionando-se um roteador, este último atenderá a maior quantidade possível de pontos clientes não conectados. Encontrada tal posição, um roteador é adicionado à solução e todos os pontos clientes dentro do raio de cobertura do roteador serão conectados ao mesmo, desde que não ultrapassem sua capacidade.

`allBestRadiusType()`: Analisa todos os roteadores da solução. Se um roteador possuir algum ponto cliente conectado com

uma demanda por qualidade de sinal acima/abaixo da necessária (ponto cliente próximo/distante do roteador), este roteador será substituído por outro com um raio de cobertura menor ou maior, se disponível, a fim de satisfazer a demanda por qualidade de sinal do ponto cliente com eficiência.

`bestLink()`: Verifica, um por um, os pontos clientes ainda não conectados da solução. Se o ponto cliente corrente está dentro do raio de cobertura de um roteador ativo e este possui uma quantidade de banda suficiente para receber o ponto cliente, o procedimento fará a ligação do ponto ao roteador em questão.

Após o método de busca local RVND, um ótimo local foi encontrado. O critério de aceitação do algoritmo é encontrar um novo ótimo local e continuar a busca por melhores soluções partindo do mesmo. Se isto acontecer, este ótimo local de melhor qualidade é armazenado (aceito) e o procedimento de perturbação é aplicado sobre este.

#### 4.3 MECANISMOS DE PERTURBAÇÃO

Nesta seção, descrevemos os procedimentos utilizados para perturbar as soluções geradas. As perturbações modificam o ótimo local gerando uma nova solução, com o objetivo de explorar outras regiões do espaço de soluções. Uma variável  $k$  controla o nível da perturbação. Cada vez que o procedimento é executado, são escolhidas  $k$  heurísticas de perturbação para serem executadas. As heurísticas utilizadas são descritas a seguir:

`randomDelRouter()`: Retira aleatoriamente um roteador da solução atual. Após a deleção, os pontos clientes que estavam conectados a este roteador ficarão desconectados e os pontos proibidos que estavam sendo cobertos ficarão descobertos.

`randomDelLink()`: Exclui a conexão entre um ponto cliente escolhido aleatoriamente e o roteador ao qual está conectado. Esta operação é executada cinco vezes, com o objetivo de garantir a eficácia da perturbação na solução atual.

`randomChangeRouterType()`: Escolhe aleatoriamente um roteador da solução e modifica o seu tipo. Este último também é escolhido aleatoriamente. Como consequência da execução desta



perturbação, raio, limite de banda e custo do roteador modificado serão alterados, podendo acarretar em pontos clientes com demanda por qualidade de sinal insatisfeita e roteadores com limite de banda excedida.

`randomMoveRouter()`: Move um roteador da solução para uma outra posição escolhida aleatoriamente. Isso fará com que os pontos conectados ao roteador, dependendo da posição escolhida, passem a possuir uma demanda por qualidade de sinal insatisfeita ou satisfeita, além de cobrir e descobrir pontos proibidos.

`randomJoinRouter()`: Dados dois roteadores `r1` e `r2` escolhidos aleatoriamente, este procedimento transfere todos os pontos conectados ao roteador `r1` para o roteador `r2`, sem se preocupar com a demanda por qualidade de sinal dos pontos clientes ou com o limite de banda do roteador alvo. Isto causará a existência de pontos clientes com demanda por qualidade de sinal insatisfeita

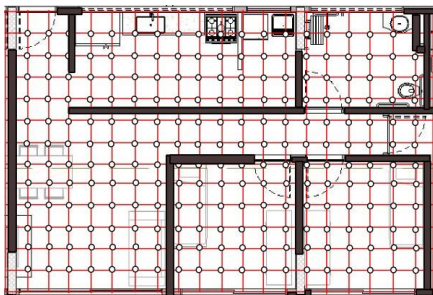
ou satisfeita e de roteadores com seu limite de capacidade de banda excedido.

## 5. GERAÇÃO DE INSTÂNCIAS

Uma vez que a literatura na área do APP é bastante fragmentada e possui distintos tipos de objetivos, se fez necessária a criação de instâncias para avaliar a eficiência dos métodos propostos. Para tal, foi utilizada a discretização do espaço em uma grade de pontos, conforme exibido na Figura 1. Cada ponto é posicionado (horizontal e verticalmente) há um metro dos outros.

O modelo de propagação utilizado para a geração de instâncias foi baseado nos estudos de Ogunjemilua et al. (2009), Androne e Palade (2010) e Sohail et al. (2013), sobre o comportamento do sinal das redes Wi-Fi na frequência de 2.4 Ghz. A Tabela 1 descreve os valores de atenuação utilizados nas instâncias geradas.

Figura 1. Discretização do espaço



Foram gerados dois tipos de instâncias: randômicas e reais, todas utilizando os três tipos de roteadores descritos na Tabela 2. Foram gerados 7 grupos de instâncias randômicas, cada um contendo 10 casos de teste em ordem crescente de dificuldade. Cada grupo possui o mesmo conjunto de pontos e de anteparos, sendo o tipo de

Tabela 1. Modelo de atenuação

| Tipo de Anteparo             | Atenuação |
|------------------------------|-----------|
| Divisória de Madeira (4cm)   | 2db       |
| Divisória de Vidro (3cm)     | 3db       |
| Drywall (5-10cm)             | 4db       |
| Parede de Tijolo (15cm)      | 6db       |
| Parede de Concreto (15-20cm) | 9db       |
| Parede de Concreto (25-30cm) | 12db      |

anteparo selecionado aleatoriamente, assim como a demanda por qualidade e por largura de banda de cada ponto. Além disso, cada grupo é identificado pela largura e altura (em metros) da área onde foram gerados os pontos, sendo o grupo 8x14 o menor e o 50x50 o maior.

Tabela 2. Tipos de roteadores utilizados

| Tipo de Roteador        | Custo   | Alcance | Banda Passante |
|-------------------------|---------|---------|----------------|
| Roteador de Entrada     | 150 R\$ | 100m    | 150 Mbps       |
| Roteador Médio          | 200 R\$ | 100m    | 300 Mbps       |
| Roteador Alta Qualidade | 300 R\$ | 400m    | 400 Mbps       |

Por fim, foram geradas 5 instâncias baseadas em plantas-baixas reais: me 100x40, referente a uma estação de metrô; ho 60x32, referente à um andar de hotel; of 100x35, referente a um escritório; sc 70x45, referente a uma escola; e sh 100x60, referente a um pequeno shopping center. Equivalente às instâncias randômicas, as instâncias reais são identificadas pelas largura e altura das plantas-baixas utilizadas como base.

## 6. EXPERIMENTOS COMPUTACIONAIS

Os experimentos computacionais foram realizados em um computador Intel Core i7 1.80 GHz com 4GB de RAM e sistema operacional Ubuntu 14.04 LTS. Todos os métodos foram implementados na linguagem C++ e compilados com o g++ na versão 4.8.4, utilizando a opção -O3. O método exato foi implementado utilizando a linguagem Concert e os testes foram executados utilizando o solver CPLEX 12.51 com tempo limite de 2 horas.

O ILS proposto foi comparado com adaptações das meta-heurísticas SA (Farkas et al., 2013), AG (Barbosa e Gouvea, 2012) e TS (Lu et al., 2006). Estas adaptações possuem a representação da solução e, conseqüentemente, o cálculo da função objetivo adequados para a versão do problema tratado neste trabalho. Cada método foi executado 10 vezes para cada instância, com um tempo limite de 10s para instâncias de tamanho 8x14 e 13x13, 30s para instâncias de tamanho 11x23, 25x25 e 30x30 e 60s para as instâncias reais, de tamanho 50x25 e 50x50. A escolha do tempo limite baseou-se na quantidade de execuções necessárias para se obter boas soluções para determinada instância.

Os resultados apresentados nas tabelas (3-10) reportam o melhor valor da função objetivo e a média aritmética, obtidos entre 10 execuções para cada instância do grupo. Para o grupo r 8x14 (tabela 3), o ILS e o GA

empatam na solução mínima encontrada, porém o GA obteve uma média melhor na instância 7. Para o grupo r 13x13 (tabela 4), o ILS consegue leve vantagem na solução mínima, com destaque para a instância 7, e perde na solução média apenas nas instâncias 0 e 6. Para o grupo r 11x23 (tabela 5), o ILS possui desempenho competitivo com relação ao AG, com destaque para a instância 1, onde houve uma diferença considerável.

Para o grupo r 25x25 (tabela 6), o ILS demonstrou maior capacidade de resolução e robustez que os demais métodos, obtendo apenas uma derrota na instância 8. Nos grupos r 30x30 e r 50x25 (tabelas 7 e 8), fica evidente a vantagem obtida pelo ILS, uma vez que o mesmo conseguiu melhores resultados em quase todas as instâncias. Para o grupo r 50x50 (tabela 9), o ILS conseguiu os melhores resultados para todas as instâncias.

Dentro do tempo limite de 2 horas, a formulação matemática proposta foi capaz de resolver todas as instâncias do grupo r 8x14, as 6 primeiras instâncias do grupo r 13x13 e apenas a primeira instância do grupo r 11x23. Quando os métodos heurísticos alcançam o resultado ótimo obtido pelo método exato, este é grafado em negrito e com um asterisco.

Os resultados obtidos para o conjunto de instâncias reais é reportado na tabela 10. Nesta, k significa 1000 vezes. Para este grupo, o ILS novamente conseguiu os melhores resultados para todas as instâncias. Os métodos da literatura, no entanto, enfrentaram dificuldades neste grupo, alcançando valores de função objetivo muito altos. Acreditamos que o tamanho das instâncias reais não permite que estes métodos executem suas buscas locais dentro do tempo limite estabelecido, reportando apenas soluções iniciais construídas aleatoriamente.

De acordo com os experimentos realizados, percebe-se que o método abordado neste

trabalho mostra-se bastante competitivo em relação aos outros métodos da literatura implementados para instâncias pequenas. Porém, à medida que as instâncias ficam

maiores, ou seja, as quantidades de pontos e de obstáculos aumenta, o ILS-RVND consegue obter resultados melhores.

Tabela 3. Resultados Computacionais para o Grupo de Instâncias r 8x14

| Inst # | ILS  |       | AS   |       | AG   |     | BT   |       |
|--------|------|-------|------|-------|------|-----|------|-------|
|        | min  | avg   | min  | avg   | min  | avg | min  | avg   |
| 0      | 339* | 339   | 339* | 339.9 | 339* | 339 | 339* | 339   |
| 1      | 181* | 181   | 181* | 290.3 | 181* | 181 | 181* | 255.5 |
| 2      | 181* | 181   | 181* | 229.9 | 181* | 181 | 181* | 242.2 |
| 3      | 333* | 333   | 333* | 351   | 333* | 333 | 333* | 333.4 |
| 4      | 326* | 326   | 339  | 339.5 | 326* | 326 | 326* | 335.6 |
| 5      | 338* | 338   | 340  | 340   | 338* | 338 | 338* | 369.6 |
| 6      | 222* | 222   | 223  | 302.9 | 222* | 222 | 222* | 263.7 |
| 7      | 336* | 336.6 | 337  | 353.4 | 336* | 336 | 338  | 377.4 |
| 8      | 229* | 229   | 330  | 338.6 | 229* | 229 | 330  | 331.8 |
| 9      | 339* | 339   | 389  | 439.9 | 339* | 339 | 339* | 394   |

Tabela 4. Resultados Computacionais para o Grupo de Instâncias r 13x13

| Inst # | ILS  |       | SA   |       | AG   |        | BT   |        |
|--------|------|-------|------|-------|------|--------|------|--------|
|        | min  | avg   | min  | avg   | min  | avg    | min  | avg    |
| 0      | 300* | 301   | 300* | 349.8 | 300* | 300    | 300* | 315.8  |
| 1      | 305* | 305   | 359  | 401.5 | 305* | 305    | 305* | 320.3  |
| 2      | 513  | 513   | 563  | 692.8 | 513  | 513    | 513  | 538    |
| 3      | 512* | 512   | 563  | 658   | 512* | 512.1  | 513  | 553    |
| 4      | 613* | 613   | 663  | 778   | 613* | 613    | 613* | 633    |
| 5      | 613* | 613   | 713  | 813   | 613* | 613    | 613* | 673    |
| 6      | 703  | 703.4 | 856  | 908.5 | 703  | 703    | 703  | 743.1  |
| 7      | 812  | 812.4 | 913  | 1088  | 813  | 813    | 813  | 898    |
| 8      | 963  | 963   | 1063 | 1243  | 963  | 977.8  | 1012 | 1072.4 |
| 9      | 1212 | 1212  | 1413 | 1543  | 1212 | 1212.6 | 1213 | 1317.8 |

Tabela 5. Resultados Computacionais para o Grupo de Instâncias r 11x23

| Inst | ILS  |        | SA   |        | AG   |        | BT   |        |
|------|------|--------|------|--------|------|--------|------|--------|
|      | #    | min    | avg  | min    | avg  | min    | avg  | min    |
| 0    | 457* | 457    | 507  | 600.6  | 457* | 457    | 457* | 523.8  |
| 1    | 522  | 522    | 573  | 702.9  | 569  | 569.2  | 569  | 640.3  |
| 2    | 618  | 619.2  | 723  | 842.9  | 617  | 617.5  | 617  | 704.5  |
| 3    | 750  | 753.6  | 859  | 933.1  | 750  | 756.3  | 800  | 842.6  |
| 4    | 823  | 823    | 1023 | 1163   | 872  | 872    | 872  | 907.4  |
| 5    | 1019 | 1094   | 1217 | 1332.4 | 1112 | 1114.6 | 1017 | 1167.3 |
| 6    | 1113 | 1113   | 1373 | 1533   | 1113 | 1113.6 | 1167 | 1198.3 |
| 7    | 1155 | 1156.1 | 1362 | 1475.9 | 1150 | 1150   | 1150 | 1211.7 |
| 8    | 1323 | 1323   | 1673 | 1753   | 1322 | 1322.2 | 1372 | 1442.6 |
| 9    | 1523 | 1523   | 1873 | 2043   | 1522 | 1532.4 | 1623 | 1662.  |

Tabela 6. Resultados Computacionais para o Grupo de Instâncias r 25x25

| Inst | ILS  |      | SA   |      | AG   |      | BT   |      |
|------|------|------|------|------|------|------|------|------|
|      | #    | min  | avg  | min  | avg  | min  | avg  | min  |
| 0    | 1300 | 1300 | 1700 | 1780 | 1300 | 1345 | 1300 | 1375 |
| 1    | 1300 | 1300 | 1600 | 1665 | 1300 | 1325 | 1300 | 1360 |
| 2    | 1600 | 1600 | 1850 | 2135 | 1600 | 1615 | 1600 | 1745 |
| 3    | 1800 | 1865 | 2100 | 2395 | 1900 | 1940 | 1800 | 1940 |
| 4    | 2000 | 2060 | 2500 | 2670 | 2050 | 2135 | 2000 | 2165 |
| 5    | 2300 | 2380 | 2650 | 3065 | 2400 | 2495 | 2350 | 2470 |
| 6    | 2500 | 2580 | 3050 | 3300 | 2600 | 2660 | 2600 | 2795 |
| 7    | 3050 | 3075 | 3750 | 3985 | 3200 | 3260 | 3050 | 3190 |
| 8    | 3800 | 3830 | 4500 | 4825 | 3950 | 4005 | 3750 | 3865 |
| 9    | 3900 | 3980 | 4600 | 4955 | 4050 | 4125 | 4000 | 4195 |

Tabela 7. Resultados Computacionais para o Grupo de Instâncias r 30x30

| Inst<br># | ILS  |        | SA   |        | AG   |        | BT   |        |
|-----------|------|--------|------|--------|------|--------|------|--------|
|           | min  | avg    | min  | avg    | min  | avg    | min  | avg    |
| 0         | 1200 | 1247.3 | 1464 | 1718.3 | 1368 | 1473.6 | 1259 | 1359.1 |
| 1         | 1333 | 1333   | 1679 | 1859   | 1442 | 1480.9 | 1392 | 1458.1 |
| 2         | 1775 | 1834.7 | 2125 | 2285   | 1875 | 1940   | 1825 | 1959.9 |
| 3         | 1600 | 1618.4 | 2193 | 2347.4 | 1894 | 1954.4 | 1600 | 1923.1 |
| 4         | 2174 | 2258.6 | 2725 | 2790   | 2375 | 2475   | 2325 | 2478.6 |
| 5         | 2451 | 2466.4 | 3325 | 3380.6 | 2623 | 2709.1 | 2551 | 2661.9 |
| 6         | 3075 | 3095   | 3375 | 3735   | 3175 | 3255   | 3125 | 3475   |
| 7         | 2924 | 2935.4 | 3625 | 3849.6 | 3125 | 3213.9 | 3175 | 3335.1 |
| 8         | 4159 | 4198.9 | 5125 | 5310   | 4375 | 4530   | 4275 | 4830   |
| 9         | 3892 | 4014.5 | 4819 | 5088.2 | 4323 | 4434.9 | 4345 | 5481.  |

Tabela 8. Resultados Computacionais para o Grupo de Instâncias r 50x25

| Inst<br># | ILS  |      | SA    |       | AG   |      | BT   |       |
|-----------|------|------|-------|-------|------|------|------|-------|
|           | min  | avg  | min   | avg   | min  | avg  | min  | avg   |
| 0         | 2200 | 2270 | 2750  | 3035  | 2400 | 2610 | 2200 | 3000  |
| 1         | 2600 | 2690 | 3150  | 3485  | 2800 | 2825 | 2650 | 2785  |
| 2         | 3600 | 3680 | 4450  | 4685  | 3850 | 3925 | 3700 | 3930  |
| 3         | 4000 | 4050 | 4900  | 5095  | 4100 | 4395 | 3950 | 4545  |
| 4         | 4900 | 4935 | 5600  | 6230  | 5250 | 5380 | 4900 | 5655  |
| 5         | 5000 | 5140 | 6200  | 6585  | 5350 | 5470 | 5250 | 5500  |
| 6         | 5900 | 6020 | 7100  | 7500  | 6300 | 6565 | 6050 | 6910  |
| 7         | 6400 | 6510 | 8000  | 8255  | 6800 | 7035 | 6450 | 7245  |
| 8         | 7600 | 7750 | 9350  | 9675  | 8250 | 8425 | 8200 | 9345  |
| 9         | 8500 | 8680 | 10350 | 10685 | 9250 | 9430 | 9850 | 10430 |

Tabela 9. Resultados Computacionais para o Grupo de Instâncias r 50x50

| Inst # | ILS   |         | SA    |         | AG    |         | BT    |         |
|--------|-------|---------|-------|---------|-------|---------|-------|---------|
|        | min   | avg     | min   | avg     | min   | avg     | min   | avg     |
| 0      | 4850  | 4930    | 6050  | 6215    | 5400  | 5650    | 6050  | 7500    |
| 1      | 5950  | 6100    | 7500  | 7800    | 6700  | 6985    | 7100  | 9050    |
| 2      | 7050  | 7190    | 8700  | 9180    | 7600  | 7935    | 7600  | 8825    |
| 3      | 8031  | 8112.8  | 9624  | 10130   | 9500  | 9784.5  | 10450 | 13784.4 |
| 4      | 9400  | 9510    | 11700 | 12080   | 10550 | 11000   | 11300 | 13095   |
| 5      | 10911 | 11015.7 | 13284 | 13695.4 | 12294 | 12687.4 | 13242 | 15593.6 |
| 6      | 11250 | 11355   | 13350 | 14135   | 13300 | 13895   | 15000 | 18065   |
| 7      | 12731 | 12960.5 | 15598 | 16300.2 | 15500 | 16180   | 16100 | 18570   |
| 8      | 16450 | 16565   | 19950 | 20395   | 18400 | 19085   | 20750 | 22210   |
| 9      | 17050 | 17230   | 20950 | 21595   | 19700 | 20170   | 21100 | 23060   |

Tabela 10. Resultados Computacionais para as Instâncias Reais

| Inst #    | ILS  |        | SA   |      | AG   |       | BT   |      |
|-----------|------|--------|------|------|------|-------|------|------|
|           | min  | avg    | min  | avg  | min  | avg   | min  | avg  |
| ho 60x32  | 770  | 770    | 920  | 970  | 820  | 850   | 820  | 880  |
| sc 70x45  | 8994 | 9247,9 | 95k  | 997k | 286k | 361k  | 417k | 458k |
| of 100x35 | 3490 | 3579,9 | 332k | 338k | 105k | 176k  | 223k | 396k |
| me 100x40 | 5250 | 5400   | 5750 | 5930 | 352k | 376k  | 428k | 456k |
| sh 100x60 | 1550 | 1625   | 930k | 935k | 4300 | 13485 | 5750 | 523k |

## 7. CONSIDERAÇÕES FINAIS E PROPOSTA DE TRABALHOS FUTUROS

Neste trabalho, propusemos um algoritmo heurístico baseado na meta-heurística Iterated Local Search (ILS), o qual apresentou bons resultados na resolução do Problema de Posicionamento de Pontos de Acesso (APPP). O método se mostrou bastante competitivo em relação aos outros métodos da literatura implementados para pequenas instâncias, enquanto que para instâncias maiores, conseguiu superá-los de forma promissora.

Como trabalhos futuros, pretendemos aprimorar o método heurístico proposto

através da criação de algoritmos de pré-processamento e de novas heurísticas de busca local, com foco nas instâncias as quais não obtivemos os resultados desejados, além de expandir os métodos descritos para a realização de posicionamento de APs em três dimensões. Acreditamos também que o AG parece promissor e, por isso, vamos tentar melhorá-lo para conseguir bons resultados para as instâncias maiores. Ainda para o futuro da pesquisa, pretendemos, adicionalmente, desenvolver um software que permita a criação de plantas-baixas para a subsequente execução do posicionamento automático dos APs.

## REFERÊNCIAS

- [1]. Androne, C. e Palade, T. Radio coverage and performance analysis for local area networks. *Electronics and Telecommunications (ISETC)*, 2010 9th International Symposium on, p. 213–216. IEEE, 2010.
- [2]. Arroyo, J. E. C. e Marques, T. B. (2006), Heurística grasp aplicado ao problema de alocação de antenas de transmissão. XXXVIII Simpósio Brasileiro de Pesquisa Operacional, Goiânia-GO.
- [3]. Barbosa, M. A. S. e Gouvea, M. M. Access point design with a genetic algorithm. *Genetic and Evolutionary Computing (ICGEC)*, 2012 Sixth International Conference on, p. 119–123. IEEE, 2012.
- [4]. Capdeville, R. M. A. e Vianna, D. S. (2013), Heurísticas grasp para o problema de alocação de pontos de acesso em uma rede sem fio em ambiente indoor. *Sistemas & Gestão*, v. 8, n. 1, p. 86–93.
- [5]. Farkas, K., Huszák, Á. e Gódor, G. (2013), Optimization of wi-fi access point placement for indoor localization. *Journal IIT (Informatics & IT Today)*, v. 1, n. 1, p. 28–33.
- [6]. Kamenetsky, M. e Unbehaun, M. Coverage planning for outdoor wireless lan systems.
- [7]. *Broadband Communications, 2002. Access, Transmission, Networking. 2002 International*

# Capítulo 10

## TEMPERATURE FIELD FORECAST IN CONCRETE DAM WITH THE USE OF ARIMA MODELS AND THE FINITE ELEMENT METHOD

*Tásia Hickmann*

*Luiz Albino Teixeira Júnior*

*Alvaro Eduardo Faria*

*Samuel Bellido Rodrigues*

*Jairo Marlon Côrrea*

*Everton Luís Garcia*

**Resumo:** Este artigo descreve um método de previsão, com a aplicação de modelos estatísticos Auto-Regressivo Integrado Média Móvel (ARIMA) e um modelo de condução de calor para prever o campo de temperatura em um bloco de contraforte da barragem de Itaipu. As séries de temperatura mensais de 2010-2014 dos termômetros da superfície ao bloco foram ajustadas por meio de splines cúbicas e a séries, agora diárias, foram utilizadas como entradas para modelos específicos ARIMA para produzir previsões como saídas. Estas saídas foram utilizadas como condições de contorno para o modelo térmico do bloco e este resolvido pelo Método dos Elementos Finitos (MEF). Obtidos assim os campos de temperatura previstos do bloco. O erro MAPE entre os valores obtidos pelo MEF e o real, em um ponto de prova, (onde está um termômetro interno) mediu o desempenho da previsão dos modelos ARIMA, e este foi satisfatório, alcançando valor próximo a 15%. O método proposto tem um caráter inovador para estruturas de análise térmica, em particular em barragens de concreto.

**Palavras-Chave.** Equação do Calor, Séries Temporais, métodos aproximados.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2016)

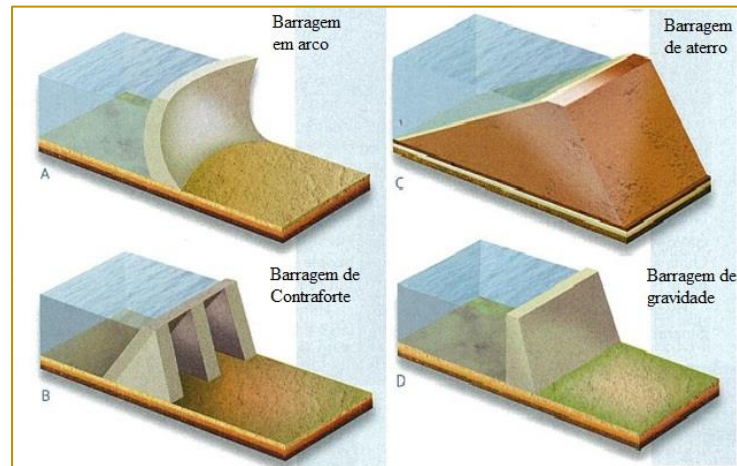


## 1. INTRODUCTION

Dams are hydraulic structures built in the course of a river to create a reservoir on its upstream side for the abstraction of water to be used in irrigation, water supply, flood control, navigation, fishing or hydroelectricity production. Dams built for hydroelectricity generation can be classified into two large

groups according to the types of material used in their construction, namely (i) concrete dams, such as gravity arches or buttresses, or (ii) conventional land and rock fill dams. The dam drawings in Figure 1A, 1B and 1D illustrate concrete dams and while the one in Figure 1C illustrates a conventional land and rock fill dam.

Figure 1. Types of dams according to the material used.



Source: Adapted from AUSTRALIAN GEOGRAPHIC

According to [Maken et al. 2013] dams are subject to changes in internal temperature due to external temperature changes and hydration heat (a heat source from concrete chemical reactions during the dam construction but which dissipates in time). The external heat load to a dam surface originates from solar radiation as well as from air, foundation and reservoir temperature variations. Those external heat loads can cause serious physical deformations and significant volumetric changes in the dam structure that often exceed the tensile strength of concrete leading to the development of cracks. Therefore the monitoring of external sources of temperature is crucial for dam structural safety. In fact, a correct evaluation of the temperature field in the dam is essential for determining stresses and deformations of thermal origin. In this regard, predictions of the temperature field are necessary for monitoring and predicting deformations in the dam structure. In this paper, statistical forecasting methods are devised to produce reliable systematic predictions of that temperature field.

To help assess the risk of dam deformation of thermal origin, we have formulated and

applied time series forecasting models to the temperatures recorded by thermometers placed in a buttress dam block of the Itaipu Hydroelectric Power Plant (IHPP) during operational phase. Since these thermometers have only monthly readings, set initially to cubic splines interpolating data to obtain daily data, needed to solve the thermal block model. With these new series, applied Auto-Regressive Integrated Moving Average (ARIMA) forecasting models and the temperature forecasts produced by those models were used as boundary conditions in a thermal model represented by a heat conduction equation for the block. The thermal model in turn was solved by the Finite Element Method (FEM) with the aid of ANSYS simulation software to model the temperature field in the block during the forecasting period.

There is a body of literature covering methods to describe the temperature field in concrete dams through the FEM. In [Daoud et al. 1997], for example, a numerical evaluation of the periodic temperature field in a concrete dam was performed that accounted for ambient temperature changes, solar radiation, snow

cover, temperature gradients and ice formation in reservoir water.

Mirzabozorg et al. 2014] investigated the effects of solar radiation on the thermal distribution on a dam arc and compared the results with those obtained from data recorded from the dam site. The development of [Noorzaei et al. 2006] focused on implementing and verifying a two-dimensional finite element code developed for thermal and structural analysis of Kinta gravity dam in Malaysia, during its construction. Local climatic conditions and thermal properties of materials were considered in the analysis. The temperatures provided by their approximate method were considered in good agreement with the actual temperatures measured by thermocouples installed in the dam body.

For applications of the ANSYS simulation software in engineering problems, particularly in thermal structure analysis, please refers to [Moaveni 2008] and [Madenci and Guven 2015].

Despite the above mentioned contributions to the literature on the thermal modelling of concrete dams, we notice the lack of research related to the forecasting analysis of dam temperature field during operational phase. With observed data from thermometers installed in concrete, the use of statistical forecasting ARIMA models of [Box and Jenkins 1970] and the FEM method (see e.g. [Madenci and Guven 2015]) for heat diffusion models to characterize the evolution of the temperature field in the dam in a future period is feasible. Therefore, prediction of the thermo-structural behavior of the dam can be investigated so that intervention in the structure may be taken, if necessary.

The mathematical modeling of conductive heat transfer process, the characterization of

FEM in the Galerkin method for solving the thermal model, the statistical forecasting of temperature from thermometers installed on the surface of a buttress concrete block of a dam and the forecast of temperature fields in this block are the objectives of this work.

This paper is structured as follows: Section 2 reviews the technical literature associated with the models that compound the proposed thermal forecasting method. It describes the heat conduction equation with the initial and boundary conditions necessary for obtaining its solution and an approximate method for solving the thermal model (the FEM) as well as the statistical forecasting ARIMA model. Section 3 describes the proposed methodology and its application to a case study of the IHPP dam. Predictions of the temperature field of a block of concrete in the dam is obtained by the method and its performance is evaluated. Section 4 concludes the paper.

## 2. ORIENTED LITERATURE REVIEW

### 2.1. HEAT CONDUCTION EQUATION

Conduction heat transfer aims to determine the temperature field of a medium resulting from conditions imposed on its borders. The classical equation for heat diffusion is determined by the first law of thermodynamics and by the Fourier law. It is an empirical relationship involving the heat flow and the temperature gradient through a proportionality constant called the thermal conductivity of the properties materials.

According to [Ozisik 1993], the heat conduction equation for an anisotropic material is given by:

$$\frac{\partial}{\partial x} \left( k_x(T) \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left( k_y(T) \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left( k_z(T) \frac{\partial T}{\partial z} \right) + G = \rho c \frac{\partial T}{\partial t}, \quad (1)$$

where  $k_x(T)$ ,  $k_y(T)$  and  $k_z(T)$  are the thermal conductivities ( $W/mK$ ) dependent on the temperature in the directions  $x$ ,  $y$  and  $z$ , respectively;  $\rho$  is the specific mass ( $kg/m^3$ ),  $c$  is the specific heat ( $J/kg K$ ) of the material

e  $G$  is the rate of internal heat generation (internal energy) per unit volume ( $W/m^3$ ). For an isotropic material in which the thermal coefficients  $k$ ,  $\rho$  and  $c$  are all constant, and there is no internal heat generation and Equation (1) can be rewritten as:

$$\nabla^2 T = \frac{1}{\alpha} \frac{\partial T}{\partial t}, \quad (2)$$

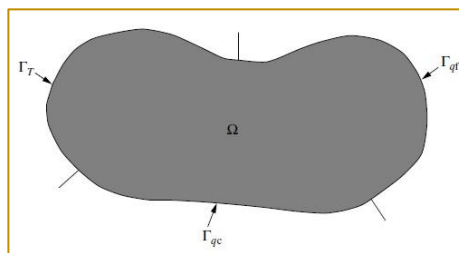
where  $\nabla^2$  is the Laplacian operator, and  $\alpha = \frac{k}{\rho c}$ , is called thermal diffusivity ( $m^2/s$ ) (see e.g. [LEWIS et al. 2004]).

### 2.1.1. INITIAL AND BOUNDARY CONDITIONS

The solution of Equation (1) is possible if the defined initial and boundary conditions are

appropriate to the analyzed physical problem. Consider an arbitrary domain over which the equation of heat conduction applies, bounded by a surface, such that, where is the surface on which a temperature  $T$  function is prescribed, and are the areas in which affect heat flow from, for example, by an external heat source and convective heat flow, respectively (see Figure 2).

Figure 2. Boundary conditions.



According to [Lewis et al. 2004], the boundary conditions for the heat equation can be of two types, or a combination thereof - of Dirichlet condition, which is the set temperature in a

portion of the domain boundary and/or Neumann condition, which is flow heat prescribed in another portion of the domain boundary.

Dirichlet Condition

$$T = T_0, \text{ in } \Gamma_T. \quad (3)$$

Neumann Condition

$$q = -k \frac{\partial T}{\partial n}, \text{ in } \Gamma_{qf}, \quad (4)$$

where  $T_0$  is a prescribed temperature ( $K$ ),  $n$  is the normal vector to the surface  $\Gamma_{qf}$  and  $q$  is the steady flow ( $W/m^2$ ).

The convective heat flux also fits as a Neumann boundary condition and can be written as:

$$-k \frac{\partial T}{\partial n} = h(T - T_a), \text{ in } \Gamma_{qc}, \quad (5)$$

where  $h$  is the heat transfer coefficient by convection ( $W/m^2$ ),  $T$  is the surface temperature in  $\Gamma_{qc}$  and  $T_a$  is the temperature of the adjacent fluid. The initial condition is the

temperature all the domain  $\Omega$  in a reference time  $t = t_0$ , usually called initial instant. This temperature can be represented as:

$$T = T_0, \text{ in } \Omega \text{ in } t = t_0. \quad (6)$$

The Equations (4) and (5) can be described in terms of the cosine directors of the normal

vector  $n = \{\tilde{l}, \tilde{m}, \tilde{n}\}$ , appropriate to the surface  $\Gamma_{af}$  and  $\Gamma_{qc}$  :

$$k_x(T) \frac{\partial T}{\partial x} \tilde{l} + k_y(T) \frac{\partial T}{\partial y} \tilde{m} + k_z(T) \frac{\partial T}{\partial z} \tilde{n} = q, \quad (7)$$

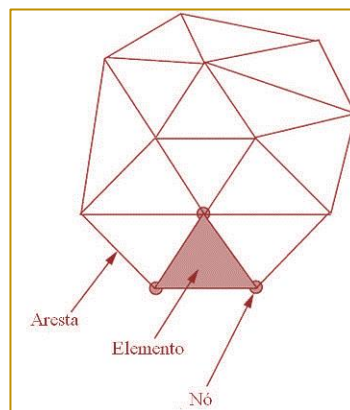
$$k_x(T) \frac{\partial T}{\partial x} \tilde{l} + k_y(T) \frac{\partial T}{\partial y} \tilde{m} + k_z(T) \frac{\partial T}{\partial z} \tilde{n} = h(T - T_a). \quad (8)$$

## 2.2. THE FINITE ELEMENT METHOD

The FEM is a methodology to obtain numerical solution of equations in differential and integral form, particularly in physical-mathematical field equations. These are approximate to single domain regions (triangles, quadrilaterals, tetrahedra, etc.) to meet together, the continuity of the field variables (displacement, temperature, pressure) at nodal points of elements of the domain (see e.g. [Gupta and Meek 1996]).

According to [Krüger 2001], in problems of Continuum Mechanics, functions as displacements, stresses, pressures and temperature are associated with infinite points of the solution region. The process of discretization in finite elements reduces the size of the problem to a finite number of unknowns by dividing the solution domain elements (mesh) and expressing the variables in terms of approximation of functions on each element. Figure 3 shows an example of a mesh of finite elements.

Figure 3. Example of finite element mesh.



Source: Adapted from [Lewis et al. 2004].

The steps to achieve an approximate solution to a physical model using FEM (according to [Lewis et al. 2004]) are:

- 1- Discretization of continuous: dividing the solution region into subdomains or non-overlapping elements;
- 2- Selection of basis functions or interpolation: such functions determine the variation of the unknown variable in study;

- 3- Formulation of the equations of nodes: take into account the individual properties of each element;

- 4- Assembling the global matrix for simultaneous solution of the elements equations;

- 5- Application of the initial and boundary conditions and loadings;

6- Resolution of the global system for determining the nodal results, such displacements and temperatures on different nodes.

There are several approaches to formulate problems in finite element, such as Direct Method, Minimum Potential Energy, Variational methods and Waste weighted. In the following, Equation (1) is discretized by a

method of the latter case, the Galerkin method.

### 2.2.1 GALERKIN METHOD

An expression that describes how temperatures are interpolated in a finite element is given by the following equation:

$$T(x, y, z, t) = \sum_{i=1}^n N_i(x, y, z) T_i(t), \quad (9)$$

where  $N_i$  are the basis functions,  $n$  is the number of nodes in an element and  $T_i(t)$  is a nodal temperature time-dependent.

The most widely used basic functions are polynomial type, in particular linear and quadratic according to the size of the mesh elements. Examples of such functions and

applications can be viewed in [Lewis et al. 2004] and [Onate 2009].

Multiplying Equation (1) a base function  $N_i$  and integrating over the domain of analysis, the representation of Galerkin is:

$$\int_{\Omega} N_i \left[ \frac{\partial}{\partial x} \left( k_x(T) \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left( k_y(T) \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left( k_z(T) \frac{\partial T}{\partial z} \right) + G - \rho c \frac{\partial T}{\partial t} \right] d\Omega = 0. \quad (10)$$

Using integration by parts on the first three terms of Equation (10) we obtain:

$$\begin{aligned} & - \int_{\Omega} \left[ k_x(T) \frac{\partial N_i}{\partial x} \frac{\partial T}{\partial x} + k_y(T) \frac{\partial N_i}{\partial y} \frac{\partial T}{\partial y} + k_z(T) \frac{\partial N_i}{\partial z} \frac{\partial T}{\partial z} \right] d\Omega + \int_{\Omega} \left[ N_i G - N_i \rho c \frac{\partial T}{\partial t} \right] d\Omega \\ & + \int_{\Gamma_q} N_i k_x(T) \frac{\partial T}{\partial x} \tilde{l} d\Gamma_q + \int_{\Gamma_q} N_i k_y(T) \frac{\partial T}{\partial y} \tilde{m} d\Gamma_q + \int_{\Gamma_q} N_i k_z(T) \frac{\partial T}{\partial z} \tilde{n} d\Gamma_q = 0, \end{aligned} \quad (11)$$

where  $\Gamma_q = \Gamma_{qc} \cup \Gamma_{qf}$  and in  $\Gamma_T$  the temperature variation is zero when considering that this boundary the

temperature is constant. Note that Equations (7) e (8),

$$\int_{\Gamma_q} N_i k_x(T) \frac{\partial T}{\partial x} \tilde{l} d\Gamma_q + \int_{\Gamma_q} N_i k_y(T) \frac{\partial T}{\partial y} \tilde{m} d\Gamma_q + \int_{\Gamma_q} N_i k_z(T) \frac{\partial T}{\partial z} \tilde{n} d\Gamma_q = - \int_{\Gamma_{qf}} N_i q d\Gamma_{qf} - \int_{\Gamma_{qc}} N_i h(T - T_a) d\Gamma_{qc}. \quad (12)$$

Through Equation (12) and approach (9), the Equation (11) can be written as:

$$\begin{aligned} & - \sum_{j=1}^n \int_{\Omega} \left[ k_x(T) \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} T_j(t) + k_y(T) \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} T_j(t) + k_z(T) \frac{\partial N_i}{\partial z} \frac{\partial N_j}{\partial z} T_j(t) \right] d\Omega \\ & + \int_{\Omega} \left[ N_i G - N_i \rho c \sum_{j=1}^n N_j \frac{\partial T_j(t)}{\partial t} \right] d\Omega - \int_{\Gamma_{qf}} N_i q d\Gamma_{qf} - \int_{\Gamma_{qc}} N_i h \left( \sum_{j=1}^n N_j T_j(t) - T_a \right) d\Gamma_{qc} = 0, \end{aligned} \quad (13)$$

where  $i$  and  $j$  represents nodes. Is possible to write in matrix terms Equation (13):

$$[C_{ij}] \left\{ \frac{\partial T_j}{\partial t} \right\} + [K_{ij}] \{T_j\} = \{f_i\}, \tag{14}$$

where the term  $\frac{\partial T_j}{\partial t}$  can be discretized using the finite difference method, as can be seen in

[Lewis et al. 2004] and the other terms of the Equation (14) are defined as:

$$[C_{ij}] = \int_{\Omega} \rho c N_i N_j d\Omega,$$

$$[K_{ij}] = \int_{\Omega} \left[ k_x(T) \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + k_y(T) \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} + k_z(T) \frac{\partial N_i}{\partial z} \frac{\partial N_j}{\partial z} \right] d\Omega + \int_{\Gamma_{qc}} h N_i N_j d\Gamma_{qc} \text{ and}$$

$$\{f_i\} = \int_{\Omega} N_i G d\Omega - \int_{\Gamma_{qf}} q N_i d\Gamma_{qf} + \int_{\Gamma_{qc}} h T_a N_i d\Gamma_{qc}.$$

### 2.3. ARIMA MODELS

The Auto-Regressive Integrated Moving Average (ARIMA) models, developed by [Box and Jenkins 1970], are widely known and used in modeling and forecasting time series. Let  $y_t (t = 1, \dots, T)$  be a time series, i.e., a set of

observations ordered in time, displaying an auto-correlation structure. According to [Liu 2009],  $y_t (t = 1, \dots, T)$  can be seen as a realization of a stochastic ARIMA  $(p, d, q)$  process, if it can be represented by

$$\nabla^d y_t = \phi_0 + \sum_{j=1}^p \phi_j \nabla^d y_{t-j} + \sum_{i=1}^q \theta_i e_{t-i} + e_t, \tag{15}$$

where  $\nabla^d := (I - B)^d$  is the difference operator, with  $d$  representing its order, and  $B$  is the delay operator, defined as  $B^k y_t := y_{t-k}$ , with  $k$  belonging to a set of integers;  $(\phi_i)_{i=0}^p$  and  $(\theta_i)_{i=0}^q$  are the sets of auto-regressive and moving average parameters, respectively, such that  $\phi_p \neq 0$  and  $\theta_q \neq 0$ , both satisfying the invertibility and the stationarity conditions (as in [Hamilton 1994]);  $e_t$  represents a state of the random variable  $\varepsilon_t (t = 1, \dots, T)$ , a stochastic white noise process, which mean and auto-covariance are stationary at zero;  $p$  and  $q$  are, respectively, the orders of the auto-regressive component, denoted by  $\sum_{j=1}^p \phi_j \nabla^d y_{t-j}$ , and the moving average component, represented by  $\sum_{i=1}^q \theta_i e_{t-i}$  in Equation (15).

plausible values for the parameters  $p$ ,  $d$ , and  $q$ , in Equation (15), which can often sometimes be obtained by profile analysis of simple autocorrelation functions and partial waste (as in [Hamilton 1994]); **(ii)** define the method to be used to estimate the ARIMA parameters (the most commonly used is the maximum likelihood estimation method); and, **(iii)** perform a diagnostic test to choose the most parsimonious and appropriate model to be used to generate both the in-sample and out-of-sample forecasts of the underlying time series  $y_t (t = 1, \dots, T)$ .

### 3. THE PROPOSED METHOD AND ITS APPLICATION TO THE ITAIPU DAM

A buttress dam block of the IHPP dam, located on the Paraná River on the border between Brazil and Paraguay, was used in the application of the proposed method to predict the temperature field in this block. Four thermometers were installed in the block's surface and their measurements of monthly temperature recorded over a period of four years from 2010 to 2014. Figure 4 shows the

According to [Liu 2009], in order to obtain the best possible ARIMA model, the following three basic steps should be taken: **(i)** obtain

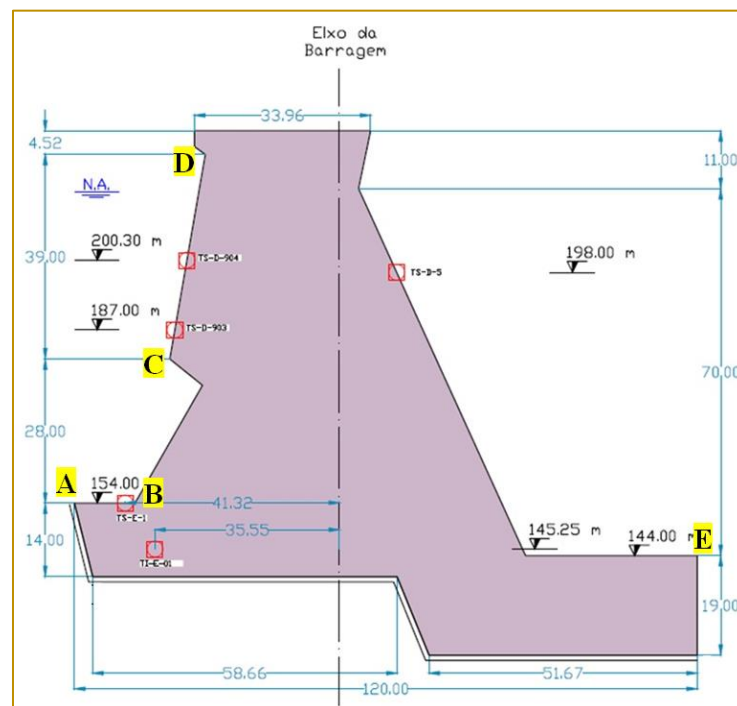
two-dimensional geometry of the block and the location of thermometers: TS-D-903, TS-D-904, TS-E-1 and TS-D-5.

Monthly temperatures for each thermometer were collected and used in the application. The data was split into a training in-sample (from 11/01/2010 to 11/11/2013) and a forecasting out-of-sample (from 09/12/2013 to 27/11/2014).

The method consists of two parts. One part deals with the statistical forecasting of temperature while the other with the heat conduction model for the concrete block. The initial stage consists in disaggregating the monthly data into daily estimations with

polynomial interpolation using cubic splines (with a MATLAB program). So, cubic splines interpolating curves were fitted to the data of each of the thermometers to approximate data daily temperatures, as the readings of thermometers in the study period were only monthly, in order that it is possible later to obtain the thermal model solution of the block. The simplicity of representation and facility the splines can be computed, cause they are popular in the data set in computer science and computer engineering, predominantly in computer graphics. The term spline is a device used by shipbuilders to draw soft forms.

Figure 4. Geometry buttress block, location of thermometers and delimitation of boundary conditions.



According to [Ruggiero and Lopes 1997], a cubic spline interpolation can be defined as follows: Let  $f(x)$  be a function defined at the points  $x_0 < x_1 < \dots < x_n$ , the function  $S_3(x)$  is

called *cubic spline* with nodes at the points  $x_i$ ,  $i = 0, \dots, n$  if it satisfies the following conditions:

I - In each subinterval  $[x_i, x_{i+1}]$ ,  $i = 0, \dots, (n-1)$ ,  $S_3(x)$  is a polynomial of degree 3:  $s_3(x)$ ;

II -  $S_3(x)$  is continuous and has continuous derivatives of the second order  $[x_0, x_n]$ ;

$$\text{III} - S_3(x_i) = f(x_i), i = 0, \dots, n.$$

This is followed by fitting an ARIMA model to each of the four daily series adjusted by splines in the training in-sample (November 2010 to October 2013) and producing out-of-sample forecasts (November 2013 to October 2014) (with an EViews program).

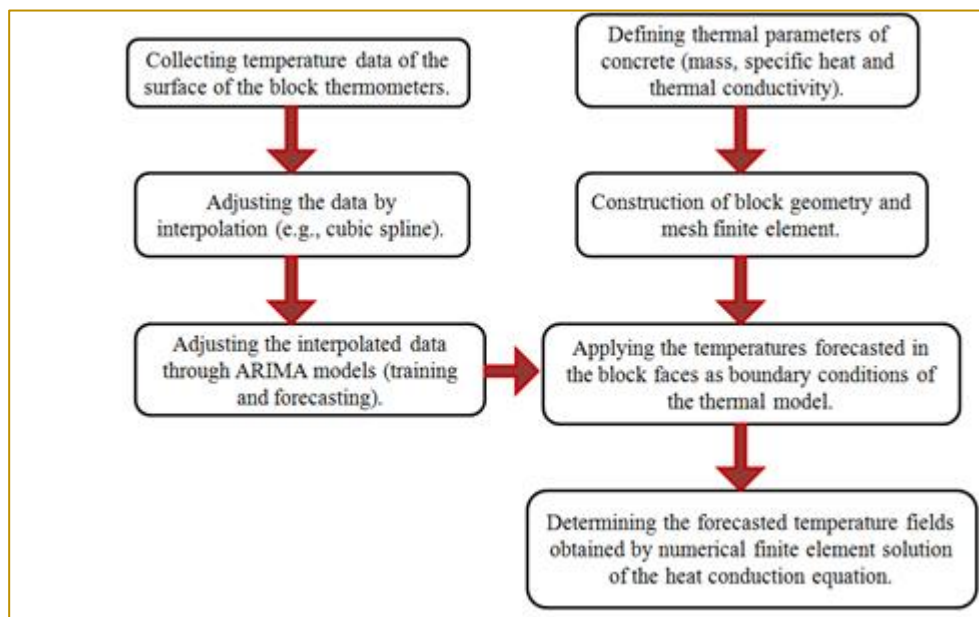
The heat conduction model part on the other hand starts by defining the concrete thermal parameters (specific mass and heat and thermal conductivity), block geometry and the mesh of finite elements (with an ANSYS program). The forecasts from the ARIMA models are then applied to the geometry contour and used to solve the heat conduction models via FEM (also with ANSYS).

After obtaining the monthly temperatures expected for a period of one year, these were taken as boundary conditions in the following block: when viewing Figure 5, from point A to

point B (clockwise) the adopted temperature function was the thermometer TS E-1, from point B to C of TS-D-903 thermometer, point C to D TS-D-904 thermometer from the point D to E TS-D-5 thermometer and finally from point D to point E, the face was considered that it had zero heat flow, ie,  $q = 0$ .

With the well-defined boundary conditions, heat conduction equation (1) can be solved by applying the FEM and the prediction of transient temperature field in the block is checked. The thermal properties of concrete, initial condition to be adopted in the thermal model, the necessary simplifications and the results obtained from the FEM are discussed in the following subsection. The Figure 5 shows a flowchart of those stages of the proposed method.

Figure 5. Flowchart of the proposed method.



### 3.1 NUMERICAL RESULTS

The Figure 6 shows for each of the four series of temperature, comparative graphs between the temperatures estimated by the cubic splines and those forecast by the ARIMA

models for the out-of-sample year. The MAPE (Mean Absolute Error Percentage) was used to estimate the error between the models, and it is given by the following expression:

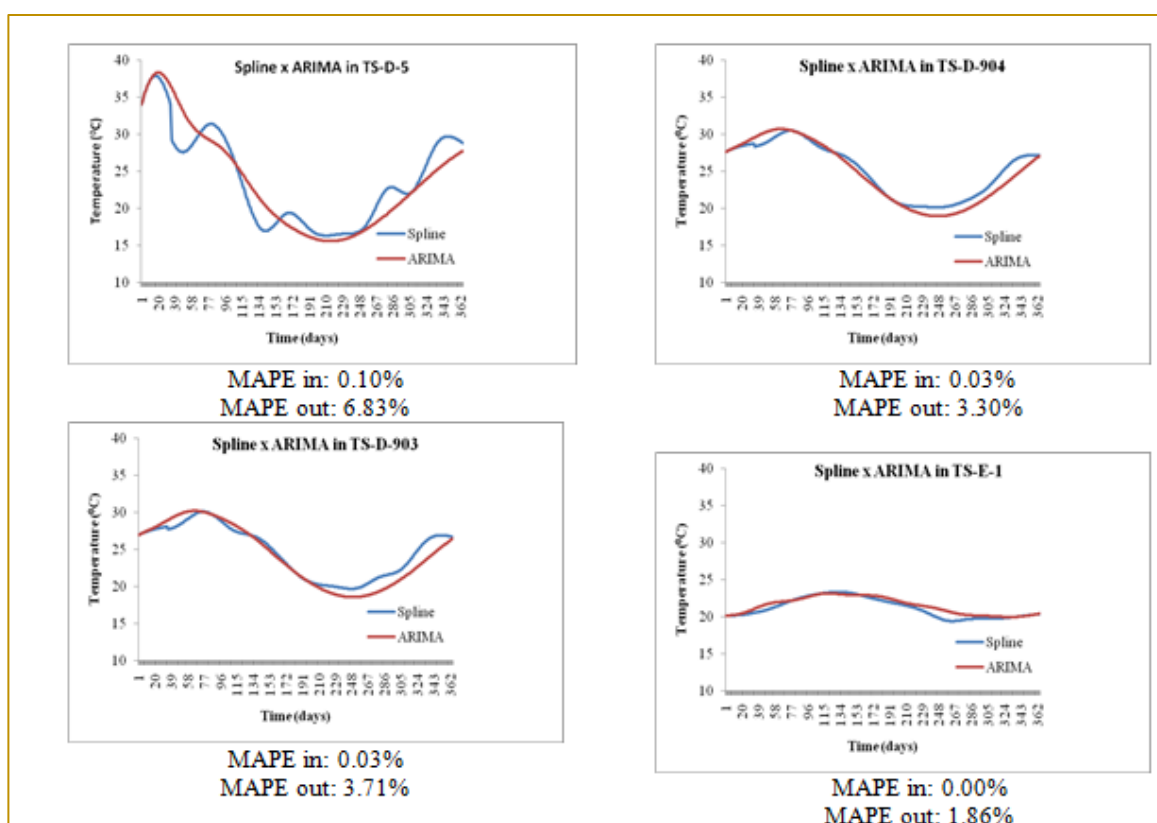


$$\text{MAPE} = \frac{\sum_{i=1}^n \left| \frac{T_i - \hat{T}_i}{T_i} \right|}{n}, \quad (166)$$

where  $n$  is the number of observed data,  $T_i$  the  $i$ -th temperature estimated by the cubic splines and  $\hat{T}_i$  the approximate value of the  $i$ -th temperature determined by the ARIMA

model. The MAPE *in* is the MAPE calculated for the in-sample period and MAPE *out* is calculated for the out-of-sample period for each of the 4 surface thermometers.

Figure 6. Cubic splines vs ARIMA for the out-of-sample period for each of the surface thermometers (TS-D-5, D-TS-903, TS-904 and TS-D-E-1).



Note from the plots in Figure 6 and the respective low MAPE in values that the forecasts by the ARIMA models are approximating well the data interpolated by cubic splines.

Some simplifications were considered in the process of generating the numerical thermal model: no source of heat, the material homogeneity, isotropic and linear elastic behavior, considered the state of plane stress and simplifications in geometric section (chamfers, drainage, supports, and other beams). The following values of thermal properties of concrete were used in the

model: a thermal conductivity of  $1.2W(mK)^{-1}$ , a specific heat of  $895.38J(kg K)^{-1}$  and a specific mass of  $2550kgm^{-3}$ .

The simulations performed in ANSYS environment - Workbench V16.1 and graphics in Thermal Transient environment. The scheme accomplished through software to block thermal analysis started with the geometry of the mesh creation, definition of the boundary conditions and the thermal properties of concrete. Initially a stationary thermal analysis was carried out in order to create a temperature field and entered in the

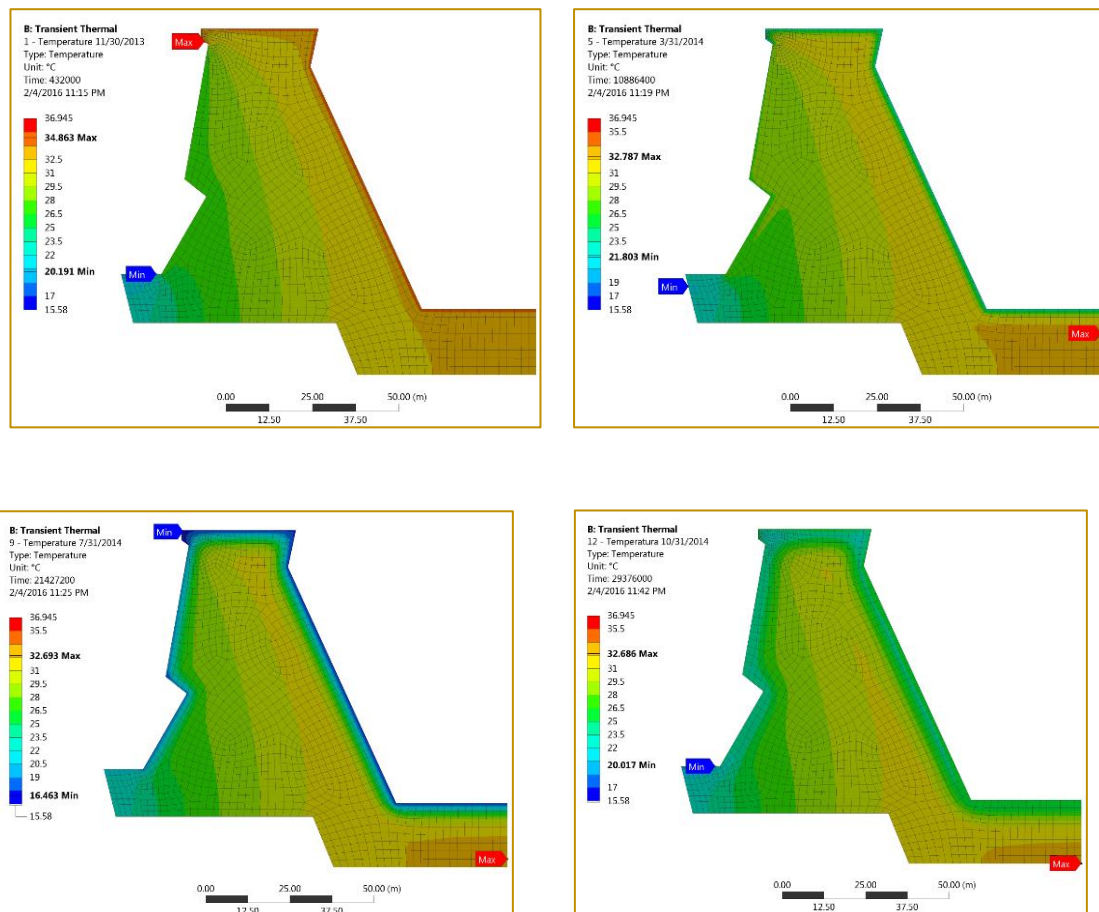
transient thermal analysis as an initial condition. Finally, monthly temperature fields were generated throughout the out-of-sample forecasting year.

The mesh used for the geometry shown in Figure 7 showed 1386 elements (SHELL 131 type, which is a shell element with four nodes and the temperature up to 32 degrees of freedom at each node and such element served for transient analyzes and stationary) and 1497 nodes, so that almost entirely the elements are quadrangular and some triangular. After the geometry of the mesh and concrete properties are defined, a stationary simulation with the temperature data

approximate temperature on 25/11/2013 was performed with the aim being to block an equilibrium temperature range and this used as initial temperature of the block to the transient analysis.

Following temperature data set of 4 thermometers obtained through the ARIMA model in the period from 26/11/2013 to 31/10/2014 served as boundary conditions to the thermal model, as mentioned earlier in this chapter. This was solved by FEM in ANSYS. Figure 7 provides trimonthly temperature fields during this period, considering the last day of the month.

Figure 7. Trimonthly transient temperature fields.



The color scale at the left hand side of each of the four diagrams in Figure 7 shows the various temperatures (in Celsius) in increasing order from minimum values at the bottom (dark blue) to maximum values at the top (red). Looking at the block in each diagram, note that the highest temperatures are in the downstream phase (contact with the environment), while warmer temperatures are located upstream (contact with the reservoir).

The simulation results as displayed in Figure 7 showed that heat storage in the dam of the soul throughout the year, since the color change in the graph is closest to the surface, and as was to be expected on the type of material, the energy dissipation is gradually slow. Moreover, it has monitored a control point within the structure where the internal thermometer TI-E-01 is located (see Figure 4). The data measured by this thermometer for

the out-of-sample forecasting year were compared with those obtained by numerical simulation (the FEM) and the calculated MAPE was 15.37%. With this, the thermal model of the block is within the engineering safety limits, according to the superintendence of the Itaipu Power Plant Engineering.

#### 4. CONCLUSIONS

Predicted temperature fields provided by ARIMA forecasting models and the finite element method for heat equation solution were determined for a case study on a buttress dam block of Itaipu Hydroelectric Power Plant. Observed data of four surface thermometers in the block have been adjusted by interpolating cubic splines functions and their ARIMA forecasting models determined.

The daily forecasts obtained one year ahead for each of the thermometers were used as boundary conditions for solving the heat

conduction equation for the concrete block. The numerical solution of the heat conduction equation was achieved through the Finite Element Method. To validate the method, the real temperatures for the forecast period on a test point, which is located an internal thermometer (TI-E-1), were compared with those obtained with the values obtained by FEM, reaching a MAPE satisfactory.

The proposed method in this paper is a novel approach to modelling heat transfer by conduction in concrete dam. One year ahead temperature field predictions in concrete dam blocks have been shown feasible in this study. Another benefit of this methodology is the fact that it encourages further studies to investigate the prediction of stresses and strains in the thermal source structure that is essential for preventive measures in relation to the life of the structure.

#### REFERENCES

- [1] Box, George and Jenkins, Gwilym (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- [2] Daoud, M., Galanis, N., and Balliiv, G. (1997). Calculation of the Periodic Temperature Field in a Concrete Dam. *Canadian Journal of Civil Engineering*, 24(5), 772–784.
- [3] Gupta, K. K., and Meek, J. L. (1996). A Brief History of the Beginning of the Finite Element Method. *International Journal for Numerical Methods in Engineering*, 39, 3761–3774.
- [4] Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.
- [5] Krüger, D. A. V. (2001). *Análise Térmica Transiente de Estruturas de Concreto Executadas por Camadas*. UFPR.
- [6] Lewis, R. W., Nithiarasu, P., and Seetharamu, K. N. (2004). *Fundamentals of the Finite Element Method for Heat and Fluid Flow*. John Wiley & Sons.
- [7] Liu, L.-M. (2009). *Time Series Analysis and Forecasting*. Scientific Computing Associates Corp.; 2nd Edition edition.
- [8] Madenci, E., & Guven, I. (2015). *The Finite Element Method and Applications in Engineering Using ANSYS*. 2. ed., Springer Science & Business Media.
- [9] Maken, D. D., Léger, P., and Roth, S. N. (2013). Seasonal thermal cracking of concrete dams in northern regions. *Journal of Performance of Constructed Facilities*, American Society of Civil Engineers.
- [10] Mirzabozorg, H., Hariri-Ardebili, M. A., Shirkhan, M., and Seyed-Kolbadi, S. M. (2014). Mathematical Modeling and Numerical Analysis of Thermal Distribution in Arch Dams Considering Solar Radiation Effect. *The Scientific World Journal*, Hindawi Publishing Corporation.
- [11] Moaveni, S. (2008). *Finite Element Analysis: Theory and Application with ANSYS*. 3. ed., Pearson Education India.
- [12] Noorzaei, J., Bayagoob, K. H., Thanoon, W. A., and Jaafar, M. S. (2006). Thermal and Stress Analysis of Kinta RCC Dam. *Engineering Structures*, Elsevier, 28(13), 1795–1802.
- [13] Onate, E. (2009). *Structural Analysis with the Finite Element Method. Linear Statics: Basis and Solids*. Springer Science & Business Media, Vol. 1.
- [14] Ozisik, M. N. (1993). *Heat Conduction*. John Wiley & Sons.
- [15] Ruggiero, M. A. G., and Lopes, V. L. R. (1997). *Cálculo Numérico: Aspectos Teóricos e Computacionais*. 2. ed., Makron Books do Brasil.

# Capítulo 11

## *APLICAÇÃO DA TÉCNICA DE ANÁLISE ENVOLTÓRIA DE DADOS E REGRESSÃO LOGÍSTICA: ANÁLISE DA EFICIÊNCIA DOS TRIBUNAIS ESTADUAIS DO BRASIL*

*Magno dos Santos Neto*

*Leandro André Cardoso de Souza*

*Pedro Paulo Gondim Simmer*

*Angela Cristina Gomes*

**Resumo:** O princípio da eficiência, consagrado na Constituição Federal de 1988 por meio da EC 19/2005, determina a busca pela melhor utilização dos recursos públicos em todas as esferas governamentais. O presente trabalho tem por objetivo medir a eficiência relativa dos Tribunais Estaduais do Brasil entre os anos de 2009 e 2014. Os dados foram analisados primeiramente aplicando-se a técnica Análise Envoltória de Dados para a análise da eficiência relativa dos Tribunais. Verificou-se que os Tribunais Estaduais dos Estados do Rio de Janeiro e Rio Grande do Sul tiveram 100% de eficiência dentro do período em análise, os demais oscilaram dentro do período de estudo. Posteriormente, verificou-se através de modelos de Regressão Logística quais variáveis contribuem para que um tribunal seja ou não classificado como eficiente. Concluiu-se que o aumento da quantidade de servidores, magistrado e a despesa total de cada tribunal não contribuem para o aumento da eficiência dos tribunais estaduais.

**Palavras Chave.** Análise Envoltória de Dados, Eficiência, Judiciário.

## 1. INTRODUÇÃO

Nas democracias contemporâneas, O sistema judiciário ocupa uma posição central entre as instituições públicas, autônoma e independente de seus governantes. Por essa razão sua eficiência é importante para garantir os direitos dos cidadãos.

A preocupação com o uso dos recursos públicos tem crescido, aumentando a pressão pública pela eficiência. [Tanzi; Schucknecht 2000], [Heller 2003], [Joumard et al. 2004]. A sociedade, cada vez mais, exige eficiência do Poder Judiciário Brasileiro. A Constituição Federal, em seu art. 37, por meio da emenda constitucional n.19/98 estabelece: A administração pública direta e indireta de qualquer dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios obedecerá aos princípios da legalidade, impessoalidade, moralidade, publicidade e eficiência.

O surgimento do princípio da eficiência no setor público abriu caminho para um maior investimento na gestão, buscando uma mudança em seus processos. Essa denominada Reforma do Judiciário procurou modificar o enfoque burocrático para um enfoque gerencial [Motta, 2010; Nogueira, 2010].

Nas organizações públicas a função essencial é a prestação de serviços com a finalidade de alcance do bem comum. No Poder Judiciário, o objetivo é a prestação de uma boa jurisdição, solucionando conflitos. Apesar disso, encontram-se diversas circunstâncias adversas, que dificultam a obtenção dessa finalidade, gerando morosidade e queda na qualidade dos serviços, por conseguinte, frustrando a sociedade.

Segundo [Boaventura Santos 2001], após a promulgação da democrática Constituição Federal de 1988 houve um crescimento da demanda pelo poder judiciário. A simplificação dos procedimentos por parte dos Juizados Especiais incentivou o acesso à justiça, além disso, diversas mudanças legislativas, como o Código de Defesa do Consumidor de 1990, contribuíram para o aumento da demanda por parte da sociedade.

Com o objetivo de obter maior eficiência na administração pública, no caso específico do Poder Judiciário Brasileiro, foi criado pela Emenda Constitucional n.45/2004 o Conselho

Nacional de Justiça. Mas foi através da Resolução n. 185/2013 que foi adotada uma plataforma única eletrônica nas cortes brasileiras. Dessa forma, estabeleceu-se um alinhamento das regras e procedimentos em todos os tribunais do país.

O objetivo deste trabalho é medir a eficiência dos tribunais estaduais do Brasil e identificar quais variáveis contribuem ou não na probabilidade de um tribunal ser classificado como eficiente ou não. Para tanto, é empregado o método de otimização linear Análise Envoltória de Dados (DEA) para medir a eficiência dos Tribunais Estaduais entre os anos de 2009 e 2014. A Regressão Logística é utilizada para verificar quais variáveis que influenciam a probabilidade de um tribunal ser classificado como eficiente. Com isso, podem-se observar quais variáveis são mais relevantes para que se obtenha uma melhor aplicação dos recursos públicos de forma a otimizar os gastos orçamentários para se atingir uma maior qualidade no atendimento à população.

O presente trabalho foi elaborado em quatro seções, além desta introdução. Na primeira seção são feitas considerações sobre os principais trabalhos relativos a aplicação da análise envoltória de dados no âmbito do poder judiciário. Na seção seguinte são apresentadas as metodologias utilizadas (técnica análise envoltória de dados e regressão logística). Na penúltima seção são apresentados os resultados obtidos no estudo. Por fim, a última seção é dedicada às considerações finais.

## 2. REVISÃO DE LITERATURA

O trabalho pioneiro de [Farrell 1957] constitui a base da literatura das metodologias do cálculo de eficiência. Partindo da definição de eficiência de uma firma como a capacidade de produzir o máximo possível para um dado conjunto de insumos, Farrell propôs que a eficiência consiste em dois componentes: a eficiência técnica, que reflete a capacidade de se obter o máximo de produto a partir de um dado conjunto de insumos, e a eficiência alocativa, que reflete a habilidade da utilização dos inputs na proporção ótima, dado seus preços e a tecnologia de produção.

A análise de eficiência tem o objetivo de construir um parâmetro de referência que permita a comparação entre as unidades para

classificá-las segundo o critério adotado. A medida de eficiência é relativa. Ela não possibilita comparações em valores absolutos. Ela não é construída de forma independente do seu grupo de comparação nem de forma ideal. As unidades são ditas eficientes sempre em relação ao grupo observado com o qual elas estão sendo comparadas.

Nos artigos internacionais sobre o tema em estudo, destacam-se as publicações de [Lewin, Morey e Cook 1982], [Kittelsen e Forsund 1992] e [Pedraja e Jiménez 1996].

No trabalho de [Lewin, Morey e Cook 1982], o principal objetivo foi mensurar a eficiência administrativa da Corte Criminal Superior da Carolina do Norte, utilizando a metodologia do DEA para o ano de 1972. Primeiramente, uma análise de regressão foi realizada com 97 jurisdições para obter a relação – se positiva, ou negativa – entre os inputs e ou outputs disponíveis. Assim como foi avaliado de que modo as variáveis de controle demográfico deveriam ser incluídas como inputs, e ainda que tipo de diferenciação, se população branca, se renda per capita, se rural-urbana entre outras. Como resultado permitiu-se observar que cada distrito pode ser tão eficiente como os que têm os mesmos níveis de recursos. Destes foram encontrados 11 distritos ineficientes e 19 eficientes na análise dos 30 Distritos Judiciais e, entre as jurisdições, a proporção foi de 63 Cortes ineficientes entre as 97 Cortes analisadas.

Em [Kittelsen e Forsund 1992] analisam as cortes distritais da Noruega, que atuam no menor nível do sistema judicial. Estas cortes provêm serviços judiciários para um ou mais municípios. Foi realizada uma análise transversal com o método DEA para mesurar a eficiência, calculado supondo as hipóteses de retornos constantes de escala e retornos variáveis. O artigo concluiu que a ineficiência era causada mais pela não utilização da escala ótima do que devido à ineficiência técnica propriamente dita. As Cortes em geral apresentaram altos índices de eficiência o que pode ser explicado pela elevada dimensão da aplicação, ou seja, quanto maior o número de inputs e outputs maior será o número de unidades eficientes.

O artigo de [Pedraja e Jiménez 1996] analisa as cortes superiores de litígios da Espanha no ano de 1991. Para a comparação entre as Cortes, foram realizados testes de homogeneidade para as hipóteses de

retornos constantes de escala e para as restrições nos pesos. Para o possível problema da falta de homogeneidade entre as unidades o autor roda duas regressões aplicando o modelo Tobit. Com a incorporação de informações sobre os custos com os juízes e funcionários foi possível calcular a eficiência global, a qual permite separar a eficiência alocativa e a eficiência técnica, gerando a classificação das Cortes em quatro grupos: as que possuem apenas a eficiência alocativa, as que possuem apenas a técnica, as que possuem ambas as eficiências e as que não são eficientes.

Na pesquisa de artigos sobre o tema em estudo, foram encontrados apenas dois trabalhos de autores brasileiros, ambos publicados no ano de 2012.

Segundo [Nogueira et al. 2012] o objetivo foi comparar a eficiência organizacional relativa dos Tribunais de Justiça do Brasil. Os dados foram analisados no enfoque de um modelo orientado para outputs utilizando-se a técnica de Análise Envoltória de Dados (DEA). O estudo foi de caráter exploratório e natureza quantitativa e qualitativa. O levantamento dos dados se deu através dos relatórios da Justiça em números nos anos de 2007 e 2008. Como resultado, verificou-se que ocorreu um aumento na quantidade de tribunais com o nível máximo de eficiência relativa nos anos de 2007 e 2009.

Em seu estudo [Yueng e Azevedo 2012] utilizam a Análise Envoltória de Dados para medir objetivamente a eficiência dos tribunais brasileiros. Seus dados foram coletados junto ao relatório anual da Justiça em números, dos anos de 2006 a 2010. Seus resultados apontam uma grande variação de eficiência relativa entre os tribunais. Observa-se que um grupo de tribunais constantemente apresenta um resultado fraco. Também se infere a partir dos resultados que a falta de recursos não é a principal explicação para a ineficiência dos tribunais.

### 3. METODOLOGIA

Nas próximas subseções serão mostrados os passos metodológicos que foram utilizados nesta pesquisa. Trata-se de um estudo empírico, a partir da abordagem quantitativa para a análise de dados secundários levantados.

### 3.1. DADOS DA PESQUISA

A coleta de dados foi realizada no sítio do Conselho Nacional de Justiça (CNJ), que através da Portaria CNJ Nº 216, de 19 de dezembro de 2012, regulamentou o acesso público ao banco de dados do Sistema de Estatística do Poder Judiciário. A base de dados compreende ao período entre o ano de 2009 e 2014.

O universo desta pesquisa é formado pelos 27 tribunais estaduais do Brasil. Buscou-se verificar o nível de eficiência e quais variáveis afetam a probabilidade de um tribunal ser considerado eficiente.

As variáveis independentes consideradas no presente estudo foram: casos pendentes, casos novos, quantidade de magistrados, quantidade de servidores, despesa total de cada tribunal (excluída a despesa com o pessoal inativo uma vez que os servidores inativos não contribuem para a produtividade dos tribunais), quantidade de processos baixados, despesa com aquisições em tecnologia de informação e comunicação da justiça estadual, despesa com custeio da tecnologia de informação e comunicação da justiça estadual, despesa com contratos da tecnologia de informação e comunicação da justiça estadual e total de despesas com tecnologia da informação.

As variáveis em questão, representam a soma de todas as instâncias dos tribunais (Primeiro Grau Comum, Juizados Especiais, Turmas Recursais e Segundo Grau).

A variável dependente binária, eficiência, para a estimação do modelo de regressão logística, será obtida através na aplicação da técnica Análise Envoltória de Dados (seção 4.2).

### 3.2. DESCRIÇÃO DA ANÁLISE ENVOLTÓRIA DE DADOS

A Análise Envoltória de Dados (Data Envelopment Analysis) é uma das ferramentas mais usuais em análises de fronteira. Esse tipo de análise realiza, por meio de um conjunto de ferramentas matemáticas, a avaliação comparativa dos resultados de um

grupo de tomadores de decisão (Decision Making Unit ou DMU). A partir disso, são estabelecidas fronteiras de eficiência entre as DMUs avaliadas [Shaw, 2009].

Segundo [Farrel 1957], o DEA é uma técnica de programação linear que teve seu surgimento na década de 50. Essa metodologia, entretanto, somente foi operacionalizada nas décadas seguintes, isto é, 70 e 80, em que houve a proposição de se medir a eficiência de processos por meio de análise de insumos (inputs) e produtos (outputs) não paramétricos de um processo [Charnes; Cooper; Rhodes 1978]; [Banker; Charnes; Cooper, 1984]. Dessa forma, o DEA passa a ser uma técnica a fim de analisar a eficiência das unidades produtivas, em que, de uma forma inicial, foi desenvolvida apenas nas linhas de produção das empresas. Nota-se, contudo, uma expansão dessa metodologia em uma gama de aspectos das diversas organizações e unidades sociais [Shaw, 2009].

Para [Banker e Morey 1986], o princípio básico do DEA é realizar uma análise das interações entre insumos e produtos das DMUs objetos de estudo. Isso porque todas as atividades ou processos envolvem certo tipo de transformação, ou seja, acrescentam-se materiais ou ideologias e ocorre uma transformação para atender às necessidades dos clientes. Assim, essa transformação envolve insumos, tais como trabalho, recursos e energia, e gera produtos ou serviços acabados que vêm a satisfazer certas demandas de clientes ou da sociedade.

Segundo [Mainardes, Alves e Raposo 2012], os inputs e outputs são o que permitem a avaliação do desempenho das operações e que as ajudam a se tornar mais produtivas e eficientes. Essa avaliação de desempenho das DMUs torna-se uma ferramenta de importância dentro do cenário atual das organizações. Isso porque, a partir dela, revelam-se pontos fortes e fracos das operações atuais bem como podem identificar-se ameaças e oportunidades dos processos vigentes. Dessa maneira, de acordo com [Lins et al. 2007], o problema de otimização do DEA para cada DMU analisada pode ser expresso da seguinte forma:

$$\frac{\sum j u_j Y_{jk}}{\sum i v_i X_{ik}} = \frac{u Y_k}{v X_k} \quad (1)$$

Em que

u e v são pesos ou multiplicadores;

X<sub>k</sub> são os insumos;

Y<sub>k</sub> são os produtos; e por convenção,  $\frac{u Y_k}{v X_k} \leq 1$ , o que gera índices de eficiência entre 0 e 1.

Segundo [Farrell 1957], a Análise Envoltória de Dados era mensurada por um único conjunto de pesos dos *inputs* e *outputs*. No estudo de [Charnes, Cooper e Rhodes 1978] reconheceram que, para legitimar a avaliação e encontrar uma eficiência relativa das unidades produtivas, deveriam ponderar os pesos a fim de encontrar uma forma mais favorável para comparação das unidades.

A operacionalização da Análise Envoltória de Dados pode seguir três passos: 1) definição das DMUs; 2) escolha do Método de DEA; e 3) seleção dos inputs e outputs que tenham relevância para estabelecer a eficiência relativa das DMUs [Ferreira e Gomes, 2009]. Como este trabalho tem como objetivo

verificar a influência da quantidade de processos baixados nos Tribunais Estaduais, delimitou-se cada Tribunal Estadual como uma DMU. Dessa forma, foram estudados 27 DMUs a fim de verificar a eficiência dos Tribunais Estaduais.

Na construção do DEA existem dois métodos básicos, que são o Constante Returns to Scale (CRS) e Variable Returns to Scale (VRS) [Charnes, Cooper e Rhodes, 1978]; [Banker; Charnes; Cooper, 1984]. O primeiro deles caracteriza-se pela redução de insumos mantendo o nível de produção, ou seja, orientação ao insumo. Para [Coelli et al. 1998], esse modelo pode ser expresso como:

$$Max_{\theta, \lambda \theta}, \quad \text{sujeito a: } -\theta Y \lambda \geq 0, \quad x_i - X \lambda \geq 0 \text{ e } \lambda \geq 0 \quad (2)$$

Em que:

$1 \leq \theta < \infty$  corresponde ao score de eficiência técnica bruto das DMUs;

$(\theta - 1)$  é o aumento proporcional na produção que poderia ser obtido pela i-ésima DMU, mantendo-se constante a utilização dos insumos;

y é o produto da DMU;

x é o insumo;

X é a matriz de insumo (n x k);

Y é a matriz de produtos (n x m); e

λ é o vetor de constantes que multiplica a matriz de insumos e produtos.

No presente estudo utilizou-se o método VSR, pois, segundo [Pedroso, Calmon e Bandeira 2009], a gestão de políticas públicas não supõe retornos constantes de escala.

Com relação à escolha das variáveis, tal procedimento procurou transmitir os efeitos mais importantes no desempenho dos Tribunais Estaduais. As variáveis escolhidas são demonstradas no Quadro 2.



Quadro 2 – Inputs e Outputs no DEA

| Nome      | Descrição   | Tipo   |
|-----------|---|--------|
| CPend     | Quantidade de casos pendentes de julgamento                           | Input  |
| CNovo     | Quantidade de casos novos   | Input  |
| DespTotal | Despesa total de cada tribunal excluída a despesa com pessoal inativo | Input  |
| Mag       | Quantidade de magistrados   | Input  |
| Serv      | Quantidade de servidores  | Input  |
| DespInf   | Despesa com tecnologia da informação                                  | Input  |
| Baixados  | Quantidade de processos baixados                                      | Output |

Fonte: Elaborado pelos autores.

A variável escolhida para ser o *Output* foi o Total de processos baixados, pois ela representa todo o ciclo da prestação jurisdicional (ajuizamento do processo judicial, julgamento, possíveis recursos e baixa do processo).

### 3.4. DESCRIÇÃO DOS MODELOS DE REGRESSÃO LOGÍSTICA

A Regressão Logística é uma técnica aplicada para estimar a probabilidade de um evento ocorrer e para identificar características de indivíduos ou elementos que pertencem a cada grupo definido com base em uma variável categórica [Fávero et al., 2009]. O objetivo da aplicação dos modelos de regressão logística é identificar a relação matemática entre as variáveis independentes e o estado eficiência dos tribunais estaduais brasileiros.

O método de seleção das variáveis aplicado no presente trabalho foi o *forward stepwise*. O método *stepwise* é baseado em um algoritmo estatístico que avalia a importância de cada variável independente e as inclui ou exclui do modelo segundo uma determinada regra. A importância de cada variável é definida em

termos de uma medida de significância estatística do seu coeficiente. Os parâmetros utilizados foram 5% de significância para a entrada das variáveis e 10% para a saída. Também foram testados o método do maior coeficiente Wald e o método da maior probabilidade condicional de máxima verossimilhança, os quais produziram resultados idênticos [Brito e Assaf Neto, 2008].

No presente trabalho são propostos dois modelos de regressão logística para analisar a eficiência dos tribunais estaduais brasileiros. No primeiro modelo, Modelo 1, a variável independente binária Eficiência ( $\text{sim}=1$ ,  $\text{não}=0$ ), só serão considerados eficientes os tribunais que alcançarem 100% de eficiência relativa (método DEA). Já no segundo modelo, Modelo 2, a variável independente (Eficiência) serão considerados eficientes os tribunais que alcançarem 80% ou mais de eficiência relativa.

Na pesquisa de literatura não foi encontrado nenhum estudo relativo à aplicação de regressão logística abordando o cenário do poder judiciário. Desta forma, os autores propõem os modelos, Modelo 1 e 2, conforme a equação geral a seguir:

$$\begin{aligned}
 P(X = 1) &= \text{Eficiência}_i \\
 &= \beta_0 + \beta_1 \text{CPend}_i + \beta_2 \text{CNovo}_i + \beta_3 \text{DespTotal}_i + \beta_4 \text{Mag}_i + \beta_5 \text{Serv}_i \\
 &\quad + \beta_6 \text{dinf}1_i + \beta_7 \text{dinf}2_i + \beta_8 \text{dinf}3_i + \beta_9 \text{Inf}_i + \beta_{10} P_{\text{Pequeno}_i} + \beta_{11} P_{\text{Médio}_i} + u_i
 \end{aligned} \quad (3)$$

Onde:  $P_{\text{pequeno}}$  e  $P_{\text{Médio}}$  são variáveis *dummies* que adotaram valor igual 1 se o tribunal for considerado de porte pequeno, médio ou grande, respectivamente, segundo classificação do Conselho Nacional de Justiça.

Tendo em vista o modelo acima especificado, o quadro abaixo apresenta a descrição das variáveis independentes utilizadas nos modelos.

Quadro 3 - Variáveis independentes dos modelos de regressão logística

| Parâmetro | Variável  | Descrição  |
|-----------|-----------|--|
| $\beta_1$ | CPend     | Quantidade de casos pendentes de julgamento  |
| $\beta_2$ | CNovo     | Quantidade de casos novos  |
| $\beta_3$ | DespTotal | Despesa total de cada tribunal excluída a despesa com pessoal inativo                |
| $\beta_4$ | Mag       | Quantidade de magistrados  |
| $\beta_5$ | Serv      | Quantidade de servidores   |
| $\beta_6$ | dinf1     | Despesa com aquisições em tecnologia de informação e comunicação da justiça estadual |
| $\beta_7$ | dinf2     | Despesa com custeio da tecnologia de informação e comunicação da justiça estadual    |
| $\beta_8$ | dinf3     | Despesa com contratos da tecnologia de informação e comunicação da justiça estadual  |
| $\beta_9$ | Inf       | Total de despesas com tecnologia da informação                                       |

Fonte: Elaborado pelos autores.

Espera-se que valores de  $\beta_1$  e  $\beta_2$  sejam estatisticamente menores do que zero, o que revela que um aumento na quantidade de casos pendentes de julgamento e casos novos diminuem a probabilidade de um tribunal ser considerado eficiente. Já as demais variáveis independentes, espera-se que contribuam de forma positiva para um tribunal ser considerado eficiente.

## 4. RESULTADOS

### 4.1. ANÁLISE ENVOLTÓRIA DE DADOS

Para a análise DEA, utilizou-se neste trabalho o modelo VRS. Segundo [Banker et al, 1984], o modelo VRS orientado a inputs determina a eficiência pela otimização da divisão entre a soma ponderada dos outputs e a soma ponderada dos *inputs*.

As Tabelas 2 e 3 apresentam os resultados de eficiência por Estado entre os anos de 2009 e

2014 de todas as instâncias (Primeiro Grau Comum, Juizados Especiais, Turmas Recursais e Segundo Grau) dos tribunais estaduais.

Ao analisar a evolução temporal da eficiência dos tribunais estaduais (Tabela 2), notamos que os tribunais dos Estados Rio de Janeiro e do Rio Grande do Sul mantiveram o nível de 100% de eficiência ao longo do período em análise. Pode-se notar que a eficiência do tribunal do Acre entre os anos de 2009 e 2010 atingiu os 100% de eficiência, e manteve 100% ao longo dos anos, caindo para 94% no ano de 2014. Já o tribunal do Mato Grosso do Sul começou com 100% de eficiência (ano 2009) e caiu o percentual ao longo dos anos em análise. Os demais Tribunais Estaduais tiveram grande variação ao longo do tempo no percentual de eficiência.

Tabela 2 - Eficiência nos Anos de 2009 a 2014

| Tribunal | Ano  |      |      |      |      |      |
|----------|------|------|------|------|------|------|
|          | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| TJAC     | 84%  | 100% | 100% | 100% | 100% | 91%  |
| TJAL     | 36%  | 68%  | 75%  | 80%  | 75%  | 77%  |
| TJAM     | 57%  | 53%  | 50%  | 74%  | 95%  | 59%  |
| TJAP     | 100% | 56%  | 75%  | 89%  | 100% | 100% |
| TJBA     | 69%  | 62%  | 51%  | 55%  | 48%  | 54%  |
| TJCE     | 80%  | 53%  | 57%  | 74%  | 71%  | 83%  |
| TJDF     | 67%  | 73%  | 78%  | 78%  | 82%  | 89%  |
| TJES     | 68%  | 45%  | 51%  | 49%  | 57%  | 68%  |
| TJGO     | 85%  | 66%  | 75%  | 87%  | 97%  | 100% |
| TJMA     | 61%  | 56%  | 78%  | 68%  | 75%  | 73%  |
| TJMG     | 79%  | 80%  | 71%  | 72%  | 73%  | 80%  |
| TJMS     | 100% | 100% | 100% | 100% | 82%  | 89%  |
| TJMT     | 42%  | 34%  | 42%  | 45%  | 62%  | 78%  |
| TJPA     | 100% | 82%  | 72%  | 76%  | 75%  | 78%  |
| TJPB     | 56%  | 49%  | 70%  | 59%  | 86%  | 84%  |
| TJPE     | 76%  | 67%  | 64%  | 59%  | 80%  | 68%  |
| TJPI     | 30%  | 31%  | 25%  | 31%  | 49%  | 55%  |
| TJPR     | 84%  | 95%  | 100% | 75%  | 100% | 100% |
| TJRJ     | 100% | 100% | 100% | 100% | 100% | 100% |
| TJRN     | 100% | 71%  | 77%  | 82%  | 69%  | 68%  |
| TJRO     | 72%  | 83%  | 95%  | 92%  | 83%  | 93%  |
| TJRR     | 63%  | 65%  | 56%  | 70%  | 55%  | 91%  |
| TJRS     | 100% | 100% | 100% | 100% | 100% | 100% |
| TJSC     | 81%  | 74%  | 83%  | 82%  | 81%  | 74%  |
| TJSE     | 86%  | 90%  | 73%  | 74%  | 100% | 86%  |
| TJSP     | 100% | 68%  | 91%  | 90%  | 81%  | 89%  |
| TJTO     | 69%  | 41%  | 56%  | 56%  | 67%  | 72%  |

Fonte: Elaborado pelos autores.

A Tabela 3 apresenta o ranking dos Tribunais Estaduais. Os tribunais do Rio de Janeiro e do Rio Grande do Sul ocuparam a primeira posição no período em análise (2009 a 2014), com eficiência de 100%. A partir do ano de 2010, o Tribunal do Estado do Acre também passou a ocupar a primeira, saindo da segunda posição em 2009, mas no ano de

2014 caiu para a terceira posição. O tribunal do Mato Grosso do Sul, entre os anos de 2009 e 2012, ocupou a primeira posição, mas no ano de 2013 caiu para a sétima posição. O Tribunal de São Paulo, no ano de 2009 ocupava a primeira posição. Mas entre os anos de 2010 e 2013, teve grande variação, e em 2014 ficou na quarta posição.

Tabela 3 - Ranking de Eficiências entre os anos de 2009 e 2014

| Tribunal | Ano  |      |      |      |      |      |
|----------|------|------|------|------|------|------|
|          | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| TJAC     | 4    | 1    | 1    | 1    | 1    | 3    |
| TJAL     | 18   | 10   | 7    | 7    | 9    | 10   |
| TJAM     | 15   | 18   | 16   | 11   | 3    | 17   |
| TJAP     | 1    | 16   | 7    | 4    | 1    | 1    |
| TJBA     | 10   | 15   | 15   | 16   | 18   | 19   |
| TJCE     | 6    | 18   | 13   | 11   | 11   | 7    |
| TJDF     | 12   | 8    | 5    | 8    | 6    | 4    |
| TJES     | 11   | 20   | 15   | 17   | 15   | 15   |
| TJGO     | 3    | 13   | 7    | 5    | 2    | 1    |
| TJMA     | 14   | 17   | 5    | 13   | 9    | 12   |
| TJMG     | 7    | 6    | 10   | 12   | 10   | 8    |
| TJMS     | 1    | 1    | 1    | 1    | 6    | 4    |
| TJMT     | 17   | 22   | 17   | 18   | 14   | 9    |
| TJPA     | 1    | 5    | 9    | 9    | 9    | 9    |
| TJPB     | 16   | 19   | 11   | 14   | 4    | 6    |
| TJPE     | 8    | 12   | 12   | 14   | 8    | 14   |
| TJPI     | 19   | 23   | 18   | 19   | 17   | 18   |
| TJPR     | 4    | 2    | 1    | 10   | 1    | 1    |
| TJRJ     | 1    | 1    | 1    | 1    | 1    | 1    |
| TJRN     | 1    | 9    | 6    | 6    | 12   | 16   |
| TJRO     | 9    | 4    | 2    | 2    | 5    | 2    |
| TJRR     | 13   | 14   | 14   | 12   | 16   | 3    |
| TJRS     | 1    | 1    | 1    | 1    | 1    | 1    |
| TJSC     | 5    | 7    | 4    | 6    | 7    | 11   |
| TJSE     | 2    | 3    | 8    | 11   | 1    | 5    |
| TJSP     | 1    | 11   | 3    | 3    | 7    | 4    |
| TJTO     | 10   | 21   | 14   | 15   | 13   | 13   |

Fonte: Elaborado pelos autores.

#### 4.3. REGRESSÃO LOGÍSTICA

Para analisar o efeito das variáveis independentes na variável dependente binária (eficiência), foram realizadas regressões logísticas. Assim, a variável binária eficiência (obtida a partir da aplicação do método DEA) foi regredida contra as

variáveis independentes conforme os modelos propostos na seção 4.2.

A Tabela 4 apresenta os resultados dos modelos de regressão logística. O Modelo 1 representa o caso em que os tribunais estaduais com eficiência relativa igual a 100% foram classificados como eficientes. Já o

Modelo 2 representa o caso em que os tribunais estaduais com eficiência relativa maior ou igual a 80% foram classificados como eficientes.

A partir da análise da estatística  $R^2$  (Nagelkerke), mostra que os modelos 1 e 2 explicam 63,1% e 83,7% a variabilidade da variável independente, respectivamente. Os modelos foram considerados estatisticamente significativos, uma vez que, os p-valores dos respectivos testes de hipóteses foram iguais a 0,001.

No Modelo 1, das variáveis inseridas, via método *stepwise*, apenas CPend, CNovo, Mag, Serv, Porte Pequeno e Porte médio, foram consideradas estatisticamente diferentes de zero. Já no Modelo 2, as variáveis casos pendentes, magistrados, servidores, despesa total de cada tribunal (excluída a despesa com pessoal inativo) e Despesa com custeio da tecnologia de informação e comunicação da justiça estadual são estatisticamente diferentes de zero. As demais variáveis foram consideradas iguais a zero.

Tabela 4 - Modelos de Regressão Logística Estimados

| Variável                  | Coeficientes $\beta$ |            |
|---------------------------|----------------------|------------|
|                           | Modelo 1             | Modelo 2   |
| CPend                     | -2,101*              | -8,674**   |
|                           | (1,135)              | (3,388)    |
| CNovo                     | -13,126**            | 1,101      |
|                           | (5,513)              | (12,357)   |
| Mag                       | -1,247**             | -5,489***  |
|                           | (0,544)              | (1,91)     |
| Serv                      | -1,475***            | -1,826**   |
|                           | (0,507)              | (0,91)     |
| despTotal                 | -0,302               | -23,943*** |
|                           | (1,908)              | (8,338)    |
| dinf1                     | 0,112**              | -0,072     |
|                           | (0,048)              | (0,127)    |
| dinf2                     | 0,109***             | 0,248*     |
|                           | (0,042)              | (0,148)    |
| dinf3                     | -0,109*              | 0,133      |
|                           | (0,064)              | (0,133)    |
| Inf                       | 0,113                | 0,934      |
|                           | (-0,985)             | (0,262)    |
| Porte_Pequeno             | 23,585****           | -7,054     |
|                           | (6,476)              | (7,002)    |
| Porte_Medio               | 19,111***            | -7,938     |
|                           | (5,182)              | (5,315)    |
| Constante                 | -24,715****          | 7,800      |
|                           | (6,863)              | (7,709)    |
| Verossimilhança de log -2 | 91,624               | 38,495     |
| R2 Cox & Snell            | 0,545                | 0,527      |
| R2 Nagelkerke             | 0,631                | 0,837      |

Fonte: Elaborado pelos autores.

Obs.: \*\*\*, \*\*, \*, significante a 1%, 5% e 10%.

Ao analisar os sinais dos  $\beta$ s, estimados no Modelo 1, constate-se que cinco foram iguais aos esperados para as variáveis: quantidade de casos pendentes de julgamento, quantidade de casos novos, quantidade de processos baixados, despesa com custeio da

tecnologia de informação e comunicação da justiça estadual e total de despesas com tecnologia da informação. Já o modelo 2, seis  $\beta$ s estimados foram iguais aos esperados para as variáveis: quantidade de casos pendentes de julgamento, despesa total de

cada tribunal (excluída a despesa com pessoal inativo), quantidade de processos baixados, despesa com custeio da tecnologia de informação e comunicação da justiça estadual, despesa com contratos da tecnologia de informação e comunicação da justiça estadual, total de despesas com tecnologia da informação. As variáveis quantidade de casos pendentes de julgamento, quantidade de processos baixados, despesa com custeio da tecnologia de informação e comunicação da justiça estadual e total de despesas com tecnologia da informação, tiveram os sinais observados iguais nos dois modelos estimados e iguais aos esperados.

O efeito de cada variável independente do modelo de regressão pode ser descrito por meio da análise dos coeficientes estimados ( $\beta$ ). Assim, no Modelo 1, os sinais negativos das variáveis casos pendentes, casos novos, magistrados, servidores e despesa com contratos da tecnologia de informação e comunicação da justiça estadual, nos mostra que quanto maior as respectivas variáveis, menor será a probabilidade de um tribunal se tornar eficiente; já as variáveis quantidade de processos baixados, despesa com aquisições em tecnologia de informação e comunicação, despesa com custeio da tecnologia de

informação e comunicação, mostra que um aumento nas respectivas variáveis, aumentará a probabilidade de um tribunal se tornar eficiente.

No Modelo 2, um aumento nas variáveis casos pendentes, magistrados, servidores e total de despesa total de cada tribunal excluída a despesa com pessoal inativo, diminui a probabilidade de um tribunal se tornar eficiente. Em sentido oposto, um aumento nas variáveis quantidade de processos baixados e despesa com custeio da tecnologia de informação e comunicação, aumentam a probabilidade de um tribunal ser eficiente.

As variáveis *dummies*, porte do tribunal (Porte Pequeno e Médio), foram consideradas estatisticamente significante, Porte Pequeno e Médio, apenas no Modelo 1.

Para avaliar a capacidade de predição da regressão logística construiu-se uma matriz de classificação que mostra a quantidade de tribunais classificados de forma correta e incorreta pelo modelo estimado. A partir dessa matriz podemos calcular os erros do tipo I e do tipo II, que correspondem a classificar um tribunal não eficiente como eficiente e um tribunal eficiente como não eficiente, respectivamente.

Tabela 6 - Matriz de classificação dos tribunais – Modelo 1

| Grupo observado | Grupo predito |     | Total | % de Acerto |        |
|-----------------|---------------|-----|-------|-------------|--------|
|                 | Eficiência    |     |       |             |        |
|                 | Não           | Sim |       |             |        |
| Eficiência      | Não           | 15  | 1     | 16          | 93,80% |
|                 | Sim           | 2   | 9     | 11          | 81,80% |
| Total           | 17            |     | 10    | 27          | 88,90% |

Fonte: Elaborado pelos autores

Tabela 7 – Matriz de classificação dos tribunais – Modelo 2

| Grupo observado | Grupo predito |     | Total | % de Acerto |        |
|-----------------|---------------|-----|-------|-------------|--------|
|                 | Eficiência    |     |       |             |        |
|                 | Não           | Sim |       |             |        |
| Eficiência      | Não           | 15  | 4     | 19          | 78,95% |
|                 | Sim           | 2   | 5     | 7           | 71,43% |
| Total           | 17            |     | 9     | 26          | 76,92% |

Fonte: Elaborado pelos autores

Conforme se observa nas Tabelas 6 e 7, os modelos 1 e 2 apresentaram maior percentual de acerto nos tribunais não eficientes (93,8% e 78,95%, respectivamente). Já para os tribunais considerados eficientes, os percentuais de acertos foram de 81,8% (Modelo 1) e 71,43% (Modelo 2). Pode-se estimar, portanto, que o percentual de acerto do modelo 1 é de 88,9%, e o Modelo 2, 76,92%.

Pode-se inferir, a partir dos resultados dos modelos de regressão logística, que aumento do quantitativo de magistrados e servidores não contribuem para um tribunal ser considerado eficiente. Por outro lado, o incremento nas despesas com tecnologia da informação contribui diretamente para a eficiência.

## 5. CONSIDERAÇÕES FINAIS

O presente trabalho teve por objetivo utilizar as técnicas Análise Envoltória de Dados e Regressão Logística para: i) apresentar uma análise da eficiência organizacional dos tribunais estaduais do Brasil entre os anos de 2009 e 2014 e, ii) identificar quais variáveis influenciam a probabilidade de um tribunal ser classificado como eficiente. Os dados utilizados foram obtidos no sítio do Conselho Nacional de Justiça.

Os resultados encontrados pela aplicação da técnica DEA, mostram que entre os anos de 2009 e 2014, apenas os tribunais estaduais

dos Estados do Rio de Janeiro e Rio Grande Sul mantiveram 100% de eficiência relativa. A eficiência relativa dos demais tribunais oscilou durante o período estudado. Dos 27 tribunais, 11 atingiram 100% de eficiência relativa, ao menos uma vez durante no período estudado.

Os resultados encontrados nos modelos de regressão logística evidenciam que: quando um tribunal é considerado eficiente, no caso da eficiência relativa obtida ser igual a 100%, as variáveis casos pendentes, casos novos, magistrados, servidores e despesa com contratos da tecnologia de informação e comunicação, reduzem a probabilidade de um tribunal ser classificado como eficiente; já as variáveis processos baixados, despesa com aquisições em tecnologia de informação e comunicação, despesa com custeio da tecnologia de informação e comunicação, aumentam a probabilidade de um tribunal ser classificado como eficiente. No caso de um tribunal ser classificado como eficiente, se a eficiência relativa for maior ou igual a 80%, as variáveis quantidade de processos baixados e despesa com custeio da tecnologia de informação e comunicação aumentam a probabilidade de um tribunal ser eficiente.

Finalmente, como sugestão para futuros trabalhos, é possível continuar analisando o efeito das variáveis independentes, utilizadas neste trabalho, sobre a variável independente binária (Eficiência), utilizando outras técnicas estatísticas como regressão logística com dados em painel para analisar também o efeito tempo.

## REFERÊNCIAS

- [1] Banker, R. D., & Maindiratta, A. (1986). Piecewise loglinear estimation of efficient production surfaces. *Management Science*, 32(1), 126-135.
- [2] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078-1092.
- [3] Banker, R. D., & Morey, R. C. (1986). Efficiency analysis for exogenously fixed inputs and outputs. *Operations research*, 34(4), 513-521.
- [4] Brito, G. A. S., & Neto, A. A. (2008). Modelo de classificação de risco de crédito de empresas. *Revista Contabilidade & Finanças*, 19(46), 18-29.
- [5] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making

units. *European journal of operational research*, 2(6), 429-444.

- [6] Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Springer Science & Business Media.

- [7] Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253-290.

- [8] FAVEIRO, L., BELFIORE, P., SILVA, F., & CHAM, B. (2009). *Análise de dados: modelagem multivariada para tomada de decisão*. São Paulo: Campus.

- [9] Ferreira, C. M. C., & Gomes, A. P. (2009). *Introdução à Análise Envoltória de Dados: Teoria. Modelos e*. (2009).

- [10] Jourmard, I., Kongsrud, P. M., Nam, Y. S., & Price, R. W. (2004). Enhancing the Effectiveness of Public Spending.
- [11] Kittelsen, S. A., & Førsund, F. R. (1992). Efficiency analysis of Norwegian district courts. *Journal of Productivity Analysis*, 3(3), 277-306.
- [12] Lewin, A. Y., Morey, R. C., & Cook, T. J. (1982). Evaluating the administrative efficiency of courts. *Omega*, 10(4), 401-411.
- [13] Lins, M. E., Lobo, M. D. C., Silva, A. C. M. D., Fiszman, R., & Ribeiro, V. D. P. (2007). O uso da Análise Envoltória de Dados (DEA) para avaliação de hospitais universitários brasileiros. *Cienc Saude Coletiva*, 12(4), 985-98.
- [14] GONÇALVES, K. (2000). *Direito Constitucional Didático*. São Paulo: Saraiva.
- [15] Mainardes, E. W., Raposo, M., & Alves, H. (2012). O desempenho das universidades públicas portuguesas segundo seus alunos: análise de eficiência por meio do Data Envelopment Analysis. *Revista Gestão Universitária na América Latina-GUAL*, 5(1), 184-215.
- [16] Pedraja-Chaparro, F., & Salinas-Jiménez, J. (1996). An assessment of the efficiency of Spanish Courts using DEA. *Applied Economics*, 28(11), 1391-1403.
- [17] Pedroso, M. D. M., Calmon, P. C. D. P., & Bandeira, L. F. (2009). O uso da análise envoltória de dados para avaliação da gestão do Programa Bolsa-Família. *Comun. ciênc. saúde*, 20(1), 37-44.
- [18] Nogueira, J. M. M., Oliveira, K. M. M., Vasconcelos, A. P., & Oliveira, L. G. L. (2012). Estudo exploratório da eficiência dos Tribunais de Justiça estaduais brasileiros usando a Análise Envoltória de Dados (DEA). *Revista de Administração Pública*, Rio de Janeiro, 46(5), 1317-1340.
- [19] Santos, B. D. S. (2001). *Direito e democracia: a reforma global da justiça. A teia global: movimentos sociais e instituições*. Porto: Edições Afrontamento.
- [20] Shaw, E. H. (2009). A general theory of systems performance criteria. *International journal of general systems*, 38(8), 851-869.
- [21] Tanzi, V., & Schuknecht, L. (2000). *Public spending in the 20th century: A global perspective*. Cambridge University Press.
- [22] Yeung, L. L. T., & Azevedo, P. F. D. (2012). Além dos "achismos" e das evidências anedóticas: medindo a eficiência dos tribunais brasileiros. *Economia Aplicada*, 16(4), 643-663.



# Capítulo 12

## *IMPROVING THE COMPUTATIONAL EFFICIENCY OF TEACHING-LEARNING BASED OPTIMIZATION FOR JOB SHOP SCHEDULING PROBLEMS BY ELIMINATING UNNECESSARY OBJECTIVE FUNCTION CALLS*

*Leonardo Ramos Rodrigues*

*João Paulo Pordeus Gomes*

*Saulo Anderson Freitas de Oliveira*

*Lucas Sousa de Oliveira*

*Takashi Yoneyama*

**Abstract:** This paper presents a method to improve the computational efficiency of TLBO in combinatorial optimization problems. Our goal is to eliminate unnecessary calls to the objective function in order to reduce simulation time. A common approach to solve combinatorial problems with TLBO is the use of priority vectors. Each priority vector is mapped into a solution by sorting its elements. However, different vectors can be mapped into the same solution. We insert a new step, called Change Verification, which consists in comparing the solutions generated from both the current and the modified priority vectors. The objective function is called only if the solutions are different from each other. Experiments with benchmark instances of the job shop scheduling problem were carried out to evaluate the proposed method. The proposed method outperformed the original TLBO in all instances, providing an average reduction in simulation time and objective function calls of 11.7% and 12.2%, respectively.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1. INTRODUCTION

Among all metaheuristic methods available in the literature, the recently proposed Teaching-Learning Based Optimization (TLBO) algorithm is one of the most promising [Rao et al., 2011]. TLBO is a population based algorithm that simulates the teaching-learning process observed in a classroom. It has achieved remarkable performances in different optimization problems (see Sahu et al. [2015], Sapathy et al. [2013] and Rao et al. [2012]) with the advantage of the absence of hyperparameters [Rao and Patel, 2012].

In recent years, several works have proposed variants of the TLBO algorithm to improve different aspects of the original formulation. In Rao and Patel [2012], the authors proposed an elitist TLBO. The elitism preserves the best candidates of each iteration and thus prevents the loss of a possible good solution. Using more than one teacher and adapting the teaching factor are some of the modifications proposed in Rao and Patel [2013] in order to improve both the exploration and the exploitation capacity of TLBO. Chen et al. [2015] incorporated local learning and self-learning methods in the TLBO algorithm, to the same end.

Although TLBO was originally proposed for optimizing mechanical design problems, it has been successfully used in combinatorial optimization problems. In Baykasoglu et al. [2014], the authors investigated the performance of TLBO in several instances of both the flow shop and the job shop scheduling problems. For the flow shop scheduling problem, the authors compared the performance of TLBO with the Novel Particle Swarm Optimization (NPSO) and the Hybrid Particle Swarm Optimization (HPSO). For the job shop scheduling problem, the authors compared the performance of TLBO with the Beam Search Algorithm (BSA) and the GRASP algorithm, among others. The performance of TLBO was comparable to the best known performances from the literature. The authors concluded that TLBO is an effective algorithm to solve combinatorial optimization problems.

Despite the numerous variants of TLBO, no previous work addressed the problem of its computational efficiency. Since TLBO is a population based algorithm, it evaluates all candidate solutions at each iteration, which can be considerably time-consuming depending on the objective function. In this

paper, we propose a variant of the TLBO with reduced computational cost by exploring particularities of combinatorial problems. Numerical experiments with twenty benchmark instances of the job shop scheduling problem are carried out to evaluate the performance of the proposed method. Experiments using the original version of the TLBO algorithm are also carried out to establish a reference baseline. Promising results were obtained.

Scheduling problems are amongst the most challenging combinatorial optimization problems. Therefore, investigating the performance of TLBO in solving these problems can give us a good idea of its capability to solve other combinatorial optimization problems.

The remaining sections of this paper are organized as follows. Section 2 introduces the teaching-learning based optimization algorithm and describes the formulation of the job shop scheduling problem. Section 3 presents the proposed method to improve the computational efficiency of the TLBO algorithm. Section 4 illustrates the application of the proposed method in twenty benchmark instances of the job shop scheduling problem. Concluding remarks are presented in section 5.

## 2. THEORETICAL BACKGROUND

### 2.1. TEACHING-LEARNING BASED OPTIMIZATION ALGORITHM

The Teaching-Learning Based Optimization (TLBO) algorithm has been recently proposed as a novel population oriented metaheuristic algorithm based on the teaching-learning process observed in a classroom [Rao et al., 2011]. This algorithm simulates the influence of a teacher on the output of a group of students in a class. The TLBO algorithm is divided into two main parts: the Teacher Phase and the Student Phase, also known as the Learner Phase [Rao and Patel, 2013]. During the Teacher Phase, students learn from the teacher, while in the Learner Phase students learn through the interaction among themselves.

There is a solution  $X$  associated with each student, which corresponds to a possible solution for the optimization problem under consideration. Also, there is a result  $f(X)$  associated with each solution (or student), which can be obtained by evaluating the

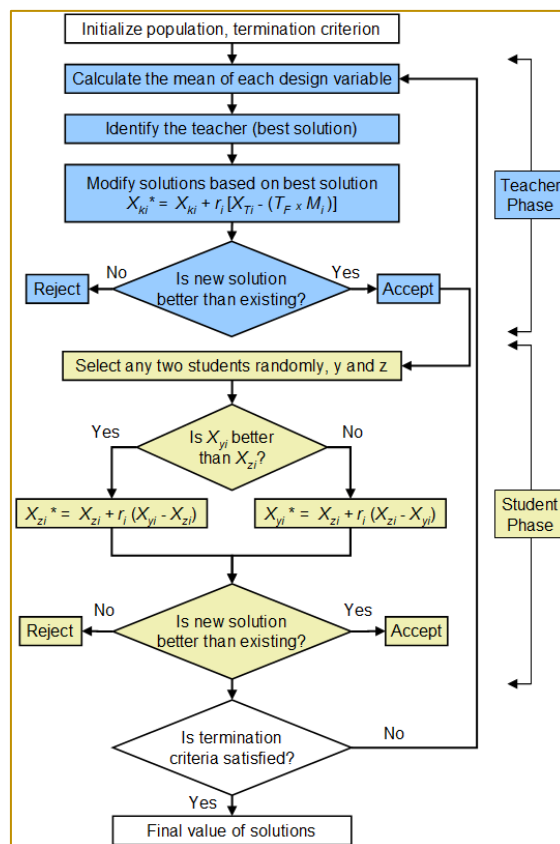
solution  $X$  using the objective function  $f(X)$ . In the job shop scheduling problem considered in this paper, a solution  $X$  corresponds to a sequence of tasks for each machine and the associated result  $f(X)$  corresponds to the makespan, which is the time spent in order to process all jobs.

At the end of each iteration, the stop criteria are checked. Different stop criteria may be adopted. Some of the most commonly adopted stop criteria are the maximum number of iterations, the maximum number of

successive iterations without any improvement, the maximum simulation time and the maximum number of function evaluations.

The student with the best solution in the population is considered as the Teacher. A flowchart describing the steps for implementing the TLBO algorithm are presented in Figure 1 [Rao et al., 2011]. The Teacher Phase and the Student Phase are described in the next sections.

Figure 1: Flowchart for the TLBO algorithm [Rao et al., 2011]



### 2.1.1. TEACHER PHASE

In this phase, the algorithm simulates the learning of the students from the teacher (best solution). During this phase, the teacher makes an effort to increase the mean result of the class.

Consider a group of  $n$  students. Let  $M_i$  be the mean solution of the students and  $T_i$  be the

teacher at iteration  $i$ . The teacher  $T_i$  will make an effort to move  $M_i$  to its own level. Knowledge is gained based on the quality of the teacher and the quality of students in the class. The difference  $D_i$  between the solution of the teacher,  $X_{Ti}$ , and the mean solution of the students,  $M_i$ , can be expressed according to Equation (1):

$$D_i = r_i(X_{Ti} - T_F \cdot M_i) \quad (1)$$

where  $r_i$  is a random number in the range [0, 1] for iteration  $i$  and  $T_F$  is a teaching factor for

iteration  $i$ , which is randomly set to either 1 or 2 according to Equation (2):

$$T_F = \text{round}(1 + \text{rand}(0,1)) \quad (2)$$

where  $\text{rand}(0,1)$  is number randomly chosen from a uniform distribution between 0 and 1, and  $\text{round}$  is a function that rounds its input to the nearest integer.

Based on the difference  $D_i$ , the existing solution of each student  $k$  in iteration  $i$ ,  $X_{ki}$ , with  $k \in \{1, 2, \dots, n\}$ , is updated in the teacher phase according to Equation (3):

$$X_{ki}^* = X_{ki} + D_i \quad (3)$$

where  $X_{ki}^*$  is the updated value of  $X_{ki}$ .

If  $f(X_{ki}^*)$  is better than  $f(X_{ki})$ ,  $X_{ki}^*$  is accepted and replaces  $X_{ki}$ . Otherwise,  $X_{ki}^*$  is discarded and  $X_{ki}$  is not changed for the next iteration.

### 2.1.2. STUDENT PHASE

In this phase, the algorithm simulates the learning of the students through interaction with one another. During this phase, students gain knowledge by discussing with other students who have more knowledge [Rao and Patel, 2013].

Consider a pair of students  $y$  and  $z$ . Let  $X_{yi}$  and  $X_{zi}$  be the solutions of students  $y$  and  $z$  in iteration  $i$ , respectively. If  $f(X_{yi})$  is better than  $f(X_{zi})$ , the solution of student  $z$  is updated according to Equation (4). Then, if  $f(X_{zi}^*)$  is better than  $f(X_{zi})$ ,  $X_{zi}^*$  is accepted and replaces  $X_{zi}$ . Otherwise,  $X_{zi}^*$  is discarded and  $X_{zi}$  is not changed for the next iteration. Similarly, if  $f(X_{zi})$  is better than  $f(X_{yi})$ , the solution of student  $y$  is updated according to Equation (5). Then, if  $f(X_{yi}^*)$  is better than  $f(X_{yi})$ ,  $X_{yi}^*$  is accepted and replaces  $X_{yi}$ . Otherwise,  $X_{yi}^*$  is discarded and  $X_{yi}$  is not changed for the next iteration.

$$X_{zi}^* = X_{zi} + r_i(X_{yi} - X_{zi}) \quad (4)$$

$$X_{yi}^* = X_{yi} + r_i(X_{zi} - X_{yi}) \quad (5)$$

During the Student Phase, a partner is randomly chosen to be paired with each student. Therefore, if a population of  $n$  students is considered, then  $n$  pairs of students are evaluated.

## 2.2. JOB SHOP SCHEDULING PROBLEM

The job shop scheduling problem is one of the best known and most difficult problem in the scheduling area [Cruz-Chavez et al., 2007]. It is classified into the NP-complete group [Garey et al., 1976].

A job shop scheduling problem consists of a set of  $j$  jobs and  $m$  machines. Each job

contains an ordered list of operations, and each operation must be processed on a specific machine. Each job must go through the machines in the specified sequence to be completely processed [Baykasoglu et al., 2014]. The operation belonging to the  $j$ -th job which is executed on machine  $m$  is denoted by  $u(m, j)$  and has a duration  $d(m, j)$ . The goal is to find a schedule that minimizes the makespan (the completion time of all jobs). The job shop scheduling problem considers the following constraints on jobs and machines [Baykasoglu et al., 2014], [Blazewicz et al., 1996]:

All jobs are independent, and available for processing at time zero.

Each job can be performed only on one machine at a time.

Processing times are deterministic and independent.

Setup times and removal times are included in processing times.

Transportation times are negligible.

Operations cannot be interrupted (preemption is not allowed).

Every machine can execute only one operation at a time.

The execution order of the operations for each job must be respected.

All machines are always available (i.e. failures do not occur).

There are no precedence constraints among operations of different jobs.

Let  $J = \{1, \dots, j\}$  be the set of jobs,  $M = \{1, \dots, m\}$  be the set of machines and  $U = \{0, 1, \dots, (j \times m) + 1\}$  be the set of operations to be scheduled, where 0 and  $(j \times m) + 1$  are dummy initial and final operations, respectively. Also, let  $du$  and  $fu$  be real numbers that denote the duration and the finish time of operation  $u$ , respectively.

The operations are interrelated by two kinds of constraints. First, the precedence constraints, which force the execution of each operation  $u$  to start only after all predecessor operations,  $Pu$ , are completed. Second, operation  $u$  can only be scheduled if the machine it requires is idle.

Let  $A(t)$  be the set of operations being processed at time  $t \in \mathbb{R}$ , and  $\alpha_{u,m}$  be a binary variable which assumes value 1 if operation  $u$  is required to be processed on machine  $m$  and zero otherwise. A schedule may be represented by a vector of finish times  $F = [f_0, \dots, f_{(j \times m) + 1}]$ . The job shop scheduling problem can be formally stated as follows [Gonçalves et al., 2005]:

$$\text{Minimize: } f_{(j \times m) + 1} \quad (6)$$

Subjected to:

$$f_u - d_u \geq f_x, \quad \forall x \in P_u \quad (7)$$

$$\sum_{u \in A(t)} \alpha_{u,m} \leq 1, \quad m \in M; t \geq 0 \quad (8)$$

$$f_u \geq 0, \quad u = 1, \dots, (j \times m) + 1 \quad (9)$$

where Equation (6) is the objective function which minimizes the makespan, Equation (7) represents the precedence relations between operations, Equation (8) imposes that one machine can only process one operation at a time, and Equation (9) imposes non negative finishing times.

The solution representation adopted in this paper for the job shop scheduling problem is based on the priority matrix approach [Baykasoglu et al., 2014]. Let  $X_k$  be an  $m \times j$  real-number matrix which represents the solution of student  $k$  for a job shop scheduling problem containing  $m$  machines and  $j$  jobs, i.e.:

### 2.2.1. SOLUTION REPRESENTATION FOR THE JOB SHOP SCHEDULING PROBLEM

$$X_k = \begin{bmatrix} x_{11} & \dots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mj} \end{bmatrix}$$

where each element  $x_{ri}$ , with  $1 \leq r \leq m$  and  $1 \leq i \leq j$  represents the priority of the operation

belonging to the  $i$ -th job that is executed by the  $r$ -th machine.

The total number of possible schedules (considering both feasible and infeasible schedules) for a job shop scheduling problem with  $m$  machines and  $j$  jobs is  $(j!)^m$ , which makes it extremely difficult to find the optimum solution using an exhaustive search even for small instances of the problem. To reduce the search space, feasible schedules are classified into four types [Gonçalves et al., 2005]:

**Inadmissible schedules:** These schedules contain excessive idle times. Inadmissible schedules may be improved by forward-shifting operations until no excess idle times exist.

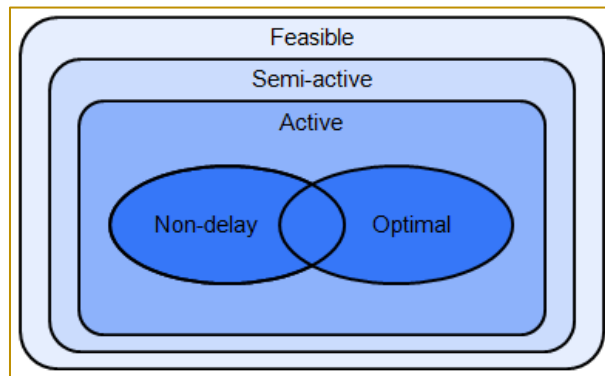
**Semi-active schedules:** These schedules can be obtained by sequencing the operations as early as possible. Semi-active schedules contain no excess idle time. However, they can be improved by shifting some operations to the front without delaying others.

**Active schedules:** In this type of schedule, no operation can be started earlier without delaying some other operation or violating a precedence constraint. An optimal schedule is always an active schedule, so the search space can be safely limited to the set of all active schedules.

**Non-delay schedules:** In this type of schedule, no machine is kept idle when it could start processing an operation. Non-delay schedules are always active.

Figure 2 shows the relationship among the different types of schedules. It is only necessary to limit the search space to the set of active schedules since optimal solutions are guaranteed to be active schedules [French, 1987]. In this paper, in order to convert the priority matrix  $X_k$  into an active schedule, the G&T algorithm is used. The G&T algorithm is briefly described below [Giffler and Thompson, 1960].

Figure 1: Relationship among schedule types



### 2.2.2. THE G&T ALGORITHM

Giffler and Thompson [1960] developed a recursive algorithm to generate active schedules systematically. The notation adopted in the G&T algorithm is as follows:

- $u(m, j)$  is the operation belonging to the  $j$ -th job that must be executed on the  $m$ -th machine.
- $S$  is the partial schedule which contains all operations already scheduled.

- $Z$  is the set of operations that can be scheduled, i.e. the set of operations without predecessor restrictions.
- $e(m, j)$  is the earliest time at which operation  $u(m, j)$  belonging to  $Z$  can be started.
- $d(m, j)$  is the duration (or execution time) of operation  $u(m, j)$ .
- $f(m, j)$  is the earliest time at which operation  $u(m, j)$  belonging to  $Z$  can be finished, i.e.  $f(m, j) = e(m, j) + d(m, j)$ .

With the notation defined, the steps to implement the G&T algorithm are described in

Algorithm 1:

**Algorithm 1** G&T algorithm

- 1: **Step 1:** Initialize  $S = \emptyset$  and  $Z$  to the subset of all operations without predecessors.
- 2: **while**  $Z \neq \emptyset$  **do**
- 3:     **Step 2:** Find the operation  $u(m,j)^+ \in Z$  with the earliest possible completion time  $f(m,j)^+$  and the machine  $m^+$  on which  $u(m,j)^+$  shall be executed.
- 4:     **Step 3a:** Identify the subset of operations  $Y \in Z$  such that all operations in  $Y$  are executed on machine  $m^+$  and for which  $e(m,j) < f^+$ .
- 5:     **Step 3b:** Choose the operation  $u(m,j)^{++}$  from  $Y$  with the largest priority value.
- 6:     **Step 3c:** Add  $u(m,j)^{++}$  to  $S$ .
- 7:     **Step 3d:** Assign  $e(m,j)^{++}$  as the starting time of operation  $u(m,j)^{++}$ .
- 8:     **Step 4:** Delete  $u(m,j)^{++}$  from  $Z$  and include its immediate successor, if any, in  $Z$ .
- 9: **end while**

To illustrate the application of the G&T algorithm, consider a job shop scheduling problem with 3 jobs and 3 machines. The

sequence of operations for each job are as follows:

$$j_1 = [(2,3), (1,7), (3,2)]$$

$$j_2 = [(1,2), (3,5), (2,4)]$$

$$j_3 = [(2,5), (3,4), (1,5)]$$

where each pair  $(m,d)$  for each job  $j_i$  represents the machine  $m$  in which the operation shall be executed and the duration  $d$  of the operation. The operations of each job must be executed in the presented order. For

instance, in this example job  $j_1$  is processed in machine  $m_2$  for 3 time units, then in machine  $m_1$  for 7 time units, and finally in machine  $m_3$  for 2 time units. Also, consider the following priority matrix:

$$X_k = \begin{bmatrix} 0.11 & 4.70 & 0.67 \\ 6.17 & 9.23 & 1.23 \\ 4.19 & 3.01 & 3.47 \end{bmatrix}$$

where each element  $x_{mj}$  represents the priority of the operation belonging to the  $j$ -th job which requires to be executed by the  $m$ -th machine.

**INITIALIZATION:**

Step 1:  $S = \emptyset$  and  $Z = \{u(2;1), u(1;2), u(2;3)\}$ .

**ITERATION 1:**

Step 2:  $f(2,1) = 3$ ;  $f(1,2) = 2$  and  $f(2,3) = 5$ . Therefore,  $f^* = 2$  and  $m^* = 1$ .

Step 3:  $Y = \{(1,2)\}$ . Operation  $u(1,2)$  is added to  $S$ , with  $e(1,2) = 0$ .

Step 4:  $u(1,2)$  is deleted from  $Z$  and  $u(3,2)$  is added to  $Z$ . For the next iteration,  $Z = \{u(2,1), u(3,2), u(2,3)\}$ .

**Iteration 2:**

Step 2:  $f(2,1) = 3$ ;  $f(3,2) = 7$  and  $f(2,3) = 5$ . Thus,  $f^* = 3$  and  $m^* = 2$ .

Step 3:  $Y = \{u(2,1), u(2,3)\}$ . Operation  $u(2,1)$  has the highest priority (6.17) and is added to  $S$ , with  $e(2,1) = 0$ .

Step 4:  $u(2,1)$  is deleted from  $Z$  and  $u(1,1)$  is added to  $Z$ . For the next iteration,  $Z = \{u(1,1), u(3,2), u(2,3)\}$ .

**ITERATION 3:**

Step 2:  $f(1,1) = 10$ ;  $f(3,2) = 7$  and  $f(2,3) = 8$ . Thus,  $f^* = 7$  and  $m^* = 3$ .

Step 3:  $Y = \{u(3,2)\}$ . Operation  $u(3,2)$  is added to  $S$ , with  $e(3,2) = 2$ .

Step 4:  $u(3,2)$  is deleted from  $Z$  and  $u(2,2)$  is added to  $Z$ . For the next iteration,  $Z = \{u(1,1), u(2,2), u(2,3)\}$ .

**ITERATION 4:**

Step 2:  $f(1,1) = 10$ ;  $f(2,2) = 11$  and  $f(2,3) = 8$ . Thus,  $f^* = 8$  and  $m^* = 2$ .

Step 3:  $Y = \{u(2,2), (2,3)\}$ . Operation  $u(2,2)$  has the highest priority (9.23) and is added to  $S$ , with  $e(2,2) = 7$ .

Step 4:  $u(2,2)$  is deleted from  $Z$ . No operation is added to  $Z$  because  $u(2,2)$  is the last operation of job  $j_2$ . For the next iteration,  $Z = \{u(1,1), u(2,3)\}$ .

**ITERATION 5:**

Step 2:  $f(1,1) = 10$  and  $f(2,3) = 16$ . Thus,  $f^* = 10$  and  $m^* = 1$ .

Step 3:  $Y = \{u(1,1)\}$ . Operation  $u(1,1)$  is added to  $S$ , with  $e(1,1) = 3$ .

Step 4:  $u(1,1)$  is deleted from  $Z$  and  $u(3,1)$  is added to  $Z$ . For the next iteration,  $Z = \{u(3,1), u(2,3)\}$ .

**ITERATION 6:**

Step 2:  $f(3,1) = 12$  and  $f(2,3) = 16$ . Thus,  $f^* = 12$  and  $m^* = 3$ .

Step 3:  $Y = \{u(3,1)\}$ . Operation  $u(3,1)$  is added to  $S$ , with  $e(3,1) = 10$ .

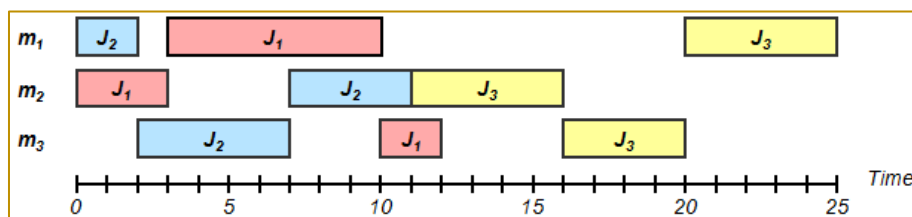
Step 4:  $u(3,1)$  is deleted from  $Z$ . No operation is added to  $Z$  because  $u(3,1)$  is the last operation of job  $j_1$ . For the next iteration,  $Z = \{u(2,3)\}$ .

**ITERATIONS 7, 8 AND 9:**

Now, all operations that have not been scheduled belong to job  $j_3$ . Therefore, in iterations 7, 8 and 9 the operations of job  $j_3$  will be scheduled following the sequence of job  $j_3$ , i.e.  $u(2,3)$ ,  $u(3,3)$  and  $u(1;3)$ .

Figure 3 shows the Gantt chart for the solution, which has a makespan of 25.

Figure 1: Gantt chart for the flow shop example

**3. PROPOSED METHOD**

When the priority vector approach is used, the priority vector associated with student  $k$ , denoted by  $X_k$ , must be converted into a job sequence  $S$  to be evaluated by the objective function, as described in section 2.2.1.

Each priority vector leads to a unique job sequence. However, two or more different priority vectors will lead to the same job sequence if their elements are in the same descending order. Consider the two priority vectors  $v_1 = [0.98, 0.23, 0.51, 0.48]$  and  $v_2 =$

$[0.67, 0.02, 0.54, 0.17]$ . Although the vectors are different, their elements are in the same descending order (i.e., the first element has the highest value, followed by the third, the fourth and the second elements). Therefore, both priority vectors  $v_1$  and  $v_2$  are mapped into the same sequence  $[1, 3, 4, 2]$ .

The main idea behind the method proposed in this paper is to insert one additional step, called Change Verification, in order to compare the job sequences obtained from the updated and the original priority vectors of each student in both the Teacher Phase and



the Student Phase. The TLBO algorithm calls the objective function only if the troubleshooting strategies differ from each other.

The goal is to eliminate unnecessary calls to the objective function. The reason is that some objective functions may have high computational costs [Hawe and Sykulski, 2006]. The Change Verification step consists in comparing two vectors. This task has a lower computational cost when compared to the computation of the makespan, which is the objective function for the job shop scheduling problem considered in this paper.

When the computational cost of the objective function is high, the computational cost saved by eliminating the unnecessary calls to the objective function compensates the

computational cost added by the Change Verification step.

#### 4. NUMERICAL EXPERIMENTS

Numerical experiments using twenty benchmark instances of the job shop scheduling problem were carried out in order to evaluate the performance of the proposed method in terms of total computational time. All the experiments reported in this paper were carried out on a personal computer with Intel Core i3, 1.9 GHz processor and 4GB RAM, running Windows 8. The algorithm was coded in Matlab. Table 1 summarizes the main characteristics of each benchmark instance used in the experiments. These instances were obtained from the OR-Library.

Table 1: Main features of the job shop benchmark instances used in the experiments

| Instance | Jobs | Machines | Instance | Jobs | Machines |
|----------|------|----------|----------|------|----------|
| abz5     | 10   | 10       | la31     | 30   | 10       |
| abz7     | 20   | 15       | la35     | 30   | 10       |
| ft06     | 6    | 6        | la36     | 15   | 15       |
| ft20     | 20   | 5        | orb05    | 10   | 10       |
| la06     | 15   | 5        | orb09    | 10   | 10       |
| la11     | 20   | 5        | swv01    | 20   | 10       |
| la12     | 20   | 5        | swv04    | 20   | 10       |
| la16     | 10   | 10       | swv05    | 20   | 10       |
| la22     | 15   | 10       | swv06    | 50   | 10       |
| la26     | 20   | 10       | yn1      | 20   | 20       |

The population size, denoted by  $PS$ , is the only parameter required by the TLBO algorithm. The stop criterion used in the experiments was the maximum number of generations, denoted by  $GM$ . The population size  $PS$  used in each benchmark instance was  $5 \times j$ , where  $j$  is the number of jobs. A maximum number of generations of 400 was used for all benchmark instances.

Table 2 shows the makespan obtained for all benchmark instances using the chosen values for  $GM$  and  $PS$ . For each benchmark instance, a Monte Carlo method with 20 iterations was carried out for both the original TLBO and the

proposed method. Column  $BKS$  shows the Best Known Solution reported in the literature for each instance. The average and the best makespan computed for each instance using both the original TLBO and proposed method are presented. The minimum values for both the average and the best makespan are shown in bold. It can be observed that both the original TLBO and the proposed method presented similar performances. This result was expected because the proposed method only eliminates the objective function calls that would not lead to better new solutions. Therefore, the quality of the final solution is not affected.

Table 2: Main features of the job shop benchmark instances used in the experiments

| Instance | BKS  | Original TLBO |      | Proposed TLBO |      |
|----------|------|---------------|------|---------------|------|
|          |      | Average       | Best | Average       | Best |
| abz5     | 1234 | 1344.3        | 1313 | 1354.9        | 1305 |
| abz7     | 656  | 831.5         | 809  | 831.8         | 818  |
| ft06     | 55   | 56.5          | 55   | 57.1          | 55   |
| ft20     | 1165 | 1231.2        | 1195 | 1227.0        | 1193 |
| la06     | 926  | 926.8         | 926  | 926.2         | 926  |
| la11     | 1222 | 1222.3        | 1222 | 1223.3        | 1222 |
| la12     | 1039 | 1042.3        | 1039 | 1039.8        | 1039 |
| la16     | 945  | 1016.8        | 995  | 1017.5        | 980  |
| la22     | 927  | 1107.9        | 1058 | 1102.1        | 1060 |
| la26     | 1218 | 1456.0        | 1397 | 1465.2        | 1410 |
| la31     | 1784 | 1950.1        | 1919 | 1952.5        | 1900 |
| la35     | 1888 | 2037.8        | 1975 | 2033.9        | 1990 |
| la36     | 1268 | 1490.3        | 1440 | 1500.3        | 1430 |
| orb05    | 887  | 966.2         | 931  | 961.9         | 933  |
| orb09    | 934  | 1011.3        | 975  | 1020.5        | 989  |
| swv01    | 1407 | 1687.4        | 1636 | 1690.0        | 1660 |
| swv04    | 1470 | 1768.1        | 1725 | 1752.0        | 1704 |
| swv05    | 1424 | 1734.3        | 1697 | 1727.8        | 1680 |
| swv06    | 2924 | 2966.6        | 2939 | 2956.8        | 2926 |
| yn1      | 884  | 1112.8        | 1098 | 1119.7        | 1094 |

Table 3 shows the average total simulation time, the average time spent running the objective function and the average number of objective function calls for both the original TLBO and the proposed method. For the latter, the average time spent running the Change Verification step is also presented.

It can be noticed that the proposed method presented a better performance for all

instances in terms of simulation time and objective function calls when compared to the original TLBO algorithm. The average reduction provided by the proposed method in total simulation time and objective function calls were 11.7% and 12.2%, respectively. Also, the average simulation time spent running the Change Verification routine is extremely low in comparison with the average total simulation time of the proposed method.

Table 3: Performance comparison between original TLBO and the proposed method in terms of computational times (in seconds) and number of objective function calls

| Instance | Original TLBO             |                                   |                          | Proposed TLBO             |                                   |                            |                          |
|----------|---------------------------|-----------------------------------|--------------------------|---------------------------|-----------------------------------|----------------------------|--------------------------|
|          | Total simulation time (s) | Obj. function simulation time (s) | Objective function calls | Total simulation time (s) | Obj. function simulation time (s) | Change simulation time (s) | Objective function calls |
| abz5     | 191.6                     | 191.0                             | 30050                    | 157.1                     | 155.7                             | 0.14                       | 25063                    |
| abz7     | 1336.6                    | 1335.3                            | 60100                    | 1237.7                    | 1233.5                            | 0.32                       | 56195                    |
| ft06     | 39.9                      | 39.5                              | 18030                    | 25.6                      | 25.1                              | 0.05                       | 11099                    |
| ft20     | 478.7                     | 477.5                             | 60100                    | 438.5                     | 435.1                             | 0.30                       | 54150                    |
| la06     | 241.7                     | 240.8                             | 45075                    | 199.5                     | 197.4                             | 0.18                       | 34956                    |
| la11     | 476.8                     | 475.7                             | 60100                    | 411.2                     | 408.0                             | 0.28                       | 51961                    |
| la12     | 475.3                     | 474.2                             | 60100                    | 423.6                     | 420.4                             | 0.27                       | 53326                    |
| la16     | 194.2                     | 193.7                             | 30050                    | 171.0                     | 169.3                             | 0.15                       | 24425                    |
| la22     | 477.8                     | 477.9                             | 45075                    | 442.9                     | 440.1                             | 0.24                       | 41504                    |
| la26     | 916.1                     | 914.8                             | 60100                    | 852.9                     | 848.9                             | 0.34                       | 56496                    |
| la31     | 2278.7                    | 2276.6                            | 90150                    | 2078.5                    | 2071.3                            | 0.57                       | 84901                    |
| la35     | 2331.4                    | 2329.2                            | 90150                    | 2101.0                    | 2093.2                            | 0.60                       | 84928                    |
| la36     | 742.2                     | 741.1                             | 45075                    | 687.2                     | 683.9                             | 0.29                       | 41572                    |
| orb05    | 187.8                     | 187.3                             | 30050                    | 157.0                     | 155.6                             | 0.12                       | 24710                    |
| orb09    | 190.2                     | 189.7                             | 30050                    | 166.4                     | 164.9                             | 0.14                       | 26248                    |
| swv01    | 960.1                     | 948.9                             | 60100                    | 874.7                     | 870.6                             | 0.35                       | 56461                    |
| swv04    | 979.7                     | 978.4                             | 60100                    | 892.4                     | 888.2                             | 0.36                       | 48991                    |
| swv05    | 948.9                     | 947.6                             | 60100                    | 889.1                     | 885.0                             | 0.34                       | 56501                    |
| swv06    | 7875.2                    | 7871.0                            | 150250                   | 7116.8                    | 7099.4                            | 1.08                       | 141677                   |
| yn1      | 1799.9                    | 1798.3                            | 60100                    | 1648.5                    | 1643.0                            | 0.44                       | 56311                    |

## 5. CONCLUSIONS

This paper presents a method to improve the computational efficiency of the TLBO algorithm in combinatorial problems in terms of simulation time. The proposed method is applicable when the priority vector approach is adopted.

Twenty different benchmark instances of the job shop scheduling problem were used in the numerical experiments to evaluate the performance of the proposed method. Simulations using the original version of the TLBO algorithm were also carried out to establish a reference baseline.

The results show that the proposed method outperformed the original TLBO algorithm in terms of simulation time and objective function calls in all benchmark instances. One

important characteristic of the proposed method is that it does not affect the final solution of the algorithm since it only eliminates those calls to the objective function that would not lead to better solutions.

The proposed method becomes more attractive as the computational cost of the objective function increases. In these situations, the computational cost saved by eliminating the unnecessary calls to the objective function becomes more relevant. The simulation time required to run the Change Verification step is very small.

A possible extension for this paper is to evaluate the performance of the proposed method in different combinatorial problems. Another possibility is to investigate the use of the Change verification step in other metaheuristic optimization algorithms.

## REFERENCES

- [1] Baykasoğlu, A., Hamzadayi, A., and Köse, S. Y. (2014). Testing the performance of teaching learning based optimization (TLBO) algorithm on combinatorial problems: Flow shop and job shop scheduling cases. *Information Sciences*, 276:204–218.
- [2] Blazewicz, J., Domschke, W., and Pesch, E. (1996). The job shop scheduling problem:

Conventional and new solution techniques. *European Journal of Operational Research*, 93:1–33.

- [3] Chen, D., Zou, F., Li, Z., Wang, J., and Li, S. (2015). An improved teaching-learning-based optimization algorithm for solving global optimization problem. *Information Sciences*, 297:171–190.

- [4] Cruz-Chavez, M. A., Martinez-Rangel, M. G., Hernandez, J. A., Zavala-Diaz, J. C., and Diaz-Parra, O. (2007). Scheduling algorithm for the job shop scheduling problem. In *Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007)*, p. 336–341.
- [5] French, S. (1987). *Sequencing and Scheduling: An Introduction to the Mathematics of the Job Shop*. Wiley & Sons, Chichester.
- [6] Garey, M. R., Johnson, D. S., and Sethi, R. (1976). The complexity of flow shop and job shop scheduling. *Mathematics of Operations Research*, 1(2):117–129.
- [7] Giffler, B. and Thompson, G. L. (1960). Algorithms for solving production-scheduling problems. *Operations Research*, 8(4):487–503.
- [8] Gonçalves, J. F., Mendes, J. J. M., and Resende, M. G. C. (2005). A hybrid genetic algorithm for the job shop scheduling problem. *European Journal of Operational Research*, 167(1):77–95.
- [9] Hawe, G. I. and Sykulski, J. K. (2006). The consideration of surrogate model accuracy in single objective electromagnetic design optimization. In *International Conference on Computational Electromagnetics (CEM)*, p. 1–2.
- [10] Rao, R. V. and Patel, V. (2012). An elitist teaching-learning based optimization algorithm for solving complex constrained optimization problems. *International Journal of Industrial Engineering Computations*, 3:535–560.
- [11] Rao, R. V. and Patel, V. (2013). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. *Scientia Iranica*, 20:710–720.
- [12] Rao, R. V., Savsani, V. J., and Vakharia, D. P. (2012). Teaching-learning-based optimization: An optimization method for continuous non-linear large scale problems. *Information Sciences*, 183:1–15.
- [13] Rao, R. V., Vakharia, D. P., and Savsani, V. J. (2011). Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43:303–315.
- [14] Sahu, B. K., Pati, S., Mohanty, P. K., and Panda, S. (2015). Teaching-learning based optimization algorithm based fuzzy-PID controller for automatic generation control of multi-area power system. *Applied Soft Computing*, 27:240–249.
- [15] Sapathy, S. C., Naik, A., and Parvathi, K. (2013). A teaching learning based optimization based on orthogonal design for solving global optimization problems. *SpringerPlus*, 2(130).

# Capítulo 13

## UTILIZAÇÃO DO MÉTODO MULTICRITÉRIO ELECTRE I COMO AUXÍLIO NA SELEÇÃO DE FORNECEDORES DE UMA EMPRESA DO RAMO CALÇADISTA

*Jordana Ramalho de Sousa*

*Thiago Yuji Eguchi*

*Maria Madalena Guerra Ferreira*

*Samuel Natividade Ferreira Junior*

*Ruy Gomes Silva*

**Resumo:** Selecionar fornecedores ideais para a empresa é uma tarefa que requer uma análise sistêmica. Os modelos de tomada de decisão multicritério exercem o papel de auxiliador para esta atividade, para que a decisão seja tomada de maneira correta. Neste sentido, este estudo visa aplicar o Método de Tomada de Decisão Multicritério, ELECTRE I, no processo de seleção de fornecedores em uma empresa do ramo calçadista, de modo que auxilie os gestores da organização para tomar a decisão. Para realização do estudo e obtenção dos resultados, avaliou-se os dados de quatro empresas fornecedoras, baseando-se em quatro critérios fornecidos pelo proprietário da empresa, com sugestões dos autores deste artigo. A utilização deste método possibilitou identificar os fornecedores que melhor atendem as necessidades da empresa, evitando qualquer tipo de preferência na escolha. Assim, o método mostra-se eficiente quando aumenta a assertividade na escolha do melhor fornecedor, gerando maior segurança na tomada de decisão.

**Palavras Chave:** ELECTRE I, Tomada de Decisão, Seleção de Fornecedores.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1. INTRODUÇÃO

A compra de suprimentos está ligada ao cotidiano de muitas organizações, talvez mais do que qualquer outra função de negócios. Se ajustando a mudanças ambientais e econômicas em todo o mundo e reconhecendo a grande importância deste setor, nenhuma empresa pode permitir que o departamento de compras fique atrás de outros departamentos [Kraljic 1983].

Sob esta perspectiva, a função compras ganha destaque nas organizações, pois esta lida diretamente com fornecedores. Destaca-se, assim, a importância estratégica das compras dentro da organização, função que deve ser corretamente gerenciada a fim de se buscar vantagens competitivas [Presoto 2012].

De Boer, Labro e Morlacchi [2001] inferem que com a crescente relevância desta função, as decisões de compra se tornam mais importantes à medida que as organizações se tornam mais dependentes dos fornecedores. Fato que implica em consequências mais graves ao se tomar uma má decisão. A partir disso, surge a importância e a necessidade de selecionar bons fornecedores, que atenderão as especificidades para que um bom produto ou serviço seja entregue ao cliente. Para tomar essa decisão, diversos métodos podem ser aplicados para auxiliar na escolha do fornecedor que atenda às necessidades da organização.

No campo das ferramentas de apoio à decisão, abordagens multicritérios estão ganhando atenção devido à sua robustez e facilidade para analisar casos mais complexos de forma eficaz. O Apoio Multicritério a Decisão possui a finalidade de dar aos tomadores de decisão as ferramentas para resolver problemas de decisão, em que diferentes pontos de vista, muitas vezes contraditórios, devem ser levados em consideração [Gonçalo e Alencar, 2014].

Os métodos de tomada de decisão da família ELECTRE têm sido abordados de maneira crescente em trabalhos científicos. Araújo e Amaral [2015], apud Ishizaka e Nemery [2013], afirmam que estes métodos têm sido

aplicados com sucesso em diversas áreas, tais como gestão ambiental, agricultura e floresta, energia, gestão da água, financiamento, os concursos, transporte e militar.

Neste contexto, o presente artigo abordará tais assuntos com o estudo de caso voltado a uma loja de calçados, localizada no centro comercial de Belém do Pará, com o objetivo de selecionar o fornecedor que melhor se adequa, no que se diz respeito aos critérios de avaliação escolhidos pelo proprietário do estabelecimento, no papel de tomador de decisão e os autores deste estudo. A escolha de tal fornecedor será feita com o auxílio do método para tomada de decisão da Escola Francesa, ELECTRE I.

## 2. REFERENCIAL TEÓRICO

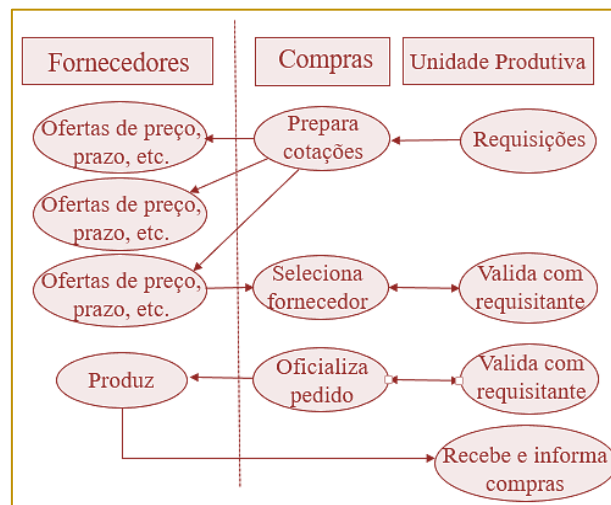
Esta seção abordará assuntos pertinentes ao desenvolvimento do estudo de caso, utilizando-se de conceitos de diversos autores, nos seguintes subtópicos, para formar o embasamento necessário que sustentará a presente pesquisa.

### 2.1 GESTÃO DE COMPRAS

A gestão de compras é definida por Lambert, Cooper e Pagh [1998] como a parte da cadeia de suprimentos responsável pela integração entre a organização em análise e seus elos a montante na cadeia. As empresas vêm se preocupando cada vez mais com o processo de compras, assim, os responsáveis do setor devem estar atentos a preço, prazo, volume e qualidade para se beneficiarem da execução eficaz deste processo [Dias 1997].

Para que haja eficácia no processo de compras de suprimentos, faz-se necessário que se compreenda as necessidades dos processos da empresa, bem como as capacitações dos fornecedores. A figura 1 a seguir demonstra uma sequência de eventos da interação da empresa com o fornecedor, que a função de compras precisa facilitar [Slack et al. 2002]:

Figura 1 - Função de compras



Fonte: Slack *et al.* (2002)

Martins e Alt [2001] afirmam que toda empresa necessita estabelecer uma interação entre todos os seus departamentos ou processos. A área de compras interage intensamente com todas as outras, havendo troca de informações úteis às suas tomadas de decisão. É necessário que os objetivos de compras estejam alinhados aos objetivos estratégicos da empresa como um todo, visando o melhor atendimento ao cliente externo e interno.

De acordo com Gaither e Frazier [2001], a missão do departamento de compras é perceber as necessidades competitivas dos produtos e serviços, sendo também fundamental no alcance dos objetivos da empresa e responsável por entregar no tempo certo, custos, qualidade e outros elementos na estratégia de operações.

As atividades relacionadas a compras envolvem uma série de fatores, como selecionar e qualificar fornecedores; avaliar o seu desempenho, negociar contratos, comparar preços, qualidade e níveis de serviço oferecidos pelos fornecedores e especificar as formas de recebimento dos bens [Senapeschi Neto 2008, apud Ballou 2001]. Assim, viabiliza-se a importância da seleção de fornecedores para uma organização, já que o resultado de uma boa escolha implica em redução de riscos e custos, aumentando assim a competitividade da organização.

## 2.2 SELEÇÃO DE FORNECEDORES

Segundo Tahriri et al. [2008], o problema de seleção de fornecedores é uma decisão que se baseia em determinação de critérios múltiplos considerados de forma coerente com as necessidades da empresa. Os critérios de seleção são pontos que vão determinar quais questões devem ser abordadas entre os fornecedores interessados na parceria ofertada. Viabilizando, assim, a importância de se desenvolver processos de seleção, os quais auxiliarão na avaliação e na decisão final da escolha dos fornecedores [Paraguassu e Macedo 2012].

Giacon [2012] afirma que na gestão moderna da cadeia de suprimentos, o processo de seleção de fornecedores é considerado uma atividade estratégica. Esta atividade exige uma avaliação abrangente de diversos critérios, não sendo baseados apenas no que se relaciona a custos, mas também de outros benefícios importantes identificados pela empresa para a decisão.

O principal objetivo da seleção de fornecedores é decidir sobre a mais apropriada fonte de fornecimento e vem acompanhada de outras discussões importantes para a empresa, como a quantidade de fornecedores que participarão da carteira de abastecimento do item e o volume de pedidos a serem alocados para cada um destes (Paraguassu e Macedo [2012], apud Weber e Current [1993]).

Segundo Lindner [2000], os compradores estão aumentando suas exigências para

poder concentrar seus esforços nas suas competências principais. Em consequência disso, a imagem das empresas frente aos seus clientes começa a ser influenciada cada vez mais pelo desempenho dos seus fornecedores. Por isso, as organizações buscam parceiros que viabilizem o atingimento de uma posição competitiva mais forte.

Escolher os fornecedores ideais para a empresa requer uma análise sistêmica e pessoas que entendam as necessidades da organização como um todo. Para auxiliar essas pessoas, que nessa atividade exercem o papel de decisor, existem os métodos de apoio multicritério a decisão, utilizados para resolver problemas em que o decisor precisa analisar um conjunto de alternativas em relação a diferentes critérios.

### 2.3 APOIO MULTICRITÉRIO A DECISÃO (AMD)

De acordo com T'Kindt e Billaut [2006], o Apoio Multicritério a Decisão (AMD) é uma abordagem conhecida como Construtivo. Não busca uma solução ótima, mas permite modelar o problema levando em conta as preferências e a experiência do tomador de decisões. Trata-se de uma abordagem flexível que, por sucessivos diálogos com o tomador de decisões, permite ao analista propor alguns elementos de resposta.

Paralelamente, o processo de auxílio à decisão contém o conjunto de elementos destacados pelo analista para ajudar o tomador de decisões. Assim, desde o início do questionamento pelo tomador de decisão, o analista pode construir modelos do problema a partir do qual pode fazer um certo número de deduções. Estes contribuem para ajudar o decisor a fazer uma escolha explícita e, portanto, tomar uma decisão [T'Kindt e Billaut 2006].

Gomes, Araya e Carignano [2011] também afirmam que os métodos do Apoio Multicritério a Decisão (AMD) ajudam na obtenção de elementos que têm como objetivo esclarecer cada decisão e recomendá-la ou favorecê-la. Trata-se de um comportamento natural que aumenta a coerência entre a avaliação do processo, os objetivos e o sistema de valores, a serviço dos quais o agente de decisão se posiciona. De acordo com Mateo [2012], o Apoio Multicritério à Decisão pode ser subdividido em métodos da Escola Francesa e Escola Americana.

Para Moreira [2007] apud Gomes [2007], a Escola Americana desenvolveu em 1970, um dos primeiros métodos para tomada de decisão multicritério, com uma metodologia baseada na análise hierárquica dos dados, o chamado método AHP (Analytic Hierarchy Process), de Thomas L.Saaty. E de acordo com Gomes, Araya e Carignano [2011], foi desenvolvido na Europa, outros métodos denominados Escola Francesa de Apoio Multicritério à Decisão. Esses métodos admitem um modelo mais flexível do problema, pois não pressupõem a comparação entre as alternativas e não impõem ao analista de decisão uma estrutura hierárquica dos critérios.

### 2.4 ELIMINATION ET CHOIX TRADUISANT LA RÉALITÉ (ELECTRE)

Segundo Gomes, Araya e Carignano [2011], os métodos ELECTRE fazem parte dos denominados Métodos de Superação, pois eles têm, como conceito teórico central, as relações de superação. Esses métodos diferenciam-se entre si pela problemática que tentam resolver, pelas informações inter e intracritérios utilizadas e pela quantidade de relações de superação construídas e pesquisadas. A família ELECTRE é composto por seis tipos de versões (ELECTRE I, II, III, IV, IS e TRI) conforme o quadro 1 a seguir:



Quadro 1 - Versões dos métodos da Família ELECTRE

| Versão | Autor           | Ano  | Tipo de Problema | Tipo de Critério | Utiliza Peso? |
|--------|-----------------|------|------------------|------------------|---------------|
| I      | Roy             | 1968 | Seleção          | Simple           | Sim           |
| II     | Roy e Bartier   | 1973 | Ordenação        | Simple           | Sim           |
| III    | Roy             | 1978 | Ordenação        | Pseudo           | Sim           |
| IV     | Roy e Hugonnard | 1982 | Ordenação        | Pseudo           | Não           |
| IS     | Roy e Skalka    | 1985 | Seleção          | Pseudo           | Sim           |
| TRI    | Yu Wei          | 1992 | Classificação    | Pseudo           | Sim           |

Fonte: Gomes, Araya e Carignano (2011)

O mesmo autor cita que os métodos ELECTRE consideram os pesos como uma medida da importância que cada critério tem para o decisor, visto que as avaliações de cada alternativa nos diferentes critérios não se reúnem em uma avaliação global. Esses métodos empregam a informação dos pesos com a finalidade de construir índices de concordância e discordância.

#### 2.4.1 ELECTRE I

Para Almeida [2013], esse método foi estabelecido para a problemática de escolha, para a qual a otimização (escolha da melhor alternativa) é um caso particular. Nessa problemática, a ideia básica é reduzir o tamanho do conjunto de alternativas  $A$ , para um subconjunto de  $A$ , com o menor número possível de alternativas.

Na primeira fase se estabelece a construção das relações de sobreclassificação, tendo

como resultado uma matriz com a comparação par a par entre as alternativas. Na segunda fase, tem-se a exploração das relações de sobreclassificação. Aplica-se um procedimento para selecionar o kernel ou subconjunto que representa a solução para o problema nessa problemática de escolha. O kernel consiste no subconjunto de alternativas  $A$  que não é sobreclassificado por nenhuma alternativa do kernel.

A primeira fase, com a construção da relação de sobreclassificação, leva às considerações que conduzem a aceitar a relação  $aSb$ , que podem ser explicitadas através de dois índices. Segundo Gomes, Araya e Carignano [2011], o índice de concordância mede os pesos dos critérios nos quais  $x_i S x_k$ , isto é, soma os pesos dos critérios pertencentes à concordância entre  $x_i$  e  $x_k$ . Calcula-se esse índice através da equação 1 seguinte:

$$C(a, b) = \sum_{\{j: g_j(a) \geq g_j(b)\}} w_j \quad (1)$$

Onde:

$g_j(a)$ : desempenho da alternativa  $a$  no critério  $j$ ;

$g_j(b)$ : desempenho da alternativa  $b$  no critério  $j$ ;

$w_j$ : peso normalizado do critério  $j$ ;

$C(a,b)$ : índice de concordância.

Por outro lado, entre os critérios podem existir alguns para os quais  $x_k S x_i$ . Isso configura um conjunto de discordância entre  $x_i$  e  $x_k$ . Pode-

se introduzir um índice de discordância, que se calcula da seguinte maneira:

$$D(a, b) = \frac{\max_{\{j: g_j(a) < g_j(b)\}} \{g_j(b) - g_j(a)\}}{\delta} \quad (2)$$

Onde:

$g_j(a)$ : desempenho da alternativa  $a$  no critério  $j$ ;

$g_j(b)$ : desempenho da alternativa  $b$  no critério  $j$ ;

$\delta$ : maior valor das amplitudes das colunas;

$D(a,b)$ : índice de discordância.

Esses dois índices não podem ser aplicados para o caso em que as avaliações das alternativas sejam qualitativas (ou com escala ordinal). Nessa situação, utiliza-se um conjunto de discordância  $D_i$  para cada critério  $i$ . Esse conjunto corresponde a pares ordenados  $(x_i, y_i)$ , tais que, quando  $g_i(a) = x_i$  e  $g_i(b) = y_i$ , então recusa-se a relação  $aSb$  [ALMEIDA, 2013].

Deve-se agora definir um limiar de concordância  $p$  e um limiar de discordância  $q$ , o que permitirá estabelecer a relação de sobreclassificação. Essas relações podem ser resumidas em uma Matriz de superação  $S$ , na qual os valores para os elementos nela contidos serão 0 ou 1, conforme for ou não observada a relação de superação. Deve-se verificar as seguintes condições:

$$aSb \text{ se, somente se } \begin{cases} C(a, b) \geq c \\ D(a, b) \leq d \end{cases}$$

Na segunda fase, tem-se o processo para selecionar o *kernel*. Uma das formas de analisar esse problema é através do uso de grafos. Três situações são indicadas no grafo, conforme a seguir:

- A seta representa a sobreclassificação de uma alternativa por outra;
- Uma ligação entre alternativas sem uma direção indica indiferença entre alternativas;
- Nenhuma ligação entre duas alternativas indica incomparabilidade.

Para aplicar o método escolhido com intuito de avaliar os fornecedores, é necessária uma escala que permita exibir uma propriedade

específica, seja um julgamento absoluto ou relativo.

## 2.5 ESCALA SAATY

Para Berzins [2009], apud Saaty [2005], a Escola Fundamental Saaty determina quantas "mais dominante" um elemento é em relação ao outro em relação a um mesmo critério. Como mecânica, a metodologia utiliza o menor elemento julgado como unidade. Os demais elementos terão um número atribuído em relação a este primeiro. Os números são determinados em termos de similitude e equivalência. Esta escala funciona similar à proporção, contudo sem ter uma unidade própria ou um zero absoluto.

Quadro 2 – Escala Fundamental Saaty

| Ponderação | Valor Normativo                          | Descrição  |
|------------|--|--|
| 1          | Igual Importância                        | As duas atividades contribuem igualmente para o objetivo.  |
| 3          | Importância pequena de uma sobre a outra | A experiência e o juízo favorecem uma atividade em relação à outra                               |
| 5          | Importância grande ou essencial          | A experiência ou juízo favorece fortemente uma atividade em relação à outra.                     |
| 7          | Importância muito grande ou demonstrada  | Uma atividade é muito fortemente favorecida em relação à outra. Pode ser demonstrada na prática. |
| 9          | Importância absoluta                     | A evidência favorece uma atividade em relação à outra, com o mais alto grau de segurança         |
| 2, 4, 6, 8 | Valores intermediários                   | Quando se procura uma condição de compromisso entre duas definições.                             |

Fonte: Almeida (2013)

### 3. PROBLEMÁTICA

A empresa pesquisada, situa-se no centro comercial da cidade de Belém no estado do Pará e atua no ramo de varejo de calçados femininos, masculinos e artigos esportivos, atuando nesse mercado a pouco mais de 10 anos. A empresa tem como foco a satisfação de seus clientes, e para isso tem como objetivo o desempenho da qualidade. A empresa busca atingir esse objetivo oferecendo produtos para seus clientes de diferentes marcas, estilos, preços competitivos e produtos que satisfaçam as necessidades de seus clientes. Dentre os tipos de produtos citados, o que possui maior procura e venda são os artigos esportivos, conforme levantamento realizado pelo proprietário do estabelecimento. Tal levantamento estatístico não será abordado neste trabalho, pois não é este o foco da pesquisa.

Como tais produtos da linha esportiva apresentam elevada demanda, os mesmos requerem maior atenção, principalmente no momento da realização da compra. Estes produtos devem estar à disposição dos clientes quando os mesmos necessitarem, pois, a falta destes poderá acarretar na perda do cliente para a concorrência. Tais produtos devem apresentar a qualidade exigida pelo cliente, bem como ter uma boa aparência e preço justo. Para isso, é necessário que sejam selecionados fornecedores que os ofereçam conforme todas as especificidades do cliente e da empresa.

Nos últimos pedidos, a empresa tem priorizado os fornecedores que oferecem os menores preços e um prazo de entrega adequado ao pedido. Entretanto, ocorrem atrasos em alguns pedidos, ocasionando transtornos para os clientes pela falta do produto. Com base nisso, o proprietário do estabelecimento percebeu que nem sempre o menor preço é fator decisório para a escolha do fornecedor.

Após estes acontecimentos, além do menor preço, o proprietário passou a considerar outros fatores, como a média do tempo de entrega dos fornecedores. Todavia, por haver muitos fornecedores, a tomada de decisão sem um instrumento matemático e de fácil uso pelo proprietário continuava sendo um fator crucial para o mesmo.

Com tais problemas citados, faz-se necessário o uso de um método multicritério de tomada de decisão, afim de diminuir as incertezas na seleção dos fornecedores e minimizar os atuais problemas enfrentados pela empresa.

### 4. RESULTADOS

A problemática focará nos quatro principais fornecedores de materiais esportivos, os quais serão avaliados de acordo com os critérios (preço, relacionamento, resolução de conflitos e qualidade) apontados pelo Gerente do Departamento de Compras. Posteriormente será atribuído o grau de importância que um critério tem em relação aos outros, através dos pesos e o

desempenho de cada fornecedor será avaliado com notas de um a nove, onde um implica em péssima performance e nove em excelente performance nos critérios que precisam ser maximizados. Para os critérios que precisam ser minimizados, a nota um significa excelente performance e a nota nove péssima.

Assim, iniciou-se o tratamento dos dados com o método ELECTRE I. Vale ressaltar que o critério preço deve ser minimizado, ou seja, quanto menor for o valor da avaliação, melhor. Os demais deverão ser maximizados, ou seja, quanto maior for o valor da avaliação, melhor. A tabela 1 contém as alternativas de fornecedores e os critérios com seus devidos pesos e notas:

Tabela 1 - Avaliação das alternativas

| Alternativa/Critério | Preço | Relacionamento | Resolução de Conflitos | Qualidade |
|----------------------|-------|----------------|------------------------|-----------|
| Fornecedor 1         | 8     | 9              | 8                      | 9         |
| Fornecedor 2         | 9     | 8              | 9                      | 8         |
| Fornecedor 3         | 6     | 7              | 7                      | 9         |
| Fornecedor 4         | 5     | 8              | 6                      | 7         |
| Pesos                | 0,40  | 0,15           | 0,20                   | 0,25      |

Fonte: Autores (2017)

Em seguida, realizou-se a normalização da tabela 1 que, de acordo com Almeida (2013), este procedimento efetua uma transformação na escala de avaliação, que passa a utilizar um intervalo (0,1). Ou seja, o elemento de menor preferência tem valor 0 e o de maior

preferência tem valor 1. Porém, o critério preço, como deseja-se minimizá-lo, deve ser transformado usando  $v'_j(a_i) = 1/v_j(a_i)$ , que mantém cardinalidade para uma escala de razão.

Tabela 2 - Normalização da avaliação das alternativas

| Alternativa/Critério | Preço  | Relacionamento | Resolução de Conflitos | Qualidade |
|----------------------|--------|----------------|------------------------|-----------|
| Fornecedor 1         | 0,2857 | 0,2813         | 0,2667                 | 0,2727    |
| Fornecedor 2         | 0,3214 | 0,2500         | 0,3000                 | 0,2424    |
| Fornecedor 3         | 0,2143 | 0,2188         | 0,2333                 | 0,2727    |
| Fornecedor 4         | 0,1786 | 0,2500         | 0,2000                 | 0,2121    |
| Pesos                | 0,40   | 0,15           | 0,20                   | 0,25      |

Fonte: Autores (2017)

Em seguida, foi utilizado a equação 4 para encontrar o índice de concordância, que mostra o quão uma alternativa é preferível à outra conforme as tabelas 3:

Tabela 3 - Índice de concordância

|              | Fornecedor 1 | Fornecedor 2 | Fornecedor 3 | Fornecedor 4 |
|--------------|--------------|--------------|--------------|--------------|
| Fornecedor 1 | —            | 0,8          | 0,6          | 0,6          |
| Fornecedor 2 | 0,2          | —            | 0,35         | 0,6          |
| Fornecedor 3 | 0,65         | 0,65         | —            | 0,45         |
| Fornecedor 4 | 0,4          | 0,55         | 0,55         | —            |

Fonte: Autores (2017)

E para encontrar o índice de discordância foi usada a equação 5, o qual exibe quando a intensidade de preferência da alternativa em

relação a outra ultrapassa um limite inaceitável como mostrado na tabela 4:

Tabela 4 - Índice de discordância

|              | Fornecedor 1 | Fornecedor 2 | Fornecedor 3 | Fornecedor 4 |
|--------------|--------------|--------------|--------------|--------------|
| Fornecedor 1 | —            | 0,2330       | 0,4997       | 0,7495       |
| Fornecedor 2 | 0,2498       | —            | 0,7495       | 1            |
| Fornecedor 3 | 0,4374       | 0,4668       | —            | 0,2498       |
| Fornecedor 4 | 0,4668       | 0,6998       | 0,4241       | —            |

Fonte: Autores (2017)

Através da análise das tabelas do índice de concordância e discordância, definiu-se o limiar de concordância ( $P = 0,5$ ) e o limiar de discordância ( $Q = 0,5$ ), que permitirá uma relação de superação. Posteriormente, obteve-se a tabela de superação, a qual a

célula com o valor 1 representa que os índices de concordância e discordância obedeceram à condição de ser maior ou igual ao limiar de concordância e menor ou igual ao limiar de discordância. Aos que não obedeceram essa regra foi atribuído o valor 0.

Tabela 5 - Tabela de superação

|              | Fornecedor 1 | Fornecedor 2 | Fornecedor 3 | Fornecedor 4 |
|--------------|--------------|--------------|--------------|--------------|
| Fornecedor 1 | —            | 1            | 1            | 0            |
| Fornecedor 2 | 0            | —            | 0            | 0            |
| Fornecedor 3 | 1            | 1            | —            | 0            |
| Fornecedor 4 | 0            | 0            | 1            | —            |

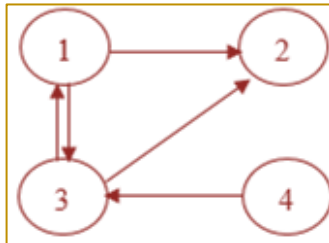
Fonte: Autores (2017)

Para concluir, fez-se a análise de grafo com o intuito de verificar qual seria a melhor opção de fornecedor, conforme a figura 2.

Através da representação da superação utilizando grafos, percebe-se que tanto o fornecedor 1 quanto o fornecedor 4 são as melhores opções de empresas fornecedoras de materiais esportivos, apontadas após

análise por meio do método ELECTRE I, pelo fato de terem maior quantidade de relação de sobreclassificação sobre as outras. Significa que essas empresas atenderam as especificações demonstradas por meio dos critérios que foram estabelecidos para este problema.

Figura 2 - Análise de grafo



Fonte: Autores (2017)

O resultado obtido neste método pode auxiliar e melhorar a maneira da empresa tomar a decisão em selecionar fornecedores. Dentre os fornecedores qualificados, para o decisor, a alternativa 4 possui maior relevância em relação aos demais, pelo fato de seu preço ser mais atrativo e gerar uma receita líquida mais satisfatória. O fornecedor 1 também se mostrou adequado às exigências de seu cliente, apesar de causar um custo mais elevado à empresa.

Caso houver necessidade de se fazer uma análise de sensibilidade da solução proposta,

pode-se alterar o limiar de concordância e o limiar de discordância e verificar o comportamento do modelo. Esse teste de sensibilidade é feito devido ao alto grau de subjetividade do método, pois as variações dos limiares modificam a relação de sobreclassificação. Em razão disso, aumentou-se o limiar de concordância para 0,6 e diminuiu-se o limiar de discordância para 0,4. Assim, foram obtidas variações nas sobreclassificações das alternativas estudadas como mostrado na tabela 6:

Tabela 6 – Tabela de Superação do Teste de Sensibilidade

|              | Fornecedor 1 | Fornecedor 2 | Fornecedor 3 | Fornecedor 4 |
|--------------|--------------|--------------|--------------|--------------|
| Fornecedor 1 | —            | 1            | 0            | 0            |
| Fornecedor 2 | 0            | —            | 0            | 0            |
| Fornecedor 3 | 0            | 0            | —            | 0            |
| Fornecedor 4 | 0            | 0            | 0            | —            |

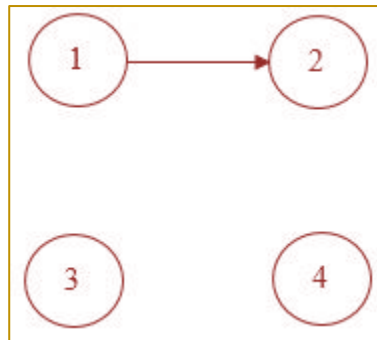
Fonte: Autores (2017)

A figura 3 representa o grafo correspondente à nova matriz de superação adquirida do teste de sensibilidade.

Com a análise de sensibilidade, verificou-se que o fornecedor 1 seria a melhor opção para a empresa por ter a maior quantidade de

sobreclassificação sobre as demais alternativas. Deste modo, mostra-se a grande importância da análise de sensibilidade para o método.

Figura 3 – Análise de Grafo do Teste de Sensibilidade



Fonte: Autores (2017)

## 5. CONCLUSÃO

Através da análise dos resultados adquiridos, percebe-se que o método multicritério ELECTRE I tem sua importância para seleção dos fornecedores. O modelo proposto identificou, pela problemática de seleção, as empresas fornecedoras de materiais esportivos que mais se destacavam em relação às outras, facilitando a tomada de decisão e com possibilidade de implementar nos demais produtos, departamentos e alternativas de expansão.

Conforme dito ao longo deste estudo, a função compras está ligada a competitividade de uma organização, na qual a seleção de fornecedores se centraliza. Para tanto, é

fundamental selecionar critérios condizentes as necessidades da empresa, de modo que auxiliem na tomada de decisão, ter capacidade de negociação e estabelecer um relacionamento com o fornecedor baseado em confiança.

Vale ressaltar que esse tipo de técnica é um suporte a tomada de decisão e selecionar fornecedores é uma decisão complexa. Em vista disto, deve-se levar em consideração todos os fatores que envolvem a organização, como variação de mercado, contratos em aberto, entre outros, para se tomar uma decisão mais assertiva, de maneira que a parceria entre comprador e fornecedor seja benéfica para ambos os lados e para os consumidores.

## REFERÊNCIAS

- [1] Acolet, T. Modelo de análise de crédito fundamentado no Electre TRI. Rio de Janeiro: Faculdades Ibmec. Dissertação de Mestrado Profissionalizante apresentada ao Programa de Pós-Graduação em Administração das Faculdades Ibmec, 2008.
- [2] Almeida, A. T. Processo de decisão nas organizações: construindo modelos de decisão multicritério. São Paulo: Atlas, 2013.
- [3] Araujo, J. J.; Amaral, T. M. Aplicação do método ELECTRE I para problemas de seleção envolvendo projetos de desenvolvimento de software livre. *Gepros. Gestão da Produção, Operações e Sistemas*, Bauru, Ano 11, nº 2, abr-jun/2016, p. 121-137.
- [4] Berzins, LORENA JACOBSON. Avaliação de Desempenho pelo AHP, através do superdecisions; Caso Inmetro. Rio de Janeiro: Faculdades Ibmec. Dissertação de Mestrado Profissionalizante apresentada ao Programa de Pós-Graduação em Administração das Faculdades Ibmec, 2009.
- [5] DE Boer, L.; Labro, E.; Morlacchi, P. A review of methods supporting supplier selection. *European Journal of Purchasing & Supply Management*, v. 7, n. 2, p. 75-89, 2001.
- [6] Dias, M. A. P. Administração de Materiais. 4 ed. São Paulo: Atlas, 1997.
- [7] Gaither, N.; Frazier, G. Administração da Produção e Operações. 8 ed. São Paulo: Pioneira, 2001.
- [8] Giacon, J. C. R. Seleção de fornecedores por análise de decisão multicritério e otimização combinatória considerando aspectos de logística e sustentabilidade. Dissertação (Mestrado), Universidade de São Paulo, São Paulo, SP, 2012.
- [9] Gomes, L. F. A. M.; Araya, M. C. G.; Carignano, C. Tomada de decisões em cenários complexos. São Paulo: Cengage Learning, 2011.
- [10] Gonçalves, T. E. E.; Alencar, L. H. A supplier selection model based on classifying its strategic impact for a company's business results. *Pesquisa Operacional*, vol.34, no.2, Rio de Janeiro May./Aug, 2014.

- [11] Kraljic, P. Purchasing must become supply management. *Harvard Business Review*, 1983.
- [12] Lambert, D.; Cooper, M.; Pagh, J. Supply chain management: implementation issues and research opportunities. *International Journal of Logistical Management*, v. 9, n. 2, p. 19, 1998.
- [13] Lindner, L. F. Z. A questão da seleção de fornecedores: um estudo de casos na indústria brasileira de autopeças. Dissertação (Mestrado em Administração), Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 2000.
- [14] Martins, P. G.; ALT, P. R. C. Administração de Materiais e Recursos Patrimoniais. Editora Saraiva, 2001.
- [15] Moreira, R. A. Análise Multicritério Dos Projetos do SEBRAE/RJ Através do ELECTRE IV. Rio de Janeiro: Faculdades Ibmec. Dissertação de Mestrado Profissionalizante Apresentada ao Programa de Pós-Graduação em Administração das Faculdades Ibmec, 2007.
- [16] Paraguassu, D. M.; Macedo, M. A. S. Seleção de fornecedores no setor público: um estudo de caso em Bio-Manguinhos/Fiocruz. XV Simpósio de Administração da Produção, Logística e Operações Internacionais, Unidade Berrini da FGV, 2012.
- [17] Presoto, A. E. Seleção de fornecedores com critérios sustentáveis na cadeia de valor varejista. Trabalho de Formatura (Graduação em Engenharia de Produção), Universidade de São Paulo, São Paulo, SP, 2012.
- [18] Senapeschi Neto, A. Gestão estratégica de compras em uma empresa do segmento de material escolar: estudo de caso longitudinal. Dissertação (Mestrado), Universidade Federal de São Carlos, São Carlos, SP, 2008.
- [19] Slack, N.; Chambers, S.; Johnston, R. Administração da produção. 2. ed. São Paulo: Atlas, 2002.
- [20] T'kindt, V.; Billaut, J-C. Multicriteria Scheduling: Theory, Models and Algorithms. New York: Springer, 2006.
- [21] Tahriri, F.; Osman, M. R.; Ali, A.; Yusuff, R. M. A review of supplier selection methods in manufacturing industries. *Suranaree J. Sci. Technol.* Vol. 15 No. 3; July - September 2008.



# Capítulo 14

## *ABORDAGEM ESTOCÁSTICA COM RECURSO EM DOIS ESTÁGIOS APLICADA A UM PROBLEMA DE PROGRAMAÇÃO LINEAR SUJEITO A INCERTEZAS*

*Solange Regina dos Santos*

*Geovana Aparecida França dos Santos*

**Resumo:** Os problemas de otimização tentam resolver de forma eficiente situações do mundo real através de modelos matemáticos que, comumente, estão associados a parâmetros incertos, como produção, demanda, custos e preços, decorrentes de erros de medição, previsão dos dados ou de caráter inerente. Na literatura encontramos diversas abordagens para incorporar e tratar estas incertezas, podendo destacar a Programação Estocástica. Tal metodologia incorpora incertezas em sua modelagem, por meio da inclusão de variáveis aleatórias com distribuição de probabilidade conhecida. Assim, buscamos nesse artigo realizar uma discussão sobre a importância de considerar as incertezas associadas a um problema de programação linear, por meio da abordagem estocástica com recurso em dois estágios.

**Palavras chave.** Otimização, Otimização Estocástica com recurso em dois estágios, Aplicação.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1. INTRODUÇÃO

Situações reais são frequentemente modeladas como problemas de otimização e estes, por sua vez, estão sujeitos a incertezas nos dados. Essas incertezas, em geral, decorrem de erros de medição ou até mesmo devido à falta de previsão de informações no momento desejado.

Nos últimos anos vários pesquisadores têm se dedicado ao desenvolvimento de abordagens destinadas ao tratamento de problemas de otimização sob incerteza. Tais abordagens podem ser classificadas em duas categorias: estocástica e robusta. A primeira delas assume que a distribuição de probabilidade das incertezas é conhecida e, por ser o foco desse estudo, será apresentada mais detalhadamente nesse artigo. Já para a segunda, informações probabilísticas não são necessárias e assume-se que as incertezas são descritas por meio de conjuntos limitados, geralmente convexos [Bertsimas e Sim, 2004].

O termo estocástico se refere aquilo cujo estado é indeterminado, com origem em eventos aleatórios, em que seu modelo possui variáveis que respondem a uma distribuição específica. Logo, segundo Bertolossi e Pagnoncelli [2006], a área de otimização estocástica estuda modelos e métodos que incorporam incertezas na modelagem através da inclusão de variáveis aleatórias com distribuição de probabilidade conhecida. O objetivo é encontrar soluções que sejam admissíveis para todas as possíveis realizações das variáveis aleatórias que são parte da modelagem, de forma a otimizar algum funcional que depende das variáveis aleatórias. Uma importante restrição dessa

metodologia é a suposição de que a distribuição de probabilidade dos dados é conhecida e não depende da decisão tomada.

Neste estudo, nos direcionamos a uma classe importante de problemas de otimização estocástica, os problemas de programação linear sujeito a incertezas com coeficientes aleatórios, em que são propostos os modelos de recursos em dois estágios. Em linhas gerais, conforme Alem e Morabito [2015], estes modelos permitem que se faça uma escolha inicial, dita de primeiro estágio, antes de se conhecer o valor de cada parâmetro incerto. Após o conhecimento dos valores dos mesmos, o agente de decisão faz novas escolhas, ditas de segundo estágio, que visam corrigir possíveis efeitos negativos gerados pela decisão do primeiro estágio. O objetivo de um modelo de programação estocástica de dois estágios é identificar uma solução de primeiro estágio que seja bem equilibrada, diante de todas as possíveis realizações das variáveis aleatórias.

## 2. OTIMIZAÇÃO ESTOCÁSTICA

Em várias aplicações, como por exemplo Mulvey et al. [1995], é comum representar as variáveis aleatórias em algum espaço de probabilidade  $(\Omega, \mathcal{F}, \Pi)$ , em que  $\Omega$  é o conjunto de possíveis estados da natureza (sendo que a realização genérica da variável aleatória é denotada por  $\omega$ ) equipado com uma  $\sigma$  – álgebra de eventos  $\mathcal{F}$  e com uma medida de probabilidade  $\Pi$ . O modelo geral linear de dois estágios com recurso pode ser escrito da seguinte maneira:

$$\begin{array}{ll}
 \text{minimizar} & c^T x + E[\min q(\omega)^T y(\omega)] \\
 \text{sujeito a} & Ax = b \\
 & T(\omega)x + W(\omega)y(\omega) \geq h(\omega) \\
 & x, y(\omega) \geq 0.
 \end{array} \tag{1}$$

No modelo (1),  $c, A$  e  $b$  são parâmetros determinísticos e definem a parte determinística do vetor de custos, da matriz tecnológica e do termo independente, respectivamente. Para cada possível realização  $\omega$ ,  $q(\omega), T(\omega), W(\omega)$  e  $h(\omega)$  definem, nessa ordem, os parâmetros estocásticos referentes ao custo, à matriz tecnológica, à matriz de recursos e ao termo independente. Além disso,  $x$  é a variável de

decisão de primeiro estágio e  $y(\omega)$  define a variável de decisão de segundo estágio, como função da realização  $\omega$ . Colocando todos parâmetros estocásticos juntos, obtém-se o valor aleatório  $\xi(\omega) \stackrel{\text{def}}{=} (q(\omega), T(\omega), W(\omega), h(\omega))$ . O valor esperado de (1) é tomado em relação à distribuição de probabilidade de  $\xi(\omega)$ , que é supostamente conhecida com expectância finita.

O modelo (1) pode ainda ser escrito como o seguinte modelo determinístico equivalente:

$$\begin{array}{ll} \text{minimizar} & c^T x + Q(x) \\ \text{sujeito a} & Ax = b \\ & x \geq 0 \end{array} \quad (2)$$

sendo que  $Q(x) = \mathbb{E}[Q(x, \xi(\omega))]$  e o valor ótimo do problema de segundo estágio é representado da seguinte forma:

$$Q(x, \xi(\omega)) = \min_{y(\omega)} \{q(\omega)^T y(\omega) | T(\omega)x + W(\omega)y(\omega) = h(\omega), y(\omega) \geq 0\}. \quad (3)$$

A representação (2) e (3) ilustra a sequência de eventos no problema de recurso. Primeiramente, são determinadas as decisões de primeiro estágio na presença de incertezas. No segundo estágio, as realizações  $\omega$  tornam-se conhecidas e as ações corretivas  $y(\omega)$  podem ser tomadas para remediar as decisões de primeiro estágio. As decisões de primeiro estágio são escolhidas, entretanto, levando em consideração seus efeitos futuros, os quais são medidos pela função recurso  $Q(x)$ .

Após estudos mais detalhados em relação aos processos estocásticos e com auxílio do software QM, analisamos aplicações de problemas clássicos de otimização estocástica disponíveis na literatura, como iremos expor a seguir.

### 3. PROBLEMA DO LOJISTA

Um lojista tem disponível determinado modelo de camiseta. Mensalmente, ele compra uma

quantia  $x$  de camiseta pelo preço  $c$  por unidade. Como possui um determinado capital para a compra de camiseta, logo obterá uma quantia limitada em um intervalo de 0 a  $u$ . As camisetas são vendidas a um preço unitário  $q$ .

O lojista possui uma incerteza em sua demanda e enfrenta um dilema: se ele comprar um número  $x$  de camiseta e a demanda for maior, ele perderá a oportunidade de vender mais camisetas. Porém, se ele comprar certa quantidade de camiseta e a demanda for menor que esta quantidade, o lojista terá um prejuízo em suas vendas, já que deve cobrir o preço de custo.

Assim, vamos supor que a demanda  $\omega$  é uma variável aleatória não-negativa com função densidade  $f$  e função de distribuição  $F$ , que  $y$  é o número de camisetas efetivamente vendidas. Sendo assim, a formulação do problema é

$$\min_{0 \leq x \leq u} \{cx + Q(x)\} \quad (4)$$

onde

$$Q(x) = \mathbb{E}_\omega[Q(x, \omega)] \quad (5)$$

e

$$\begin{array}{ll} Q(x, \omega) = & \text{minimizar} \quad -q y(\omega) + c b(\omega) \\ & \text{sujeito a} \quad y(\omega) + b(\omega) \leq x \\ & y(\omega), b(\omega) \geq 0 \end{array} \quad (6)$$

em que  $\mathbb{E}_\omega$  representa a esperança com respeito a demanda  $\omega$ . Para a quantidade  $x$  de camisetas compradas, a função  $-Q(x, \omega)$  denota o lucro obtido com a venda das camisetas com um valor  $\omega$  fixo.

Esse problema se baseia em dois estágios. No primeiro estágio o lojista precisa decidir quantas camisetas irá comprar através da variável  $x$ . Após esta decisão, ele irá vender as camisetas para uma demanda  $\omega$ . As variáveis de segundo estágio são a quantidade de camisetas que ele vendeu ( $y(\omega)$ ) e a não vendidas que permanecem no estoque ( $b(\omega)$ ). Sendo assim, o problema do lojista busca determinar a quantidade exata de camisetas que irá comprar de forma a

maximizar o lucro esperado sob incerteza de demanda.

#### 4. RESOLUÇÃO DO PROBLEMA

Para resolver esse problema, precisamos primeiramente encontrar a solução de segundo estágio, a qual é felizmente imediata: se a demanda  $\omega$  for menor do que o número de camisetas compradas, então  $y^*(\omega) = \omega$ . Se for maior, então  $y^*(\omega) = x$ . Para encontrar o valor de  $b^*(\omega)$ , basta notar que só haverá camisetas no estoque se a demanda for menor do que o número de camisetas compradas. Logo,

$$\begin{aligned} y^*(\omega) &= \min \{\omega, x\} \\ b^*(\omega) &= \max \{x - \omega, 0\}. \end{aligned} \quad (7)$$

Deste modo, podemos reescrever a formulação (5) da seguinte forma

$$Q(x) = \mathbb{E}_\omega [-q \min\{\omega, x\} + c \max\{x - \omega, 0\}]. \quad (8)$$

Logo, se desejamos encontrar uma solução ótima para este problema de otimização, podemos utilizar o método da derivada primeira da função  $Q(x)$  e analisarmos o seu sinal, buscando uma solução em seus extremos, ou em seus pontos críticos. Para isto, lembramos que a função analisada está limitada no intervalo de  $[0, u]$ . A fim de definir melhor a função  $Q(x)$ , precisamos definir um dos conceitos utilizados para se calcular a esperança de uma variável aleatória contínua.

**Definição 1:** Seja  $g$  uma função contínua e  $X$  uma variável aleatória contínua com função densidade  $f$ . Então,  $\mathbb{E}[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$ , caso a integral exista.

Assim, considerando a equação (5) e a Definição 1, podemos dizer que

$$\mathbb{E}[g(t)] = \int_{-\infty}^x g(t)f(t)dt + \int_x^{+\infty} g(t)f(t)dt. \quad (9)$$

Sabendo que a variável  $\omega$  é incerta e que  $x$  é fixo, caso  $\omega < x$ , pela equação (8) temos que

$Q(x) = \mathbb{E}_\omega [-q\omega + c(x - \omega)]$ . Porém, caso  $\omega > x$ , temos que  $Q(x) = \mathbb{E}_\omega [-qx]$ . Sendo assim, substituindo estes valores na equação (9), segue que

$$Q(x) = \mathbb{E}_\omega [Q(\omega, x)] = \int_{-\infty}^x (-qt + c(x - t))f(t)dt + \int_x^{+\infty} (-qx)f(t)dt. \quad (10)$$

Resolvendo a equação (10),

$$\begin{aligned}
 Q(x) &= \int_{-\infty}^x -qt f(t) + c(x-t)f(t) dt + \int_x^{+\infty} (-qx)f(t) dt \\
 &= -qx + (q+c) \int_{-\infty}^x F(t) dt.
 \end{aligned}
 \tag{11}$$

A função  $Q$  é definida pela soma da função distribuição acumulada da variável aleatória  $t$  com a função polinomial  $-qx$ . Portanto, os candidatos a solução do problema (4) estarão

nos extremos do intervalo  $[0, u]$  ou nos pontos críticos encontrados ao calcular  $Q'(x) = 0$ .

Assim, temos que  $Q'(x) = -q + (q+c)F(x)$ , sendo as soluções possíveis dadas por

|                                |                 |                            |                          |  |
|--------------------------------|-----------------|----------------------------|--------------------------|--|
| Extremos do Intervalo $[0, u]$ | $c + Q'(0) > 0$ | $c + [-q + (q-r)F(0)] > 0$ | $F(0) > \frac{q-c}{q-r}$ | $x^* = 0$                                  |
|                                | $c + Q'(u) < 0$ | $c + [-q + (q-r)F(u)] < 0$ | $F(u) < \frac{q-c}{q-r}$ | $x^* = u$                                  |
| Ponto Crítico                  | $Q'(x) = 0$     | $c + [-q + (q-r)F(x)] = 0$ | $F(x) = \frac{q-c}{q-r}$ | $x^* = F^{-1}\left(\frac{q-c}{q-r}\right)$ |

Resumidamente,

$$\begin{cases}
 x^* = 0, & \text{se } F(0) > \frac{q-c}{q+c} \\
 x^* = u, & \text{se } F(u) < \frac{q-c}{q+c} \\
 x^* = F^{-1}\left(\frac{q-c}{q+c}\right), & \text{caso contrário.}
 \end{cases}
 \tag{12}$$

Vimos, por hipótese, que a função  $Q$  é convexa.

**Definição 2:** Seja  $I$  um intervalo não vazio de  $\mathbb{R}$ . A função  $f: I \rightarrow \mathbb{R}$  é dita convexa em  $I$  quando

$$f(\alpha x_0 + (1-\alpha)x_1) \leq \alpha f(x_0) + (1-\alpha)f(x_1),
 \tag{13}$$

para todos os pares de pontos  $(x_0, x_1)$  em  $I$  e todo  $\alpha \in (0,1)$ .

**Resultado:** Com base na Definição 2, vamos provar que a função  $Q$  é convexa.

**Prova:** Primeiramente, tomemos um  $\alpha \in (0,1)$ , com  $a \leq x_1 < x_2 \leq b$  e  $x := \alpha x_1 + (1-\alpha)x_2$ . Consideremos,

$$\Phi := f(x) - \alpha f(x_1) - (1-\alpha)f(x_2),$$

que deve ser não-positiva ( $\Phi \leq 0$ ), para que tenhamos

$$f(x) \leq \alpha f(x_1) + (1-\alpha)f(x_2),$$

ou seja, queremos que  $f$  seja convexa. Temos

$$\Phi = f(x) - \alpha f(x_1) - (1 - \alpha)f(x_2).$$

Substituímos  $f(x)$  por  $-qx + (q + c) \int_{-\infty}^x F(t) dt$ :

$$\begin{aligned} \Phi &= -qx + (q + c) \int_{-\infty}^x F(t) dt + \alpha qx_1 - \alpha(q + c) \int_{-\infty}^{x_1} F(t) dt + (1 - \alpha)qx_2 + \\ &\quad - (1 - \alpha)(q + c) \int_{-\infty}^{x_2} F(t) dt \\ &= q(\alpha x_1 - x) + (1 - \alpha)qx_2 - (q + c) \int_x^{x_2} F(t) dt + \alpha(q + c) \int_{x_1}^{x_2} F(t) dt. \end{aligned}$$

Considere a função  $F$  não decrescente, então temos que

$$x \leq y \Rightarrow \{X \leq x\} \subset \{X \leq y\} \Rightarrow F(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq y) = F(y).$$

Além disso, temos que  $a \leq x_1 < x < x_2 \leq b$ . Como  $F$  é crescente em  $[a, b]$ , se  $u \in (x_1, x)$ , então  $F(u) \leq F(x)$ . Logo, vale

$$\begin{aligned} \Phi &= q(\alpha x_1 - x) + (1 - \alpha)qx_2 - (q + c) \int_x^{x_2} F(t) dt + \alpha(q + c) \int_{x_1}^{x_2} F(t) dt \\ &\leq q(\alpha x_1 - x) + (1 - \alpha)qx_2 - (q + c) \int_x^{x_2} F(x) dt + \alpha(q + c) \int_{x_1}^{x_2} F(x) dt \\ \Phi &\leq q(\alpha x_1 - x) + (1 - \alpha)qx_2 - (q + c)F(x)(x_2 - x) + \alpha(q + c)F(x)(x_2 - x_1). \end{aligned}$$

Substituindo  $x := \alpha x_1 + (1 - \alpha)x_2$ .

$$\begin{aligned} \Phi &\leq q(\alpha x_1 - \alpha x_1 - (1 - \alpha)x_2) + (1 - \alpha)qx_2 + \\ &\quad - (q + c)F(x)(x_2 - \alpha x_1 - (1 - \alpha)x_2) + \alpha(q + c)F(x)(x_2 - x_1) \\ &\leq -a(x_2 - x_1)(q + c)F(x) + \alpha(q + c)F(x)(x_2 - x_1) \\ \Phi &\leq 0. \end{aligned}$$

Logo,

$$f(x) - \alpha f(x_1) - (1 - \alpha)f(x_2) \leq 0.$$

## 5 UM EXEMPLO NUMÉRICO

Suponha que o custo por camiseta para o lojista seja  $c = R\$26,90$ , que o preço de venda seja  $q = R\$49,90$  e que o poder de compra seja  $u = 500$  camisetas. Além disso, considere que a demanda  $\omega$  é dada por uma

variável aleatória uniforme contínua definida no intervalo  $[0,500]$ . Por definição, uma variável aleatória uniforme possui função densidade de probabilidade dada por

$$f(x|a,b) = \frac{1}{b-a}, a \leq x \leq b \quad (14)$$

com esperança  $\mathbb{E}[X] = (a+b)/2$  e variância  $\sigma^2(X) = (b-a)^2/12$ . Logo, integrando-se a

densidade de  $\omega$ , obtemos a função de distribuição da demanda:

$$F(\omega) = \begin{cases} \frac{\omega}{500}, & \text{se } 0 \leq x \leq 500 \\ 1, & \text{se } x > 500 \\ 0, & \text{caso contrário.} \end{cases} \quad (15)$$

A inversa dessa função é  $F^{-1}(y) = 500y$  no intervalo  $[0,500]$ . Usando (1.6), temos que a solução do problema é  $x^* = F^{-1}\left(\frac{115}{384}\right) = 149,74$ . Contudo, como é um problema real e não podemos produzir 149,74 camisetas,

assumiremos  $x^* = 149$ . Assim, utilizando a formulação (1.1) e a equação (1.9) podemos dizer que o lucro esperado para este valor é

$$\begin{aligned} \mathbb{E}_\omega[cx + Q(x, \omega)] &= cx - qx + (q+c) \int_0^x \frac{\omega}{500} d\omega \\ \mathbb{E}_\omega[26,9 \cdot 149 + Q(149, \omega)] &= 4008,1 - 7435,1 + 76,8 \int_0^{149} \frac{\omega}{500} d\omega \\ &= -3427 + \frac{96}{625} \left( \frac{22201}{2} \right) = R\$-1721,96. \end{aligned} \quad (16)$$

Portanto, o lojista teria que comprar 149 camisetas mensalmente para obter um lucro esperado de  $R\$ 1721,96$ . Vamos atribuir ao valor encontrado na equação (3.3) o nome de Valor do Problema Estocástico, ou RP (*Recourse Problem* - RP). Na próxima seção, iremos mensurar o ganho por considerarmos o problema estocástico bem como a quantidade deixada de lucrar por não conhecer com exatidão o futuro.

## 6. EVPI E VSS

Existem alguns conceitos na Otimização Estocástica que servem de apoio para a compreensão e estudo dos resultados obtidos na seção anterior. Um deles é o Valor Esperado de Informação Perfeita (*Expected Value of Perfect information* - EVPI), o qual é obtido pela diferença entre o Valor do Problema Estocástico (*Recourse Problem* - RP) e o valor esperado das soluções (*Wait and See* - WS). Assim

$$EVPI = RP - WS.$$

Outra forma de tentar explicar estes termos seria de que o EVPI, segundo Birge e Louveaux [1997], mede o quanto o agente de decisão estaria disposto a pagar para se obter informação perfeita e precisa sobre o futuro. Neste problema, seria o quanto o Lojista estaria disposto a pagar para se obter com precisão o valor da demanda  $\omega$ . Para Kall e Wallace [1994], o EVPI também representa quanto se esperaria ganhar se fosse possível determinar de antemão o valor das variáveis aleatórias. Ainda na ideia destes autores, o EVPI é importante, porque mostra se é importante considerar a aleatoriedade do problema, pois não necessariamente um valor alto para o EVPI possa indicar a necessidade de se resolver o problema estocástico. Porém, quando o EVPI é baixo, temos uma indicação

de que não é tão importante considerar a aleatoriedade do problema e, portanto, aproximações podem funcionar bem.

O RP representa o valor do problema estocástico, o qual foi encontrado na seção anterior. Já o WS representa o valor ótimo do problema para cada cenário, o qual neste caso está definido no intervalo de [0,500].

Como o valor de RP já foi calculado, basta encontrar o valor de WS. Para isto, precisamos lembrar que o valor de  $\omega$  está definido no intervalo [0,500]. Assim, como o EVPI supõe que o valor do coeficiente aleatório é conhecido, para um determinado valor  $\omega$ , a solução é obviamente  $x^* = \omega$ . Logo,

$$WS = \mathbb{E}_\omega[c\omega - q\omega] = \mathbb{E}_\omega[-23\omega] = -23\mathbb{E}_\omega[\omega] = -23 \cdot \frac{0 + 500}{2} = -5750.$$

O valor da esperança de  $\omega$  foi calculado lembrando que  $E[X] = (a + b)/2$ . Por fim, o valor do EVPI será dado por

$$EVPI = R\$ 5750,00 - R\$ 1721,96 = R\$4028,04.$$

Outro valor importante em se considerar na Otimização Estocástica é o Valor da Solução Estocástica (*Value of Stochastic Solution* -

VSS), o qual é obtido pela diferença entre a Solução do Valor Esperado (*Expectation of the Expected Solution* - EEV) e o RP. Assim,

$$VSS = EEV - RP.$$

Para se encontrar o valor de EEV, inicialmente calcula-se a solução ótima do problema para  $\omega = 250$ , ou seja, com demanda constante igual a média de  $\omega$ . Em seguida, o valor de  $x^*$  encontrado é substituído na equação (3.3) para se obter um novo lucro ótimo. Novamente, como o valor de RP já foi obtido, vamos calcular o valor de EEV. Para isto,

vamos inicialmente encontrar o valor ótimo para  $\omega = 250$ , a qual é obtido facilmente, uma vez que conhecendo-se o valor de  $\omega$  basta comprar  $x^* = 250$  camisetas para maximizar o lucro. O próximo passo é usar o valor de  $x^*$  na equação (3.3). A solução após os cálculos será  $-R\$ 950,00$ . Deste modo

$$VSS = R\$ 1721,96 - R\$ 950,00 = R\$ 771,96.$$

Assim, concluímos ser essencial o modelo estocástico nesse problema, pois quando os valores do EVPI são elevados, como nesse caso  $EVPI = R\$4028,04$ , indicam a necessidade de resolver tal modelo,

confirmando essa necessidade pelo resultado do VSS, já que nos fornece o quanto estamos ganhando em considerar o modelo estocástico, ao invés de simplesmente supor que o lucro de vendas é dado por sua média.



## 7. CONSIDERAÇÕES FINAIS

O método de otimização estocástica, por meio do exemplo abordado, mostrou-se uma alternativa adequada e flexível na configuração do plano de produção, pois permitem modelar as variáveis aleatórias de forma bastante natural com a utilização de cenários que podem ser gerados com diferentes estruturas e distribuições de probabilidade. Também, a possibilidade de utilizar as variáveis de decisão de segundo estágio se apresentou uma estratégia interessante, visto que tais variáveis podem ser utilizadas para se remediar dos parâmetros estocásticos, ajustando e corrigindo decisões do primeiro estágio tomadas antes da realização das variáveis aleatórias.

Em linhas gerais, a otimização estocástica em dois estágios permite que se faça uma escolha inicial, dita de primeiro estágio, antes de se conhecer o valor de cada parâmetro incerto. Após o conhecimento dos valores dos mesmos, o agente de decisão faz novas escolhas, ditas de segundo estágio, que visam corrigir possíveis efeitos negativos

## REFERÊNCIAS

- [1] Alem, D.; Morabito, R. (2015) Planejamento da produção sob incerteza: programação estocástica versus otimização robusta. *Gestão da Produção*, 22(3): 539-551.
- [2] Bertsimas, D.; Sim, M. (2004) The price of robustness. *Operations Research*, 52(1): 35-53.
- [3] BIRGE, John R.; LOUVEAUX, Francois. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

gerados pela decisão do primeiro estágio. Além disso, essa abordagem é considerada uma metodologia muito rica, por empregar conceitos e resultados de diversas áreas como, programação linear, probabilidade e estatística.

Além disso, ao considerarmos o método estocástico em relação ao determinístico, o valor ótimo do determinístico corresponde a um único custo mínimo, o valor ótimo do estocástico refere-se a uma composição de custos mínimos, um para cada cenário, ponderados pelas probabilidades desses cenários, comumente denominado como custo mínimo esperado. Ainda, o modelo determinístico gera um único plano de produção, ao passo que o modelo estocástico de dois estágios gera plano de produção dependentes dos cenários.

## 8. AGRADECIMENTOS

Agradecemos a Fundação Araucária e a PRPPG/UNESPAR pelo incentivo ao desenvolvimento deste trabalho e o apoio financeiro concedido.

- [4] Bortolossi, H. J.; Pagnoncelli, B. K. (2006) Uma introdução à otimização sob incerteza. In: III Bienal da Sociedade Brasileira de Matemática. Universidade Federal de Goiás.
- [5] Kali, Peter; WALLACE, Stein W. (1994) *Stochastic programming*. New York: Springer.
- [6] Mulvey, J. M.; Vanderbei, R. J.; Zenios, S. A. (1995) *Operations Research. Robust optimization of large-scale systems*. 43, 264-281.

# Capítulo 15

## *A COMPARATIVE STUDY OF METAHEURISTICS APPLIED TO TROUBLESHOOTING OPTIMIZATION PROBLEMS*

*Lucas Sousa de Oliveira*

*Takashi Yoneyama*

*Leonardo Ramos Rodrigues*

**Abstract:** Metaheuristics are powerful in that they efficiently address complex, real-life-sized problems. Many algorithms have been proposed in the last few years, each one with different features, but few stand out. The purpose of this paper is to compare the recently proposed Teaching-Learning Based Optimization (TLBO) with two of the most well-known methods, namely Simulated Annealing (SA) and Genetic Algorithm (GA). Numerical experiments were conducted using three different instances of the troubleshooting optimization problem with different complexity levels. A comparison among the algorithms was made in terms of convergence speed, accuracy and precision. Algorithm complexity was also taken into account, in order to make the comparison useful for practical application. A thorough analysis was made for each algorithm, considering different perspectives of the data collected during the experiments. The final results showed that TLBO has advantages over the others.

**Keywords.** Metaheuristics. Troubleshooting Problems. TLBO.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1. INTRODUCTION

Metaheuristics are powerful in the way they efficiently provide good solutions for real-world problems [Simon, 2013]. They have been successfully used when the search space is too large, the problem is too complex, the evaluation function is too complicated, the solution is too constrained, or there is no hint of how to solve the problem [Michalewicz and Fogel, 2004]. A metaheuristic can be defined as a generalization of a heuristic that can be applied to a wide variety of problems. The growing interest in metaheuristic algorithms is due to their flexibility (which allows them to be used for a great variety of problems) and good performance (which allows them to efficiently solve big-size instances of problems).

Metaheuristics suffer from their generality: they lack restrictions on their parameters. Therefore, it is necessary to finely tune the parameters such that computational resources and algorithm accuracy are not affected. The Genetic Algorithm (GA), for example, has a high sensitivity to mutation probability, crossover probability and population size [Pinel et al., 2012; Srinivas et al., 2014].

The selection of good parameters is usually done using parameter tuning or parameter control [Eiben and Smit, 2012]. The former denotes the set of methods that define the parameters before the optimization is run, while the latter denotes the methods in which the parameters are dynamically defined during the execution of the optimization algorithm. Parameter control methods can still be categorized into deterministic, adaptive, or self-adaptive.

A set of good parameters that provides good results for a specific problem is not guaranteed to provide good results for other problems. Parameters are usually mutually sensitive and very dependent on the models and problems at hand. Those reasons strengthen the need for parameter control, even though it is harder to achieve [Wolpert and Macready, 1997]. The Teaching-Learning Based Optimization (TLBO) algorithm is a novel metaheuristic that has been recently proposed by Rao et al. [2011]. It has the population size as its unique parameter, besides the stop criteria. TLBO is based on the teaching-learning process observed in a classroom and simulates the influence of a

teacher on the output of a group of students in a class.

TLBO has achieved remarkable performances in different types of problems such as constrained [Rao and Patel, 2012], and unconstrained [Rao and Patel, 2013] optimization problems. It has also been successfully used in combinatorial optimization problems such as the flow shop and the job shop scheduling problems [Baykasoglu et al., 2014] and the set covering problem [Crawford et al., 2015].

In this paper, we compare the performance of TLBO with the performance of Simulated Annealing (SA) and Genetic Algorithm (GA) in terms of convergence speed, accuracy and precision. All the algorithms are used in their original form to provide a fair comparison.

Numerical experiments are conducted using the troubleshooting optimization problem. Troubleshooting optimization problem is a combinatorial optimization problem which is known to be NP-hard for most of real applications [Vomvel'a, 2003]. Three troubleshooting models with different complexity levels are used.

The remaining sections of this paper are organized as follows. Section 2 describes the troubleshooting problem. Section 3 presents the basic principles of the metaheuristic algorithms used. Section 4 presents the proposed metric used to compare the results. Section 5.1 presents the results observed during the numerical experiments. Concluding remarks are presented in section 6.

## 2. TROUBLESHOOTING OPTIMIZATION PROBLEMS

Troubleshooting is the name given for the sequence of actions performed in order to fix a given system. These actions are commonly separated into two categories: diagnostic and repair [Vomvel'a, 2003].

The first type of action, usually less expensive and time consuming, can be inconclusive, i.e., represent some cost for no advancement toward an actual solution. Repair actions, on the other hand, are a step in the direction of solving the problem. Several actions might be required to fix a single fault, as well as several faults might have occurred simultaneously.

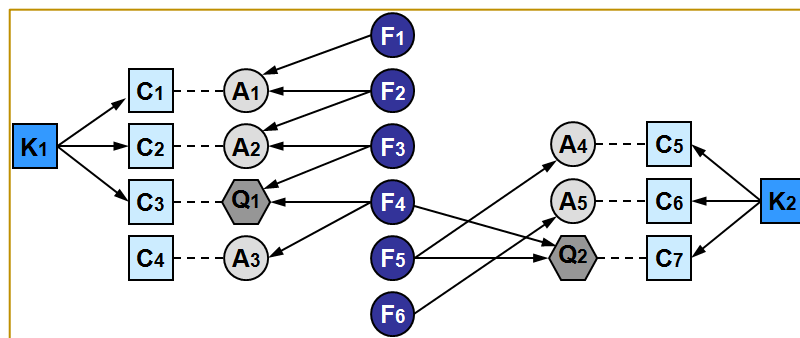
These types of problems are very practical and common in our daily lives. The decision theoretic troubleshooting comes as a way to

model and study our way of making decisions to optimize the troubleshooting process and thus minimize the costs involved.

A common way to represent this problem is through oriented graphs, such as the one presented in Figure 1. Nodes represent the possible failure modes ( $F_i$ ), diagnose questions ( $Q_i$ ) and repair actions ( $A_i$ ), while the oriented vertices represent the sequence of actions that can be executed [Vianna et al., 2016].

Costs are represented by  $C_i$ , while cost clusters are represented by  $K_i$ . A fault caused by failure mode  $F_i$  is only fixed when all repair actions connected to it are executed. A diagnostic question isolates a subset of possible failure modes. Each action and question has an associated cost, which is incurred if the action or question is executed. A cluster cost is incurred if at least one action or question connected to it is executed.

Figure 1: Troubleshooting Model 1 - Bayesian network representing a simple troubleshooting problem.



Several assumptions need to be made to ensure optimality of the sequence found [Langseth

and Jensen, 2001]:

- Single Fault: only one fault can be present at the system at a time.
- Perfect Repair: repair actions are always effective.
- Fixed Cost: costs do not vary with time.
- No questions: diagnose actions are not allowed.

- Independent Actions: each action addresses exactly one fault.

Without these restrictions there is no polynomial approximation algorithm for such problem [Lin, 2014].

As mentioned earlier, every troubleshooting strategy has an associated Expected Cost of Repair (ECR). This cost is calculated in terms of the action costs  $C_i$ , the cluster costs  $K_i$  and the failure mode probabilities  $P(F_i)$ . The ECR is defined as showed in Equation (1).

$$ECR(s) = \sum_{i=1}^n \left( P(F_i) \cdot \sum_{j \in I_{F_i}} C(A_j) + C(K_j) \right) \quad (1)$$

where  $P(Fi)$  is the probability of occurrence of  $F_i$ ,  $n$  is the number of possible failure modes,  $C(A_j)$  with  $j \in lFi$  is the cost of each action (repair or diagnose) and its cluster in the path  $lFi$ ,  $lFi$  is the path to repair  $F_i$  according to the strategy  $s$ . This will be used as the objective function of the metaheuristic optimization algorithm. Clusters are activated only once for every set of actions that use them.

### 3. METAHEURISTICS

#### 3.1. TEACHING-LEARNING-BASED OPTIMIZATION

The TLBO algorithm is divided into two main parts: the Teacher Phase and the Student Phase, which is also known as the Learner Phase [Rao and Patel, 2013]. During the Teacher Phase, students learn from the teacher, while in the Learner Phase students learn through the interaction among themselves.

There is a solution  $X$  associated with each student, which corresponds to a possible solution to the optimization problem under

consideration. Also, there is a result  $f(X)$  associated with each solution (or student), which can be obtained by evaluating the solution  $X$  using the objective function  $f$ . In the troubleshooting problem considered in this paper, a solution  $X$  corresponds to a troubleshooting strategy  $S$  and the associated result  $f(X)$  corresponds to its Expected Cost of Repair, denoted by  $ECR(S)$  [Vianna et al., 2016].

#### 3.1.1. TEACHER PHASE

In this phase, the algorithm simulates the learning of the students from the teacher (best solution). During this phase, the teacher makes an effort to increase the mean result of the class. Consider a group of  $n$  students. Let  $M_i$  be the mean solution of the students and  $T_i$  be the teacher at iteration  $i$ . The teacher  $T_i$  will make an effort to move  $M_i$  to its own level. Knowledge is gained based on the quality of the teacher and the quality of students in the class. The difference  $D_i$  between the solution of the teacher,  $XT_i$ , and the mean solution of the students,  $M_i$ , can be expressed according to Equation (2):

$$D_i = r_i(X_{T_i} - T_F \cdot M_i) \quad (2)$$

where  $r_i$  is a random number in the range [0; 1] for iteration  $i$  and  $T_F$  is a teaching factor for iteration

$i$ , which is randomly set to either 1 or 2 according to Equation (3):

$$T_F = \text{round}(1 + \text{rand}(0,1)) \quad (3)$$

Based on the difference  $D_i$ , the existing solution of student  $k$  in iteration  $i$ ,  $X_{ki}$ , with

$k \in \{1, 2, \dots, n\}$ , is updated in the teacher phase according to Equation (4):

$$X_{ki}^* = X_{ki} + D_i \quad (4)$$

where  $X_{ki}^*$  is the updated value of  $X_{ki}$ .

If  $f(X_{ki}^*)$  is better than  $f(X_{ki})$ ,  $X_{ki}^*$  is accepted and replaces  $X_{ki}$ . Otherwise,  $X_{ki}^*$  is discarded.

#### 3.1.2. STUDENT PHASE

In this phase, the algorithm simulates the learning of the students through interaction with one another. During this phase, students gain knowledge by discussing with another students who have more knowledge [Rao and Patel, 2013].

Consider a pair of students  $y$  and  $z$ . Let  $X_{yi}$  and  $X_{zi}$  be the solutions of students  $y$  and  $z$  at iteration  $i$ , respectively. If  $f(X_{yi})$  is better than  $f(X_{zi})$ , the solution of student  $z$  is updated according to Equation (5). Then,  $X_{zi}^*$  will replace  $X_{zi}$  if  $f(X_{zi}^*)$  is better than  $f(X_{zi})$ .

Similarly, if  $f(X_{zi})$  is better than  $f(X_{yi})$ , the solution of student  $y$  is updated according to Equation (6). Then,  $X_{yi}^*$  will replace  $X_{yi}$  if  $f(X_{yi}^*)$  is better than  $f(X_{yi})$ .

$$X_{zi}^* = r_i(X_{yi} - X_{zi}) \quad (5)$$

$$X_{yi}^* = r_i(X_{zi} - X_{yi}) \quad (6)$$

At the end of each iteration, the stop criteria must be checked. Different stop criteria may be adopted. Some of the most commonly adopted stop criteria are the maximum number of iterations, the maximum number of successive iterations without any improvement and the maximum simulation time. In this paper, the maximum simulation time is adopted as the stop criterion.

The SA algorithm is mainly divided into three blocks: temperature management, neighborhood generation and acceptance criteria. These blocks are briefly described in the following lines.

### 3.2. SIMULATED ANNEALING

The Simulated Annealing (SA) algorithm was developed by Khachaturyan et al. [1979] and uses the annealing process present in metallurgy to improve the traditional greedy search. Besides simple, SA is very fast and can be applied to both continuous and combinatorial applications. It is commonly used when finding a good solution fast is essential, but finding the optimal solution is not.

#### 3.2.1. TEMPERATURE MANAGEMENT

The SA algorithm has three important temperatures that must be defined: the initial temperature, denoted by  $T_0$ , the current temperature, denoted by  $T_i$  (where  $i$  is the current iteration), and the final temperature, denoted by  $T_F$  (where  $N$  is the maximum number of iterations).

The behavior of  $T_i$  can follow any profile desired, but it is typically chosen to be a monotonic descending function. This paper uses a linear function to describe the temperature profile, as shown in Equation (7).

$$T_i = (T_0 - t_{i-1}) \frac{T_0 - T_F}{t_F} \quad (7)$$

where  $t_{i-1}$  is the time since the beginning of the simulation until the start of the current iteration, and  $t_F$  is the expected final simulation time.

In the troubleshooting optimization problem considered in this paper, a candidate solution is a permutation of the available repair actions  $A_i$  and diagnostic questions  $Q_i$ . Thus, new individuals are generated by swapping the position of two elements in the original solution.

The stop criteria is also attributed to this block because when  $t_i > t_F$  the simulation ends.

#### 3.2.2. NEIGHBORHOOD GENERATION

The neighborhood generation block can assume different forms. The only requirement is the capability of providing a solution vector slightly different from the previous one.

#### 3.2.3. ACCEPTANCE CRITERIA

The acceptance criteria in SA governs its ability to, from time to time, accept a solution that is worse than the previous one. In this paper, the acceptance criteria is based on the commonly adopted Metropolis' Algorithm [Metropolis et al., 1953]. For minimization

problems, the probability of a solution to be accepted is defined according to Equation

(8).

$$P_A = \begin{cases} 1 & \text{if } C_i < C_{i-1} \\ e^{\frac{(C_i)-(C_{i-1})}{T_i}} & \text{if } C_i > C_{i-1} \end{cases} \quad (8)$$

where  $C_i$  is the cost of the current solution,  $C_{i-1}$  is the cost for the previous solution,  $T_i$  is the temperature for the current iteration.

paper, the initial population was built randomly.

### 3.3. GENETIC ALGORITHMS

Genetic Algorithms (GA) are one of oldest class of metaheuristics and one of the first Evolutionary Algorithms that became available. GA were introduced by Holland [1975]. Genetic Algorithms use concepts observed in human genetics to combine, mutate and select chromosomes until a stop criteria is met. They have been used successfully in both continuous and combinatorial optimization, and their success is mainly attributed to their simplicity and flexibility.

#### 3.3.2. EVALUATION

The evaluation stage is highly dependent of the problem at hand since it attributes a score to each of the individuals in the population. It is also usually delegated to this stage the control of the stop criteria. The objective function used in this paper is the Expected Cost of Repair (ECR) defined in Equation (1).

The first step required by the algorithm is the definition of a way to encode the solution vector into a chromosome, which is composed by genes. A collection of chromosomes forms the population in the algorithm. In this paper, the chromosome is a sequence of repair actions and diagnose questions identified by non-repeating integer numbers.

#### 3.3.3. SELECTION

The selection stage mimics the natural selection (or survival of the fittest) process, in which only the strongest individuals survive. This process receives the score from the evaluation stage in order to properly select the individuals.

The basic algorithm proposed by Holland takes the population of chromosomes and passes it through a series of five stages: initialization, evaluation, selection, crossover and mutation. A brief description on each stage is presented in the following lines.

There are several methods available to implement this stage. They can be classified as deterministic, stochastic or hybrid. Deterministic methods are more well-behaved and easier to debug. The stochastic methods mimic the natural selection process more ideally, but tend to be less well-behaved. The hybrid approaches are usually preferred because they tend to present the benefits of the previous methods without their disadvantages.

#### 3.3.1. INITIALIZATION

The initialization stage defines a way to generate a valid initial population. In this

In this paper, we used a hybrid approach in which a few elite individuals survived between generations while others were picked randomly. Non-elite individuals were selected based on a rank probability, in which they are ranked according to their fitness. The probabilities assigned depend on a given Pr. Each individual receives a probability  $P_i$  according to Equation (9).

$$P_i = P_r \cdot (1 - P_r)^i \quad (9)$$

### 3.3.4. CROSSOVER

The crossover stage mimics the reproduction process, in which the genes from two parents are mixed to form new individuals. This process usually takes a probability  $PC$ , also known as crossover rate, to determine if the parents will be crossed or kept.

Several implementations are available to implement this stage. In this paper, we use the Ordered Crossover (OX) method. The OX method starts by randomly defining two cut points. Then, the parents are broken into three sections. A queue is then built by selecting, in this order, the second section of the first parent, the second section of the second parent, the third section of the second parent and the first section of the second parent. The queue is then scanned for duplicated and they are removed leaving only the first occurrence.

### 3.3.5. MUTATION

The mutation stage mimics the mutation process that occurs with the genes in our cells. It allows us to evolve from one generation to another and is considered to be one of the main reasons for the diversity of the species. The individuals in the population are

mutated according to a mutation probability  $P_M$ .

## 4. METRIC

This section presents the metric proposed to compare the performance of the different optimizations

for solving the troubleshooting optimization problem. In each simulation, the following

outputs are recorded:

- the Estimated Cost of Repair (ECR), which indicates the quality of a given solution;
- the iteration time, which indicates the exact time it took for every iteration to run;
- the iteration count, which indicates how many iterations occurred before the current one; and
- the solutions vector for every iteration.

A simple metric, named Score and denoted by  $S$ , is proposed to evaluate the performance of the algorithms. The computation of  $S$  for continuous and discrete problems are presented in Equations (10) and (11), respectively.

$$S = \int_0^{t_{max}} ECR dt \quad (10)$$

$$S \approx - \sum_1^N \left( \min(ECR_i, ECR_{i-1}) + \frac{|ECR_i - ECR_{i-1}|}{2} \right) \cdot (t_i - t_{i-1}) \quad (11)$$

where  $ECR_i$  is the Expected Cost of Repair in iteration  $i$ ,  $t_i$  is the simulation time at iteration  $i$ .

## 5. NUMERICAL EXPERIMENTS

A limitation faced when performing such simulations was the restricted computational time available, a restriction is often observed in commercial applications. One consequence of this restriction is that it is not viable to sample the parameter space extensively, and so a subset of conditions will have to be studied and the conclusions will have to be extrapolated. Design of Experiments (DOE), which is a procedural way of analyzing some

experiment, was used to define few experiments that would help map the parameter space and help draw conclusive results.

A simple algorithm derived from the steepest descend was used to guide the DOE through the parameter space exploration. Each algorithm had a initial parameter setting arbitrarily chosen according to current best practices. Full-factorial form, which uses  $2^k$  points where  $k$  is the amount of parameters defined a Region of Interest (ROI) to be analyzed, was used since aliasing (using  $2^{k-n}$ ) would reduce the statistical significance of the data. As an optimization algorithm itself, the stop criterion used for the DOE was based



on signal-to-noise ratio concept, where the algorithm would stop if  $\sigma/\mu > 1$  for S or if 20 DOE iteration were executed.

A Monte Carlo approach was also used. The simulations conducted in this paper were repeated 5 times for each parameter/algorithm

setting. This value was empirically set in order for the whole simulation to be completed in approximately one full day, as shown in Table 1.

Table 1: Complete simulation time estimate

|                                 | TLBO | SA          | GA   |
|---------------------------------|------|-------------|------|
| Number of Models                |      | 3           |      |
| Maximum DOE Iterations          |      | 20          |      |
| # Input Parameters              | 1    | 2           | 5    |
| DOE Experiments per Iteration   | 2    | 4           | 32   |
| Fixed Simulation Time (Seconds) |      | 10          |      |
| Estimated Simulation Overhead   |      | 10 %        |      |
| Estimated Time (Hours)          | 0.4  | 0.7         | 5.9  |
| Total Estimated Time (Hours)    |      | <b>7.0</b>  |      |
| Monte Carlo Repetitions         |      | 5           |      |
| Estimated Time (Hours)          | 1.8  | 3.7         | 29.3 |
| Total Estimated Time (Hours)    |      | <b>34.8</b> |      |

The algorithms were implemented in Matlab® R2016a and ran on a HP ENVY dv6-7200 CTO with Intel Core i7-3630QM CPU @ 2.40 GHz, 12 GB of RAM, Nvidia 920M GPU with 2 GB of dedicated memory and Windows 10 Pro OS.

Speed was calculated directly from the ECR by comparing the time taken for the algorithm the 105% of the best solutions found. The best solutions found had ECRs equal to 78.5, 199.0 and 206.3 for Models 1, 2 and 3 respectively.

## 5.1. SIMULATION RESULTS

The results were measured in terms of speed, accuracy and precision, each calculated over a dimension of the data collected here. Accuracy was measured by how far the final values were from the best solution found between all the algorithms. Precision was measured by the standard deviation of the value in the best experiment performed.

### 5.1.1. SIMULATED ANNEALING

A initial ROI was set, as hard limits to encapsulate the ROIs, is shown in Table 2. These values where chosen after a few experiments that indicated they would perform fairly for the problem at hand. The hard limits, values for which the ROIs would be truncated, were defined by the simulated annealing intrinsic features.

Table 2: Initial Region of Interest and Hard Limits for the SA algorithm.

| Parameters | Initial Region |             | Hard Limits   |               |
|------------|----------------|-------------|---------------|---------------|
|            | Lower Bound    | Upper Bound | Lower Bound   | Upper Bound   |
| $T_0$      | 10             | 50          | Unconstrained | Unconstrained |
| $\Delta T$ | 0              | 10          | 0             | Unconstrained |

Figure 2 shows an example of the simulations performed for Model 1. The results are summarized in Table 3. These simulation were unique in that they did not reach the stop criterion for two of the three models, which happened because the standard deviation would decrease as much as the mean slope.

The SA reached the best mean score for Model 3, even though its precision dropped considerably. Although its score was the best for Model 3, its best solution did not reach the 5% value above the best solution (216.6 for Model 3) to be considered fast.

Figure 2: Example of slopes of each ROI as the DOE searched for a better parameter set for Model 1. Black thick lines represent the

mean slope, while the red thin line represents the slope for each Monte Carlo repetition of the ROI gradient.

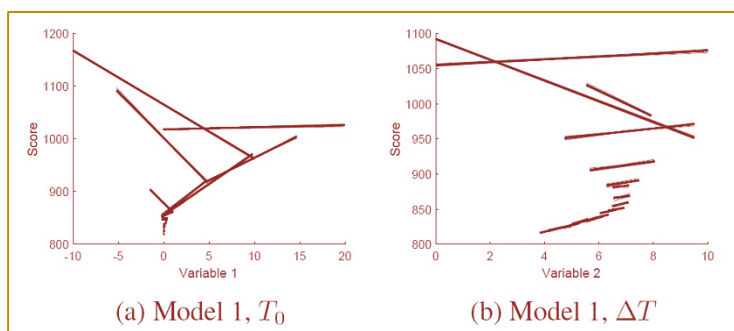


Table 3: Final simulation results for SA algorithm.

| Parameter                             | Model  |         |           |
|---------------------------------------|--------|---------|-----------|
|                                       | 1      | 2       | 3         |
| $T_0$                                 | 0.00   | -0.02   | 0.10      |
| $\Delta T$                            | 3.54   | 4.67    | 4.80      |
| Best Mean Score                       | 812.67 | 2080.72 | 2245.93   |
| Max DOE Iteration                     | Max    | Max     | 12        |
| Accuracy (% Worse than Best Solution) | 3.5%   | 4.6%    | 8.9%      |
| Precision (Standard Deviation as %)   | 0.1%   | 0.2%    | 1.0%      |
| Speed (5% to Best, in seconds)        | 0.06   | 1.23    | $\geq 10$ |

### 5.1.2. GENETIC ALGORITHM

The initial ROI and its hard limits is shown in Table 2. These values were chosen after a few experiments that indicated they would perform fairly for the problem at hand. The

hard limits, values for which the ROIs would be truncated, were defined by the Genetic Algorithm intrinsic features. The lower population bound, for example, was chosen because with less than 2 individuals there would not be enough parents to crossover.

Table 4: Initial Region of Interest and Hard Limits for the GA algorithm.

| Parameters                      | Initial Region |             | Hard Limits |               |
|---------------------------------|----------------|-------------|-------------|---------------|
|                                 | Lower Bound    | Upper Bound | Lower Bound | Upper Bound   |
| Population Size                 | 10             | 50          | 2           | Unconstrained |
| Mutation Rate                   | 10%            | 20%         | 0%          | 100%          |
| Crossover Rate                  | 10%            | 20%         | 0%          | 100%          |
| Rank Probability                | 10%            | 20%         | 0%          | 100%          |
| Proportion of Elite Individuals | 50%            | 60%         | 0%          | 100%          |

The simulation summary is presented in Table 5. GA presented the slowest, less precise and less accurate performance among the three algorithms. Models greater than Model 1 could not have its minimum value found in the

appointed time. Its accuracy deviation increased exponentially as the models got more complex. Its speed was also very poor, making it not able to reach the best solution in the appointed time.

Table 5: Final simulation results for GA algorithm.

| Parameter                             | Model  |         |         |
|---------------------------------------|--------|---------|---------|
|                                       | 1      | 2       | 3       |
| Population Size                       | 217    | 179     | 134     |
| Mutation Rate                         | 41.35% | 43.10%  | 41.73%  |
| Crossover Rate                        | 19.92% | 21.15%  | 16.57%  |
| Rank Probability                      | 2.13%  | 0.36%   | 2.16%   |
| Proportion of Elite Individuals       | 75.07% | 74.97%  | 66.03%  |
| Best Mean Score                       | 808.4  | 2268.57 | 2895.92 |
| Max DOE Iteration                     | 14     | 13      | 10      |
| Accuracy (% Worse than Best Solution) | 3.0%   | 14.0%   | 40.4%   |
| Precision (Standard Deviation as %)   | 0.9%   | 1.5%    | 7.3%    |
| Speed (5% to Best, seconds)           | 1.6987 | ≥ 10    | ≥ 10    |

### 5.1.3. TLBO

The initial ROI and Hard Limit chosen for TLBO are those shown in Table 6. The lower

bound was chosen because 2 individuals because there needs to be a learner in the algorithm besides the teacher.

Table 6: Initial Region of Interest for the TLBO algorithm

| Parameters      | Initial Region |             | Hard Limits |               |
|-----------------|----------------|-------------|-------------|---------------|
|                 | Lower Bound    | Upper Bound | Lower Bound | Upper Bound   |
| Population Size | 2              | 10          | 2           | Unconstrained |

The Table 7 summarizes the results for the TLBO simulations. Its accuracy, precision and speed equals the SA algorithm for the smaller

models, but its accuracy and precision gets slightly worse than SA for Model 3.

Table 7: Final simulation results for TLBO algorithm.

| Parameter                             | Model  |         |         |
|---------------------------------------|--------|---------|---------|
|                                       | 1      | 2       | 3       |
| Population Size                       | 11     | 8       | 2       |
| Best Mean Score                       | 785.22 | 2008.37 | 2297.13 |
| Max DOE Iteration                     | 1      | 1       | 1       |
| Accuracy (% Worse than Best Solution) | 0.0%   | 0.9%    | 11.3%   |
| Precision (Standard Deviation as %)   | 0.0%   | 1.0%    | 2.2%    |
| Speed (5% to Best, seconds)           | 0.03   | 2.46    | 9.25    |

## 6. CONCLUSION

This work was set out with the purpose of comparing metaheuristic optimization algorithms to the newly proposed Teaching-Learning Based Optimization. To do so, the Troubleshooting problem was coded and simulated within a common-ground framework. This comparison was made pragmatically and focused on convergence

speed, accuracy and precision. The algorithms complexity was also taken into account, in order to make the comparison useful for practical application.

All three algorithms are considered to be conceptually simple, but they allow a lot of design choices that can complicate optimization task. GA was the method with the most choices to be made which, summed to

its highly sensitive parameter tuning, made it barely practical. SA and TLBO presented the opposite characteristic: they seemed very practical to set. Even though SA has one parameter more than TLBO, it was not hard to set it to a good value. One preoccupation towards SA is that its Metropolis' Criterion can make the algorithm perform very poorly if not properly contained, tending to a Random Search.

TLBO was more accurate for Model 1 but quickly got worse for Model 3. SA's was more accurate for Model 3. The accuracy of GA got a lot worse as the model became more complex.

Although TLBO and SA present the same level of precision for small models, SA performed better for Model 3. One point to note is that TLBO took very few DOE iteration to tune.

## REFERENCES

- [1] Baykasoglu, A., Hamzadayi, A., and KOSE, S. Y. (2014). Testing the performance of teaching learning based optimization (TLBO) algorithm on combinatorial problems: Flow shop and job shop scheduling cases. *Information Sciences*, 276:204–218.
- [2] Crawford, B., Soto, R., Leiva, F. A., Johnson, F., and Paredes, F. (2015). Problema del conjunto de cobertura resuelto mediante el algoritmo binario de optimización basado en enseñanzaaprendizaje.
- [3] In Conferência Ibérica de Sistemas e Tecnologias de Informação, p. 106–109, Águeda. IEEE.
- [4] Eiben, A. E. and Smit, S. K. (2012). Evolutionary algorithm parameters and methods to tune them. In Hamadi, Y., Monfroy, E., and Saubion, F., editors, *Autonomous Search*, p. 15–36. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [5] Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 2 edition.
- [6] Khachatryan, A., Semenovskaya, S., and Vainshtein, B. (1979). Statistical-thermodynamic approach to determination of structure amplitude phases. p. 519–524.
- [7] Langseth, H. and Jensen, F. V. (2001). Heuristics for two extensions of basic troubleshooting. *Frontiers in Artificial Intelligence and Applications*, 66:80–89.
- [8] Lin, V. (2014). Decision-theoretic troubleshooting: Hardness of approximation. *International Journal of Approximate Reasoning*, 55(4):977–988. Special issue on the sixth European Workshop on Probabilistic Graphical Models.
- [9] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087.
- [10] Michalewicz, Z. and Fogel, D. B. (2004). *How to Solve It: Modern Heuristics*. Springer, enlarged 2nd edition.
- [11] Pinel, F., Danoy, G., and Bouvry, P. (2012). Evolutionary algorithm parameter tuning with sensitivity analysis. In *Proceedings of the 2011 International Conference on Security and Intelligent Information Systems (SIIS'11)*, p. 204–216, Berlin, Heidelberg. Springer-Verlag.
- [12] Rao, R. V. and Patel, V. (2012). An elitist teaching-learning based optimization algorithm for solving complex constrained optimization problems. *International Journal of Industrial Engineering Computations*, 3:535–560.
- [13] Rao, R. V. and Patel, V. (2013). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. *Scientia Iranica*, 20:710–720.
- [14] Rao, R. V., Vakharia, D. P., and Savsani, V. J. (2011). Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43: 303–315.
- [15] Simon, D. (2013). *Evolutionary Optimization Algorithms*. Wiley. URL <https://books.google.com.br/books?id=gwUwIEPqk30C>.
- [16] Srinivas, C., Reddy, B. R., Ramji, K., and Naveen, R. (2014). Sensitivity analysis to determine the parameters of genetic algorithm for machine layout. *Procedia Materials Science*, 6:866–876. 3<sup>rd</sup> International Conference on Materials Processing and Characterisation (ICMPC 2014).

[17] Vianna, W. O. L., Rodrigues, L. R., Yoneyama, T., and Mattos, D. I. (2016). Troubleshooting optimization using multi-start simulated annealing. In 10th Annual IEEE Systems Conference, Orlando, FL, USA.

[18] Vomvel'a, M. (2003). Complexity of decision-theoretic troubleshooting. *International Journal of Intelligent Systems*, 18(2):267–277.

[19] Wolpert, D. H. and Macready, W. G. (1997). No free lunch theorems for optimization. *Trans. Evol. Comp*, 1(1):67–82.

# Capítulo 16

## OTIMIZAÇÃO DO ROTEIRO DE DISTRIBUIÇÃO DE MATERIAIS NO SISTEMA ÚNICO DE SAÚDE EM JOINVILLE/SC

*Flavia Haweroth*

*Vanina Macowski Durski Silva*

*Ana Paula Nunes Duarte*

**Resumo:** O conceito de Gerenciamento da Cadeia de Suprimentos, que já é amplamente conhecido e aplicado por empresas do setor privado, também pode ser adaptado ao setor público, no entanto a sistemática de contratação e os entraves do sistema licitatório são barreiras para seu efetivo uso. As atividades de transporte e administração de estoques representam um alto custo para qualquer empresa, inclusive para um almoxarifado público, e as ferramentas de roteirização de veículos são fundamentais na redução destes custos. Este artigo se propõe a aplicar uma heurística de resolução do Problema de Roteirização de Veículos em um almoxarifado do sistema público de saúde na cidade de Joinville, no Brasil, propondo assim uma rota alternativa de distribuição física. As rotas atuais e propostas pela heurística são comparadas a fim de encontrar melhora. O resultado foi satisfatório, e mostrou uma redução de 18,3% da distância percorrida no mês.

\* Artigo originalmente apresentado no Encontro Nacional de Engenharia de Produção (ENEGEP, 2017)

## 1 INTRODUÇÃO

O Gerenciamento da Cadeia de Suprimentos (GCS) para Ballou (2006) trata do conjunto de atividades logísticas, que se repetem inúmeras vezes ao longo do canal pelo qual matérias-primas vão sendo convertidas em produtos acabados. Esse conceito também é aplicável na Administração Pública, onde o cliente final seria a população através do uso dos serviços públicos fornecidos, atendendo os princípios de economicidade e qualidade exigidas desses serviços (TRIDAPALLI, 2008).

Dentre as atividades logísticas que compõem o GCS, a administração de estoque e o transporte são as que se destacam como mais custosas a qualquer empresa, pública ou privada, por isso para a redução de custos deve-se estudar os procedimentos dessas duas atividades, que numa visão mais ampla tem reflexos em todos os outros departamentos (ASSINI *et al.*, 2012). Segundo Tridapalli (2008, p. 42):

“Trabalhar com o conceito de ‘chão de fábrica’ no setor público, lidando com redução de custos e aumento de eficiência, é a bandeira de vários segmentos sociais para combater a falência do Estado. O poder público brasileiro precisa de qualidade na gestão e a tecnologia de racionalização de custos usada pelo setor privado, para aumentar a eficiência.”

Entre as decisões a serem tomadas no transporte de mercadorias destaca-se a escolha do modal e a roteirização de veículos ao longo da rede de transporte a fim de minimizar os tempos ou distâncias (BALLOU, 2006). Diante desse contexto o objetivo deste trabalho é propor uma alternativa para a rota de distribuição de materiais ambulatoriais no setor da saúde do município de Joinville. Para isso será aplicado um método heurístico de resolução do Problema de Roteirização de Veículos (PRV) implementado em linguagem de programação. A rota obtida será posteriormente comparada com a atual levando em considerando as limitações da modelagem e dos dados.

O trabalho está estruturado como segue: o presente capítulo, que contém a parte introdutória, em seguida no capítulo 2 apresenta-se o referencial teórico dos temas pertinentes ao trabalho, o capítulo 3 contempla a definição do problema a ser estudado. O quarto capítulo apresenta a metodologia aplicada, o quinto capítulo apresenta a análise dos resultados seguido do capítulo final de conclusão.

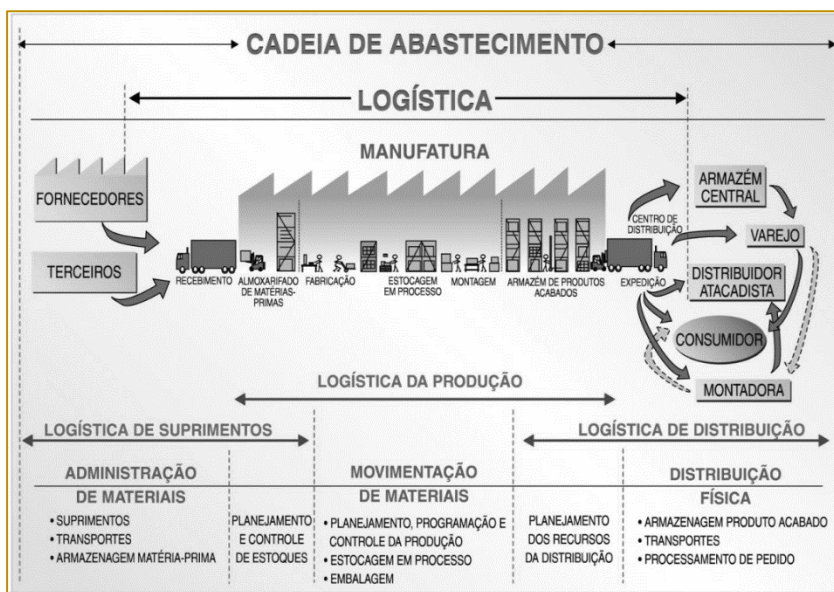
## 2. REFERENCIAL TEÓRICO

### 2.1. DISTRIBUIÇÃO FÍSICA DE PRODUTOS

A logística é definida pelo Conselho de Gerenciamento da Cadeia de Suprimentos citado por Tadeu (2008, p. 17) como “processo de planejamento, implementação e controle eficiente e eficaz do fluxo e armazenagem de mercadorias, serviços e informações relacionadas desde o ponto de origem até o ponto de consumo, com objetivo de atender às necessidades do cliente”. A logística de distribuição, conforme pode ser visto na Figura 1, compõe o elo final da Cadeia de Suprimentos fazendo a interface entre empresa e consumidor. Esse elo envolve transferência de produtos entre fábrica e os armazéns próprios ou de terceiros seus estoques, os subsistemas de entrega urbana e interurbana (ALVARENGA e NOVAES, 2000).

De acordo com Coêlho (2010), na área da saúde pública os processos logísticos, devido às suas características muito específicas, precisam ser encarados com uma abordagem orientada não só para a racionalização de custos, mas também como elemento fundamental de apoio à prestação de cuidados de saúde aos pacientes. Para tanto, deve-se haver um eficiente esquema de planejamento das atividades de compras, armazenagem e gerenciamento de materiais em estoque, bem como na distribuição desses materiais destinados ao uso em atividades hospitalares, uma vez que a falta de um mecanismo eficiente de controle impede a apuração das reais necessidades de abastecimento.

Figura 1 - Modelo de Cadeia de Suprimentos



Fonte: ASSINI *et al.* (2008)

Para tanto, deve-se haver um eficiente esquema de planejamento das atividades de compras, armazenagem e gerenciamento de materiais em estoque, bem como na distribuição desses materiais destinados ao uso em atividades hospitalares, uma vez que a falta de um mecanismo eficiente de controle impede a apuração das reais necessidades de abastecimento.

Segundo Novaes (2004) o sistema de distribuição pode ser dividido em “um para um” e “um para muitos”. Na distribuição “um para um”, o veículo é carregado totalmente, com o intuito de ter o melhor aproveitamento possível de sua capacidade. O carregamento do veículo é feito no depósito ou num centro de distribuição e transportada para um outro ponto de destino, podendo ser outro centro de distribuição ou qualquer outro tipo de instalação. Na distribuição “um para muitos”, o veículo é carregado no centro de distribuição com mercadorias, com destino a diversas lojas ou diversos clientes, executando um roteiro de entregas predeterminado. Este é o tipo de distribuição física que caracteriza o ambiente em que foi realizado o estudo de caso.

## 2.2. PROBLEMA DE ROTEIRIZAÇÃO DE VEÍCULOS (PRV)

O PRV é um problema de roteirização em nós onde o objetivo é a definição de rotas de

entrega (sequência de nós) a serem percorridas por uma frota de veículos com um custo mínimo, originando e terminando em uma garagem ou depósito central. A frota deve servir um conjunto de clientes, cada qual caracterizado por sua localização no espaço bidimensional, demanda e tempo de serviço. Cada cliente tem sua demanda suprida por exatamente um veículo e a demanda total dos clientes não deve exceder a capacidade total do veículo. O tempo total de qualquer rota não pode exceder um limite previamente estabelecido, incluindo os tempos de viagem entre clientes e tempos de serviço em cada cliente (BARBOSA, 2005).

Uma formulação de programação matemática para o PRV foi dada por Fisher e Jaikumar (1981), apresentada pelas Equações (1) à (8).

As restrições (2) asseguram que um veículo não visite mais de uma vez um cliente. As restrições (3) garantem que o depósito recebe uma visita de todos os veículos. As restrições (4) obrigam que a capacidade dos veículos não seja ultrapassada. As restrições (5) garantem que os veículos não interrompam suas rotas em um cliente e as restrições (6) são as tradicionais restrições de eliminação de subrotas.



$$\text{Min } Z = \sum (c_{ij} \sum x_{ijk}) \quad (1)$$

Sujeito a:

$$\sum y_{ik} = 1 \quad i = 2, \dots, n \quad (2)$$

$$\sum x_{ijk} = m \quad i = 1 \quad (3)$$

$$q_i y_{ik} \leq Q_k \quad k = 1, \dots, m \quad (4)$$

$$\sum x_{ijk} = \sum x_{jik} = y_{ik} \quad i = 1, \dots, n \quad k = 1, \dots, m \quad (5)$$

$$\sum x_{ijk} \leq |S| - 1 \quad \forall S \subseteq \{2, \dots, n\}, k = 1, \dots, m \quad (6)$$

$$y_{ik} \in \{0,1\} \quad i = 1, \dots, n \quad k = 1, \dots, m \quad (7)$$

$$x_{ijk} \in \{0,1\} \quad i, j = 1, \dots, n \quad k = 1, \dots, m \quad (8)$$

Em que:

$x_{ijk}$  é uma variável binária que assume o valor 1 quando o veículo  $k$  visita o cliente  $j$  imediatamente após o cliente  $i$ , 0 em caso contrário;

$y_{ik}$  é uma variável binária que assume valor 1 se o cliente  $i$  é visitado pelo veículo  $k$ , 0 em caso contrário;

$q_i$  é a demanda do cliente  $i$ ;

$Q_k$  é a capacidade do veículo  $k$ ;

$c_{ij}$  é o custo de percorrer o trecho que vai do cliente  $i$  ao cliente  $j$ .

### 2.3. MÉTODO DE CLARKE E WRIGHT

O PRV pertence à categoria de problemas NP - difíceis, cuja dificuldade de encontrar a solução aumenta rapidamente à medida que o número de nós cresce (OLIVEIRA, 2007). Laporte (1992) cita e exemplifica uma série de algoritmos exatos, mas reconhece a limitação que estes apresentam ao lidar com problemas maiores e mais complexos, indicando as meta-heurísticas como um campo de estudos mais promissor. Apesar dos métodos heurísticos não garantirem que uma solução ótima seja encontrada, os benefícios, tanto em termos computacionais como por apresentar boa representação da realidade, obtendo soluções satisfatórias, são suficientes para adotar a abordagem heurística em problemas de otimização.

O algoritmo de Clarke e Wright ou método das economias foi proposto pela primeira vez em 1964 por Clarke e Wright para resolver o PRV em que o número de veículos é livre. O método inicia com  $n$  roteiros contendo o depósito e um cliente, onde  $n$  é o número de

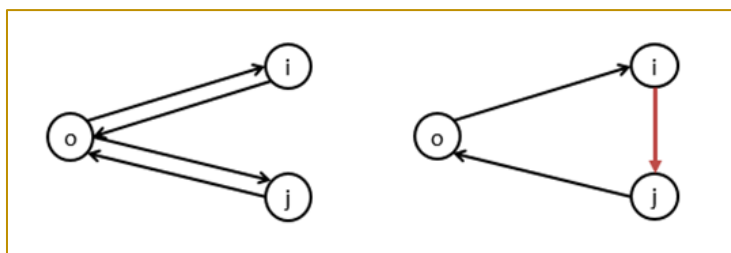
clientes a serem atendidos, e a cada etapa um par de rotas é mesclado de acordo com a maior economia que pode ser gerada (LAPORTE, 1992).

A economia  $s_{ij}$  gerada na união de dois roteiros é calculada através da Equação (9), onde  $d_{oi}$ ,  $d_{oj}$  e  $d_{ij}$  são as distâncias entre os pontos. A partir disso, lista-se em ordem decrescente as economias de todos os pares de vértices e a cada iteração avalia-se o topo da lista para fazer a união das rotas. Um par de rotas não será mesclado se as restrições de capacidade do veículo e tempo total de percurso forem ultrapassadas. A Figura 2 exemplifica uma união de duas rotas.

$$s_{ij} = d_{oi} + d_{oj} - d_{ij}$$

Segundo Ballou (2006), este método é relativamente rápido, em termos computacionais, para problemas com um número moderado de paradas, e capaz de gerar soluções que são quase ótimas. Através da avaliação dos métodos existentes para resolução do PRV, este se mostrou mais apropriado para este artigo.

Figura 2 - Combinação de rotas na heurística de Clarke e Wright



Fonte: Autor

### 3. METODOLOGIA

Neste artigo, o local em estudo é o almoxarifado da Secretaria da Saúde do município de Joinville, chamado de Serviço de Administração de Materiais Ambulatoriais (SAMA) e todas as informações necessárias foram obtidas por entrevista e visita ao local. A SAMA recebe dos fornecedores e armazena até a distribuição, que é terceirizada através de um contrato de quilometragem fixa de 3000 km por mês. O contrato inclui um veículo (modelo Ducato Cargo da Fiat) e um motorista que trabalha 8 horas diárias. Nesse tipo de contrato o valor é pago integralmente mesmo que a quilometragem real não atinja o limite e esses 3000 km são usados nas atividades de distribuição e também em outras que serão desconsideradas por este artigo. A distribuição segue um cronograma mensal organizado em 4 semanas que serão tratadas separadamente na resolução do problema. A

Tabela 1 mostra o cronograma das duas primeiras semanas do mês de agosto de 2015.

Com essas questões levantadas, surgiu o interesse em analisar se há um melhor caminho a ser realizado pelo motorista a fim de minimizar a quilometragem gasta mensalmente e, assim possibilitar uma redução da quilometragem do contrato.

O método de resolução a ser aplicado é a heurística de Clarke e Wright implementada em linguagem C++ pelas autoras. Para aplicar o método foram necessárias a matriz de distâncias, as demandas dos clientes e os tempos de percurso e parada. A matriz de distâncias foi obtida através do *Software ArcGIS 10* usando a ferramenta *network analysis*. As distâncias são assimétricas e relativas as rotas mais curtas entre um ponto e outro.

Tabela 1 - Duas primeiras semanas do cronograma de entregas de agosto de 2015

| Cronograma |               |               |               |                 |
|------------|---------------|---------------|---------------|-----------------|
|            | Semana 1      |               | Semana 2      |                 |
|            | Manhã         | Tarde         | Manhã         | Tarde           |
| Segunda    | Pirabeiraba   | SAMU          | Jarivatuba    | Bmw I           |
|            | Rio Bonito    | Centrinho     | Floresta      | Bmw II          |
|            | Rio Da Prata  | Bakita        |               | Orestes Guim.   |
|            | Canela        | Resid. Terap. |               | Pq. Guarani     |
|            | Odonto III    |               |               | Presídio        |
| Terça      | PA Sul        | CAPS II       | PA Sul        | Itinga          |
|            | Siavo         | CAPS I        | Ad. Garcia    | Itinga Cont.    |
|            | CEO           | CAPS AD       |               | Profipo         |
|            | PAM Odonto    | CAPS III      |               | Km 4            |
|            | NAIPE         | CAF           |               |                 |
|            |               | Patrimônio    |               |                 |
|            |               | UAA           |               |                 |
| Quarta     | PA Leste      | Comasa        | PA Leste      | Jd. Edilene     |
|            | Iririu        | Jd. Iririu    | Gloria        | Paranaguamirim  |
|            |               |               | Vila Nova     | E. De Matos     |
|            |               |               | Bom Retiro    |                 |
| Quinta     | PA Norte      | Unid. Sanit.  | PA Norte      | Itaum           |
|            | PAM Bucarein  | CTA           | Costa E Silva | Dengue          |
|            | Laboratório   | Imunização    | SOIS          | GUVS            |
|            |               | Epidemiologia | Pq. Douat     |                 |
|            |               | Ambulatório   |               |                 |
| Sexta      | Rio Do Ferro  | Saguaçu       | Fatima        | Farmácia Escola |
|            | Pq. Joinville | PAM Boa Vista | Petropolis    | Zeladoria       |
|            | Aventureiro I |               |               |                 |

Fonte: Autor

A demanda de cada unidade de saúde é variável e os materiais entregues não são unitizados, sendo assim difícil definir

demanda e capacidade do veículo. Para tratar desse problema foi solicitada a ajuda a uma funcionária na unidade que é responsável

pela separação dos pedidos. A mesma classificou as unidades em quatro categorias qualitativas de tamanho, onde a maior seria a que ocuparia toda a capacidade do veículo. Essa classificação qualitativa foi quantificada através de pesos ponderados com base em roteiros do cronograma com vários pontos, conforme mostra a Tabela 2. O tempo de

percurso foi obtido assumindo a velocidade de 30 km/h e o máximo percurso foi de 4 horas que corresponde a jornada cumprida pelo motorista em cada período. Os tempos de parada nas unidades para descarregar foram considerados constantes com valor de 30 minutos, sugerido pelo motorista.

Tabela 2 - Classificação das unidades e seus pesos na ocupação do veículo

| Classificação | Peso |
|---------------|------|
| GG            | 1    |
| G             | 0,3  |
| M             | 0,15 |
| P             | 0,1  |

Fonte: Autor

Tendo sido definido todos os valores e limites necessários no programa, o mesmo foi executado em computador com processador Intel Core I5 de 3,10 GHz, 4 GB de memória RAM e sistema operacional Windows 7 Ultimate.

#### 4. RESULTADOS E DISCUSSÕES

Na Tabela 3 podemos ver que a redução das distâncias em cada uma das semanas foi em média de 18,0% e no mês inteiro de 18,3%.

Tabela 3 - Comparação das rotas atuais com as obtidas no programa

|                  | Distâncias Percorridas (km) |                           |          |
|------------------|-----------------------------|---------------------------|----------|
|                  | Rotas Atuais                | Rotas Obtidas No Programa | Melhoria |
| Semana 1         | 214                         | 192                       | 10,3%    |
| Semana 2         | 225                         | 165                       | 26,7%    |
| Semana 3         | 208                         | 182                       | 12,5%    |
| Semana 4         | 246                         | 191                       | 22,4%    |
| Total Mês        | 893                         | 730                       | 18,3%    |
| Média por Semana |                             |                           | 18,0%    |

Fonte: Autor

Em posse desses resultados foi também feito uma reorganização do cronograma utilizando as novas rotas encontradas, conforme pode ser visto na Tabela 4, que mostra essa reorganização para a primeira semana. Algumas alterações se deram apenas por mudanças nas ordens de visita de uma mesma rota. Em outros casos as unidades

foram completamente recombinaadas. Ocorreram também algumas rotas inadequadas do ponto de vista geográfico, unidades distantes umas das outras combinadas em uma única rota. Isso se deve às características da heurística usada, e podem ser resolvidas com alguma alteração manual.

Tabela 4 - Reorganização do cronograma da semana 1

| 1ª SEMANA |               |                 |               |                 |
|-----------|---------------|-----------------|---------------|-----------------|
|           | ATUAL         |                 | PROPOSTA      |                 |
|           | Manhã         | Tarde           | Manhã         | Tarde           |
| Segunda   | Pirabeiraba   | SAMU            | Rio Da Prata  | Saguaçu         |
|           | Rio Bonito    | Centrinho       | Odont III     | Pirabeiraba     |
|           | Rio Da Prata  | Bakita          | Canela        | SAMU            |
|           | Canela        | Resid. Terap.   | Rio Bonito    | Laboratório     |
|           | Odont III     |                 |               |                 |
| Terça     | PA Sul        | CAPS II         | PAM Bucarein  | NAIPE           |
|           | SIAVO         | CAPS I          | SIAVO         | CAPS I          |
|           | CEO           | CAPS AD         | PAM Odonto    | CAPS AD         |
|           | PAM Odonto    | CAPS III        | CEO           | UAA             |
|           | NAIPE         | CAF             | Patrimônio    | CAPS II         |
|           |               | Patrimônio      | CAF           | Resid. Terap.   |
|           |               | UAA             |               | Centrinho       |
| Quarta    | PA Leste      | Comasa          | PA Leste      | Jd. Iririu      |
|           | Iririu        | Jd. Iririu      | Iririu        | Comasa          |
|           |               |                 |               | Bakita          |
| Quinta    | PA Norte      | Unid. Sanitária | PA Norte      | Ambulatório     |
|           | PAM Bucarein  | CTA             |               | Unid. Sanitária |
|           | Laboratório   | Imunização      |               | Imunização      |
|           |               | Epidemiologia   |               | Epidemiologia   |
|           |               | Ambulatório     |               | CTA             |
|           |               |                 |               | CAPS III        |
| Sexta     | Rio Do Ferro  | Saguaçu         | P. Joinville  | PA Sul          |
|           | P. Joinville  | PAM Boa Vista   | Rio Do Ferro  | PAM Boa Vista   |
|           | Aventureiro I |                 | Aventureiro I |                 |

Fonte: Autor

## 5. CONCLUSÃO

Os resultados obtidos neste artigo mostraram a possibilidade de melhoria no desempenho logístico do setor público através de investimento em tecnologia e uso de sistemas de gerenciamento e otimização dos processos, como a roteirização de veículos. O setor público tem dificuldades de implementar as mesmas soluções da iniciativa privada, devido à dependência do sistema licitatório. A proposta de um novo cronograma de entregas baseado no resultado desse artigo

poderia ser implementada na prática, mas não necessariamente implicaria na alteração do contrato ao término de sua vigência.

Como sugestão para trabalhos futuros seria necessário um acompanhamento das entregas para se obter dados reais de tempo dos percursos e o tempo de parada nas unidades. Assim como estudar a possibilidade de trabalhar com o cronograma mensal como um todo, utilizando restrições de janela de tempo.

## REFERÊNCIAS

- [1]. ALVARENGA, Antônio Carlos; NOVAES, Antônio Galvão. Logística Aplicada. 3ª Edição. São Paulo, 2000.
- [2]. ASSINI, Danilo José; *et al.* Logística no Setor Público. CESUMAR. Maringá, 2012.
- [3]. BALLOU, Ronald H. Gerenciamento da Cadeia de Suprimentos: Logística Empresarial. Editora Bookman, 2006.
- [4]. BARBOSA, M. R. Aplicação de uma Abordagem Adaptativa de Busca Tabu a Problemas de Roteirização e Programação de Veículos. São Carlos – SP, 2005.
- [5]. COÊLHO, Eugênio Pacceli de Freitas. Logística de dispensação na rede de saúde pública. In: III CONGRESSO CONSAD DE GESTÃO PÚBLICA, 2010, Brasília. Anais. Brasília, 2010.

- [6]. FISHER, Marshall L.; JAIKUMAR, Ramchandran. A generalized assignment heuristic for vehicle routing. *Networks*, v. 11, n. 2, p. 109-124, 1981.
- [7]. LAPORTE, Gilbert. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, v. 59, n. 3, p. 345-358, 1992.
- [8]. NOVAES, Antônio Galvão. *Gerenciamento da Cadeia de Distribuição*. 2ª edição. São Paulo: Atlas, 2004.
- [9]. OLIVEIRA, Humberto César Brandão de. *Um Modelo Híbrido Estocástico para Tratamento de um Problema de Roteamento de Veículos com Janela de Tempo*. Tese de Mestrado. UFPE, Recife, 2007.
- [10]. TADEU, Hugo Ferreira Braga. *Logística Empresarial*. Belo Horizonte: Fundac-BH, 17 p. 2008.
- [11]. TRIPADALLI, Juarez Paulo. *Comércio Eletrônico: uma Perspectiva no Setor Público para Melhoria da Logística*. UFRJ. Rio de Janeiro – RJ, 2008.

# Capítulo 17

## MODELAGEM DE UM PROCESSO PRODUTIVO UTILIZANDO REDES NEURAS ARTIFICIAIS

*Simone Massulini Acosta*

*Anderson Levati Amoroso*

*Ângelo Márcio Oliveira Sant'Anna*

**Resumo:** Para melhorar a qualidade de seus produtos ou serviços as empresas buscam reduzir a variabilidade nos seus processos e produtos para que estes atendam as especificações preestabelecidas. Em um processo industrial muitas vezes não se consegue controlar todas as causas ou fatores de variação que produzem determinado efeito sobre uma ou mais características da qualidade dos produtos. Como resultado pode-se ter a produção de produtos não conformes às especificações, que podem ser mensuradas através da fração de produtos não conformes. A modelagem da fração de produtos não conformes pode ser realizada utilizando-se diferentes técnicas e modelos de regressão. As redes neurais artificiais são muito utilizadas para a modelagem de dados de processos. O objetivo deste artigo é modelar a fração de produtos não conformes às especificações de uma indústria curtidora de couro utilizando redes neurais artificiais e comparar os resultados obtidos com outros modelos de regressão desenvolvidos para o mesmo processo produtivo. O modelo de regressão utilizando redes neurais artificiais apresentou melhor desempenho no ajuste do modelo aos dados do processo quando comparado ao modelo de regressão de Quase-verossimilhança e ao modelo de regressão Beta e desempenho próximo ao modelo utilizando regressão por vetores suporte.

**Palavras-chave:** Inteligência computacional, redes neurais artificiais, modelagem de dados.

## 1. INTRODUÇÃO

Os processos produtivos estão sujeitos a variabilidade, que está relacionada com as diferenças existentes entre as unidades produzidas. Segundo Costa, Epprecht e Carpinetti (2010), se a variabilidade do processo for grande, as diferenças entre as unidades produzidas serão fáceis de observar mas, se a variabilidade for pequena, tais diferenças serão difíceis de observar.

O conceito de qualidade tem sido definido de diferentes maneiras na literatura. Para Montgomery (2004) a qualidade é inversamente proporcional à variabilidade e a melhoria da qualidade é a redução da variabilidade nos processos e produtos. Taguchi define que um produto de qualidade é aquele que atende perfeitamente às especificações atingindo o valor alvo com a menor variabilidade possível em torno dele. Os elementos que, em conjunto, descrevem a adequação de um produto ao uso podem ser denominados características da qualidade ou indicadores de desempenho (RIBEIRO; CATEN, 2012).

Em um processo industrial muitas vezes não se consegue controlar todas as causas ou fatores de variação que produzem determinado efeito sobre as características da qualidade dos produtos, pois certas causas são inerentes ao processo. As causas de variação que podem ser controladas podem resultar na produção de produtos com características da qualidade não conformes às especificações preestabelecidas, que podem ser mensuradas através da fração de produtos não conformes (MONTGOMERY, 2004).

A fração não conforme é definida como a razão entre o número de unidades não conformes da amostra e o tamanho da amostra e podem compreender a razão entre dois números discretos, denominada de percentual, ou a razão entre dois números contínuos, denominada de proporção. As características da qualidade do tipo fração possuem as observações no intervalo  $[0,1]$  e geralmente não possuem distribuição Normal (SANT'ANNA; CATEN, 2010).

A análise de regressão consiste de uma técnica estatística para modelar e investigar a relação entre as características da qualidade com os fatores controláveis que podem afetá-las (CHARNET et al., 2008). A modelagem da fração de produtos não conformes é uma relevante estratégia para investigar a relação

entre as causas de variação que interferem no processo e as características da qualidade dos produtos.

Segundo Sant'Anna e Caten (2010) os modelos estatísticos de regressão linear não são adequados quando se monitoram características da qualidade do tipo fração de produtos não conformes, pois esses pressupõem que a característica da qualidade possui distribuição Normal e possibilita a previsão de valores fora do limite do intervalo  $[0,1]$ . Os modelos estatísticos pertencentes à classe dos Modelos Lineares Generalizados e o modelo de regressão Beta são alternativas para a modelagem de variáveis com distribuição não Normal.

Um método muito utilizado para a modelagem de processos é baseado em Rede Neural Artificial (RNA). As RNAs são modelos de processamento de dados que emulam redes de neurônios biológicos, capazes de recuperar rapidamente uma grande quantidade de dados e reconhecer atributos baseados na experiência (HAYKIN, 2009).

As RNAs utilizam o princípio de minimização do risco empírico (*Empirical Risk Minimization*, ERM) baseado no princípio da indução para inferir uma função que minimize o erro sobre os dados de treinamento conhecidos e, desta forma, espera-se que esse procedimento leve também a um menor erro sobre os dados de teste (VAPNIK, 1998).

A aplicação de RNAs para modelar diferentes processos químicos e industriais tem sido apresentado na literatura nos últimos anos. Bakar et al. (2013) realizaram um estudo para identificar os parâmetros principais de um trocador de calor em uma planta de produção de bio metanol. Osulale e Zhang (2014) avaliaram diferentes modelos de RNAs para a previsão da carga em um mercado de eletricidade. Yong et al. (2015) utilizaram RNAs para minimizar o projeto de um permutador de calor, permitindo menor tempo e custo de fabricação. Acosta, Sant'Anna e Canciglieri (2016) analisaram a concentração de fósforo nas ligas de aço produzidas por uma empresa siderúrgica, pois altos níveis de fósforo causam danos as características físicas das ligas de aço.

O objetivo deste artigo é modelar a fração de produtos não conformes às especificações de uma indústria curtidora de couro utilizando Redes Neurais Artificiais e comparar os resultados obtidos com outros modelos de

regressão desenvolvidos para o mesmo processo produtivo.

## 2. REDES NEURAIS ARTIFICIAIS

Segundo Braga, Carvalho e Ludermir (2012), as RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples (neurônios artificiais), dispostas em uma ou mais camadas e interligadas por um grande número de conexões, que calculam determinadas funções matemáticas (normalmente não lineares). Na maioria dos modelos as conexões estão associadas a pesos, que armazenam o conhecimento adquirido pelo modelo e servem para ponderar a entrada recebida por cada neurônio da rede.

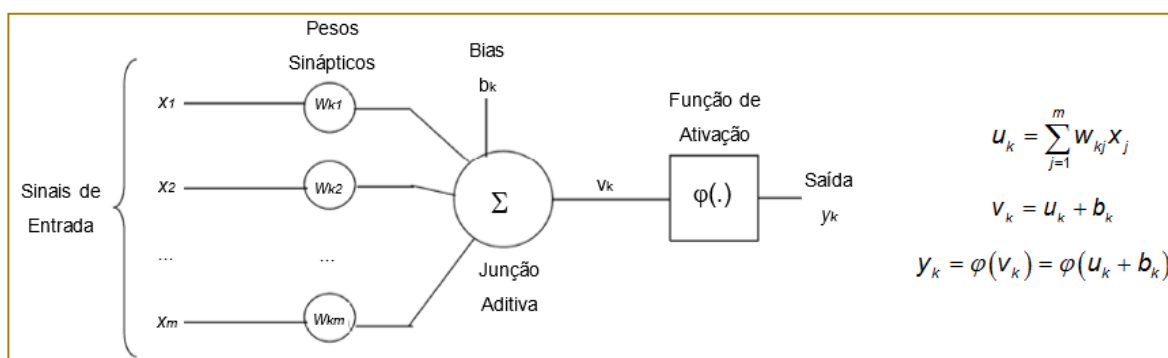
O procedimento usual na solução de problemas utilizando RNAs inicia por uma fase de aprendizagem, em que um conjunto de exemplos é apresentado para a rede, que extrai as características necessárias para

representar a informação fornecida. Essas características são utilizadas para gerar respostas para o problema (BRAGA; CARVALHO; LUDERMIR, 2012).

De acordo com De Castro (2006), as RNAs podem ser caracterizadas por três aspectos principais: (1) um grupo de neurônios artificiais, denominados nós, unidades ou neurônios; (2) um padrão de conectividade entre os neurônios, denominado de arquitetura ou estrutura; e (3) um método para determinar os valores dos pesos, denominado de algoritmo de treinamento ou de aprendizado.

Cada neurônio em uma rede neural recebe uma série de entradas. Uma função de ativação é aplicada a esses valores de entrada, o que resulta no nível de ativação do neurônio, que é o valor de saída do neurônio. A Figura 1 apresenta um modelo não linear de um neurônio artificial e sua representação matemática (HAYKIN, 2009).

Figura 1 - Modelo não linear de um neurônio artificial e sua representação matemática



Fonte: Adaptado de Haykin (2009)

Existem várias funções de ativação que podem ser utilizadas para o projeto de RNAs. As funções de base radial (por exemplo, gaussiana) e sigmoide (por exemplo, logística e tangente hiperbólica) são totalmente diferenciáveis e facilitam o processo de aprendizado da rede. A função sigmoide é a mais utilizada em RNAs (HAYKIN, 2009).

A arquitetura de uma RNA define a forma como seus neurônios estão arranjados, uns em relação aos outros, através do direcionamento das conexões sinápticas destes neurônios. A topologia de uma RNA pode ser definida como sendo as diferentes formas de composições estruturais que esta

poderá assumir (SILVA; SPATTI, FLAUZINO, 2010).

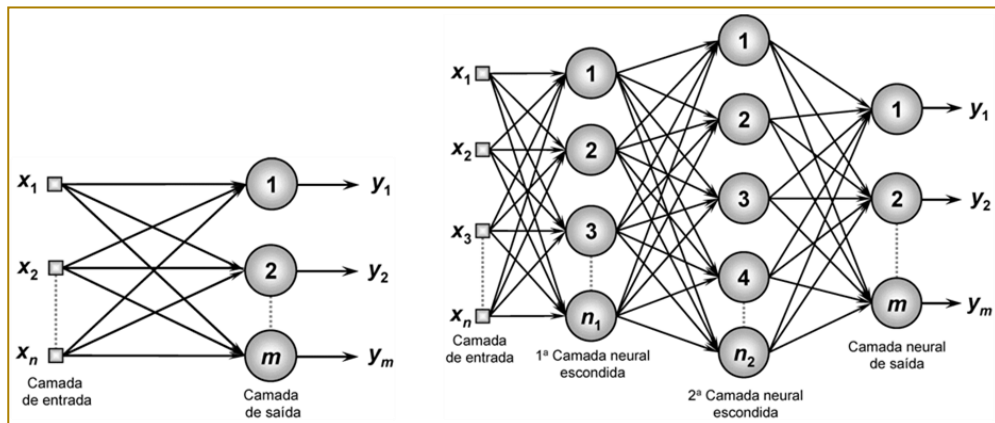
De acordo com De Castro (2006), em geral é possível distinguir três tipos principais de estruturas de RNAs: redes *feedforward* de camada única, redes *feedforward* de camadas múltiplas e redes recorrentes. Nas redes *feedforward* não existe conexão entre os neurônios da mesma camada e não existe realimentação entre as camadas. A Figura 2 apresenta um exemplo de uma rede *feedforward* de camada única e de uma rede *feedforward* de camadas múltiplas (SILVA; SPATTI; FLAUZINO, 2010).



Figura 2 - Exemplos de rede *feedforward*

(a) – Camada única

(b) - camadas múltiplas



Fonte: Silva, Spatti e Flauzino (2010)

Para resolver problemas não linearmente separáveis utilizando RNAs, a alternativa mais utilizada são as redes *feedforward* de camadas múltiplas. As camadas escondidas utilizam funções de ativação não lineares. Entre os principais tipos de redes com esta arquitetura estão o Perceptron Multicamadas (*Multilayer Perceptron*, MLP) e as Redes de Base Radial (*Radial Basis Function*, RBF) (FACELI et al., 2011).

Segundo Du e Swamy (2013), o teorema de Kolmogorov garante que uma rede *feedforward* com uma camada escondida, com um número suficiente de neurônios, pode aproximar qualquer função contínua com qualquer grau de precisão. Baseado no teorema de Kolmogorov, o teorema de Hecht-Nielsen apresenta que qualquer função contínua pode ser aproximada por uma rede *feedforward* com  $n$  nós de entrada,  $2n+1$  neurônios escondidos e  $m$  neurônios de saída. A utilização de duas camadas escondidas permite a aproximação de qualquer função (SCARCELLI; TSOI, 1998; FACELI et al., 2011).

Segundo Haykin (2009), a propriedade de importância primordial para uma RNA é a sua habilidade de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem. Uma RNA aprende através de um processo iterativo de ajustes dos pesos sinápticos e níveis de bias, tornando-se mais instruída sobre seu ambiente após cada iteração do processo de aprendizagem. Um conjunto preestabelecido de regras para a solução de um problema de aprendizagem é denominado de algoritmo de aprendizagem, tendo-se como objetivo final a

generalização de soluções a serem produzidas pelas saídas.

Os algoritmos de aprendizagem diferem entre si na maneira de obter o ajuste do peso sináptico de um neurônio, podendo ser agrupados em aprendizado supervisionado, não supervisionado e por reforço (BRAGA; CARVALHO; LUDERMIR, 2012).

No treinamento de redes *feedforward* de camada única por meio do aprendizado supervisionado e correção de erros, o erro é obtido diretamente através da diferença entre a saída desejada e a saída atual da rede. No entanto, nas redes *feedforward* de camadas múltiplas esse procedimento pode ser aplicado somente para a camada de saída, pois não existem saídas desejadas definidas para as camadas escondidas (DE CASTRO, 2006).

Conforme Braga, Carvalho e Ludermir (2012) para as redes MLP, que se caracterizam por utilizarem funções de ativação sigmóides nas camadas escondidas e sigmóides ou lineares na saída, é utilizado o algoritmo de retropropagação de erros ou *error backpropagation*. Este algoritmo utiliza o gradiente descendente e estima o erro das camadas escondidas por meio de uma estimativa do efeito que estas causam no erro da camada de saída. Existem várias modificações do algoritmo *backpropagation* que visam acelerar seu tempo de treinamento e melhorar seu desempenho, sendo uma destas modificações a Levenberg-Marquardt.

Normalmente, o procedimento para o desenvolvimento do modelo RNA segue os seguintes passos: obtenção dos dados,

análise e preparação dos dados, divisão dos dados, escolha da arquitetura da rede, seleção dos melhores parâmetros para a rede, simulações de treinamento e de teste e análise do desempenho. A Figura 3 apresenta um fluxograma geral do procedimento para a obtenção do modelo RNA.

Para análise do desempenho e avaliação da capacidade de generalização do modelo desenvolvido com RNAs podem ser utilizadas estratégias de minimização do erro de

predição ou erro residual. O erro residual ( $e_t$ ) é a diferença entre o valor real ( $y_i$ ) e o valor estimado pelo modelo ( $y_t$ ), conforme apresentado na equação (1). A equação (2) apresenta o erro médio absoluto (*Mean Absolute Error*, MAE), a equação (3) apresenta o erro médio quadrático (*Mean Squared Error*, MSE) e a equação (4) apresenta a raiz do erro médio quadrático (*Root Mean Squared Error*, RMSE), onde  $n$  representa o número de amostras.

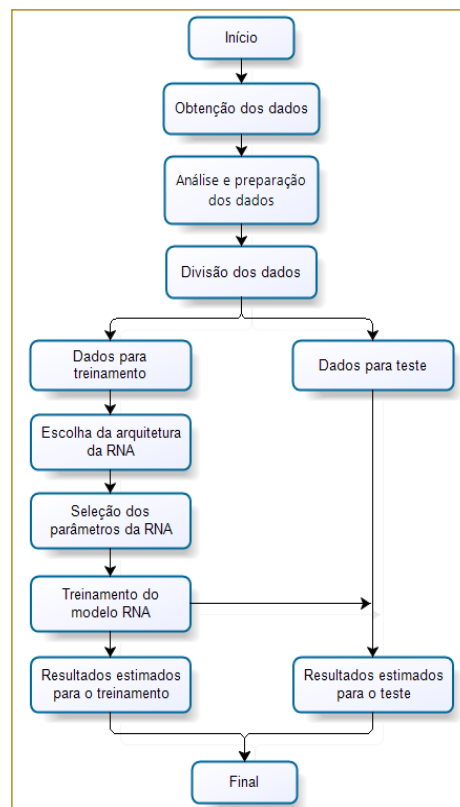
$$e_t = y_i - y_t \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_t| = \frac{1}{n} \sum_{i=1}^n |e_t| \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_t)^2 = \frac{1}{n} \sum_{i=1}^n (e_t)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_t)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_t)^2} \quad (4)$$

Figura 3 - Fluxograma para o desenvolvimento do modelo RNA



Fonte: Autoria própria

Outro critério que pode ser utilizado para verificar a adequação do modelo é o

Coefficiente de Informação de Akaike corrigido (*corrected Akaike Information Criterion*, AICc),

equação (5) (SANT'ANNA; CATEN, 2010).

$$AICc = \ln \left( \frac{\sum_{t=1}^n e_t^2}{n} \right) + \frac{2n(W+1)}{n-W-2} \quad (5)$$

onde  $n$  representa o número de amostras e  $W$  representa o número de parâmetros do modelo. Para as RNAs o número de parâmetros do modelo corresponde ao número de pesos e de *bias* a serem ajustados, sendo  $W = k(j+1) + i(j+1)$ , onde  $k$  é o número de entradas,  $j$  é o número de neurônios na camada oculta e  $i$  é o número de neurônios na camada de saída (YU; WANG; LAI, 2006).

### 3. ESTUDO APLICADO

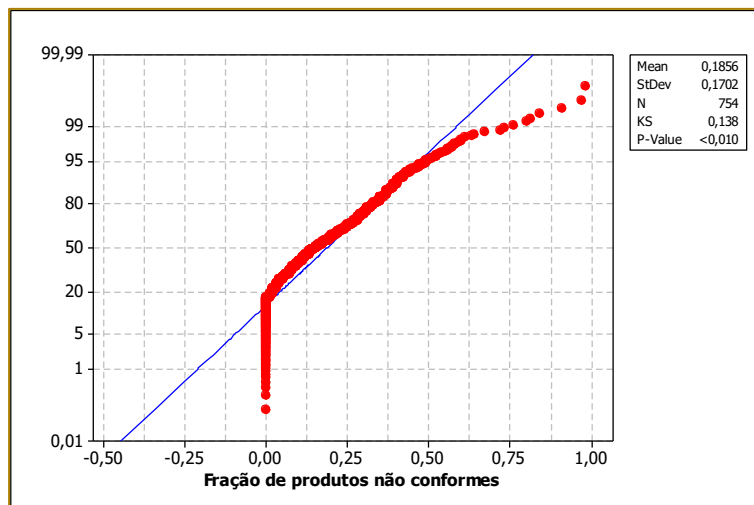
Sant'Anna e Caten (2010) apresentam a modelagem dos dados de uma empresa curtidora de couro, produtora de couro acabado e fornecedora para as indústrias de calçados e artefatos em couro, utilizando o modelo de regressão de Quase-verossimilhança, pertencente à classe dos Modelos Lineares Generalizados, e o modelo de regressão Beta.

A etapa *wet blue* do processo produtivo avaliado consiste em: o classificador recebe

um lote de diferentes tamanhos contendo as matérias-primas e verifica se as características de qualidade satisfazem às especificações, por métodos cognitivos. As matérias-primas que não satisfazem às especificações são classificadas como produtos não conformes e a fração de produtos não conformes às especificações, por lote, é considerada a variável dependente. Os dados coletados contemplam uma amostra de 754 lotes.

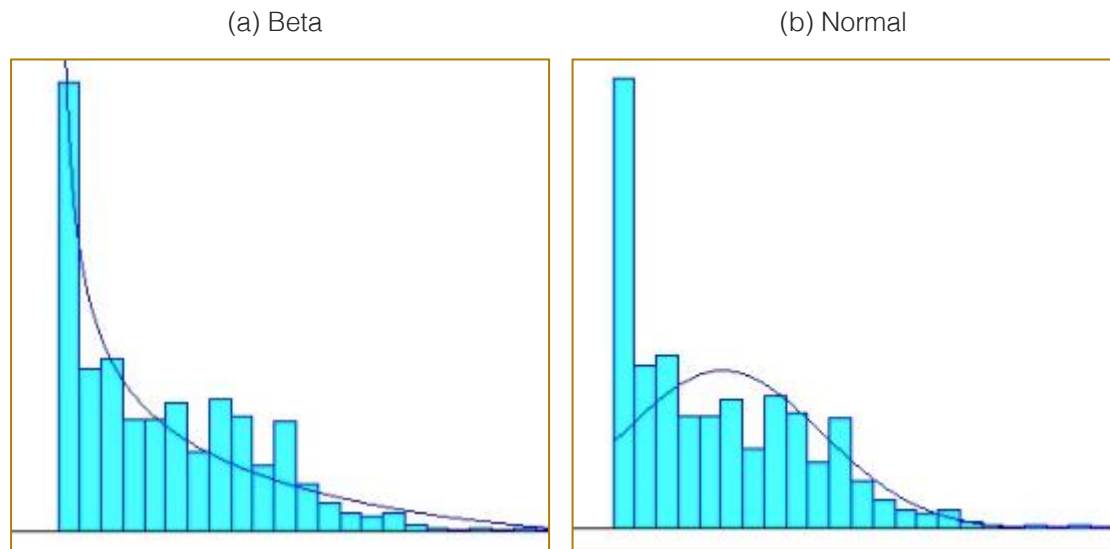
A fração de produtos não conformes do processo não segue a distribuição Normal. A Figura 4 apresenta o teste de normalidade dos dados do processo baseado no método Kolmogorov-Smirnov com 95% de confiança ( $\alpha = 0,05$ ), sendo o valor obtido para *p-value* menor que 0,01. A Figura 5 apresenta o histograma da fração de produtos não conformes, onde se observa que o comportamento dos dados não segue a distribuição Normal e segue a distribuição Beta.

Figura 4 - Teste de normalidade da fração de produtos não conformes



Fonte: Autoria própria

Figura 5 - Distribuição da fração de produtos não conformes com distribuição sobreposta:



Fonte: Autoria própria

Os fatores controláveis definidos como variáveis independentes para a modelagem da fração de produtos não conformes foram: a seleção da matéria prima conforme qualidade e preço ( $x_1$ ), com cinco níveis diferentes; a procedência da matéria-prima adquirida pela empresa ( $x_2$ ), com cinco níveis; o classificador que inspeciona as matérias-primas ( $x_3$ ), com três níveis, e o estado de rebaixamento da matéria-prima ( $x_4$ ), com dois níveis.

Para a construção dos modelos de regressão as variáveis independentes qualitatvas seleção, procedência, classificador e rebaixamento foram substituídas pelas variáveis *dummy*. As novas variáveis independentes foram definidas como: seleção tipo 2 ( $x_1$ ), seleção tipo 3 ( $x_2$ ), seleção tipo 4 ( $x_3$ ), seleção tipo 5 ( $x_4$ ), procedência 2 ( $x_5$ ), procedência 3 ( $x_6$ ), procedência 4 ( $x_7$ ), procedência 5 ( $x_8$ ), classificador 2 ( $x_9$ ), classificador 3 ( $x_{10}$ ) e rebaixamento ( $x_{11}$ ).

Para a definição das variáveis independentes relevantes para a construção do modelo de regressão foi realizado o teste de correlação de Pearson. Após o teste foram selecionadas as variáveis seleção (tipo 2, tipo 3, tipo 4 e tipo 5), classificador (2 e 3) e rebaixamento como estatisticamente significativas para explicar a variável dependente fração de produtos não conformes.

Para o desenvolvimento dos modelos baseados em RNAs foi utilizada a biblioteca para redes neurais do programa Matlab®. Após a análise e preparação dos dados o conjunto total de dados foi dividido randomicamente em dois subconjuntos: 70% para o grupo de treinamento e 30% para o grupo de teste.

A arquitetura selecionada para a RNA foi a MLP com sete variáveis de entrada, uma camada oculta e uma variável de saída. Após diversas simulações os parâmetros selecionados para as RNAs foram: algoritmo de treinamento Levenberg-Marquardt, função de ativação logística para a camada oculta e linear para a camada de saída e taxa de aprendizado de 0,01.

Para a obtenção do número de neurônios da camada oculta foi utilizado o teorema de Hecht-Nielsen e foram analisados MSE, RMSE, MAE e AICc para redes com diferentes números de neurônios na camada oculta. A Tabela 1 apresenta os resultados de algumas simulações com os dados de treinamento para a obtenção do número de neurônios da camada oculta. Em função dos resultados obtidos para os erros e para AICc foi selecionada a RNA com 19 neurônios na camada oculta.

Tabela 1 - Resultados das simulações para obtenção do número de neurônios da camada oculta

| Neurônios | 14      | 15      | 16      | 17      | 18      | 19      | 20      | 21      |
|-----------|---------|---------|---------|---------|---------|---------|---------|---------|
| MSE       | 0,01407 | 0,01285 | 0,01288 | 0,01281 | 0,01852 | 0,01262 | 0,01269 | 0,01271 |
| RMSE      | 0,1186  | 0,1134  | 0,11349 | 0,1132  | 0,1361  | 0,11235 | 0,11265 | 0,11275 |
| MAE       | 0,09502 | 0,08939 | 0,08934 | 0,08763 | 0,1138  | 0,08850 | 0,08727 | 0,08878 |
| AICc      | 339,10  | 372,46  | 407,66  | 444,705 | 484,28  | 524,935 | 568,49  | 614,61  |

Fonte: Autoria própria

Após a seleção da arquitetura e dos parâmetros, a RNA foi treinada e, após a fase de treinamento, o modelo obtido foi utilizado para estimar as saídas utilizando os dados do grupo de teste. A Figura 6 apresenta que os valores estimados pelo modelo RNA se aproximam dos valores dos dados experimentais do processo do grupo de teste.

A Tabela 2 apresenta os valores dos erros calculados para os dados do grupo de

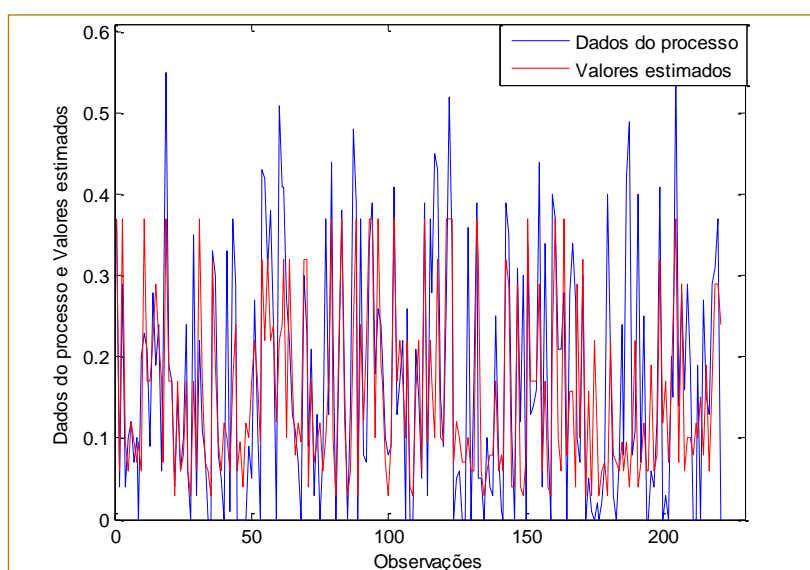
treinamento e do grupo de teste. Os valores dos erros obtidos pelo modelo RNA para os grupos de treinamento e de teste não diferem significativamente, indicando não haver sobreajuste do modelo. Analisando a Figura 6 e a Tabela 2 pode-se considerar que o modelo RNA representa adequadamente a fração de produtos não conforme do processo.

Tabela 2 - Valores dos erros calculados para os grupos de treinamento e de teste

| Erro | Grupo de treinamento | Grupo de teste |
|------|----------------------|----------------|
| MSE  | 0,0126               | 0,0138         |
| RMSE | 0,1123               | 0,1176         |
| MAE  | 0,0885               | 0,0937         |

Fonte: Autoria própria

Figura 6 - Valores dos dados do processo do grupo de teste e valores estimados pelo modelo RNA



Fonte: Autoria própria

Na sequência foi realizada a comparação do modelo RNA com o modelo de regressão de Quase-verossimilhança e modelo de regressão Beta apresentados em Sant'Anna e Caten (2010) e com o modelo utilizando regressão por vetores suporte apresentado em Acosta, Amoroso e Sant'Anna (2016).

A forma de regressão para o modelo de Quase-verossimilhança é descrita pela equação (6) e para o modelo Beta é descrita pela equação (7) (SANT'ANNA; CATEN, 2010).

$$g(\hat{\mu}) = -3,7692 + 2,0475 * x_1 + 2,2857 * x_2 + 2,5226 * x_3 + 2,7583 * x_4 + 0,2672 * x_9 + 0,4483 * x_{10} - 0,8357 * x_{11} \quad (6)$$

$$g(\hat{\mu}) = -2,4225 + 1,5822 * x_1 + 1,1202 * x_2 + 1,2295 * x_3 + 1,4469 * x_4 + 0,2883 * x_9 + 0,3740 * x_{10} - 0,6807 * x_{11} \quad (7)$$

A aplicação de máquinas de vetores suporte (*Support Vector Machine, SVM*) foi originalmente desenvolvida para problemas de classificação de padrões, sendo posteriormente estendida para o tratamento de problemas de regressão e denominada regressão por vetores suporte (*Support Vector Regression, SVR*). O modelo utilizando SVR possui o parâmetro C igual a 10, o parâmetro  $\epsilon$  igual a 0,03 e a função *kernel* RBF com parâmetro  $\gamma$  igual a 0,2 (ACOSTA; AMOROSO; SANT'ANNA, 2016).

Os dados do grupo de teste utilizados no modelo RNA foram utilizados com o modelo Quase-verossimilhança, modelo Beta e modelo SVR. A Tabela 3 apresenta os valores dos erros calculados para os quatro modelos. Analisando os valores dos erros verifica-se que o modelo RNA possui um desempenho melhor para a modelagem da fração de produtos não conformes às especificações do que o modelo Quase-verossimilhança e modelo Beta. O modelo RNA possui um desempenho um pouco menor, porém próximo, ao do modelo SVR.

Tabela 3 - Valores dos erros calculados para o grupo de teste nos modelos

| Modelo                | MSE    | RMSE   | MAE    |
|-----------------------|--------|--------|--------|
| Quase-verossimilhança | 0,0161 | 0,1268 | 0,0998 |
| Beta                  | 0,0151 | 0,1231 | 0,0961 |
| SVR                   | 0,0131 | 0,1144 | 0,0853 |
| RNA                   | 0,0138 | 0,1176 | 0,0937 |

Fonte: Autoria própria

As máquinas de vetores suporte utilizam o princípio de minimização do risco estrutural (*Structural Risk Minimization, SRM*) que envolve a minimização de um limite superior sobre o erro de generalização. Vapnik, Golowich e Smola (1996) mostraram que o princípio SRM é superior ao princípio ERM, utilizado em redes neurais artificiais. Desta forma, modelos de aprendizado de máquina baseados no princípio SRM tendem a

apresentar uma maior habilidade para generalizar bem frente a dados de teste não observados (VAPNIK, 1998).

#### 4. CONCLUSÕES

Os efeitos das variáveis independentes (fatores controláveis) sobre a variável dependente (característica da qualidade) em processos industriais podem ser analisados

através da modelagem dos dados destes processos usando modelos de regressão. Para esta modelagem podem ser utilizadas técnicas estatísticas, como os modelos lineares generalizados e o modelo de regressão Beta, e técnicas de inteligência computacional, como as redes neurais artificiais e a regressão por vetores suporte.

Neste artigo foram utilizadas as redes neurais artificiais para a modelagem da fração de produtos não conformes às especificações de uma indústria curtidora de couro em que a característica da qualidade é mensurada no intervalo  $[0,1]$ . O modelo obtido foi comparado com outros modelos de regressão desenvolvidos para o mesmo processo produtivo.

Analisando os valores dos erros obtidos para os modelos verifica-se que o modelo RNA apresentou melhor desempenho no ajuste do modelo aos dados do que o modelo Quase-verossimilhança e o modelo Beta e um desempenho próximo ao do modelo SVR.

Através dos resultados verifica-se que os modelos obtidos utilizando as duas técnicas de inteligência computacional obtiveram um desempenho melhor do que os modelos utilizando as duas técnicas estatísticas. Estes resultados estão de acordo com trabalhos disponíveis na literatura onde os autores concluem que as redes neurais artificiais e a regressão por vetores suporte apresentam desempenho similar ou superior quando comparadas as técnicas estatísticas convencionais. Apesar destas técnicas de inteligência computacional apresentarem bons resultados no modelamento de processos, estas ainda são mais complexas que as técnicas estatísticas convencionais.

Para trabalhos futuros sugere-se a modelagem da fração de produtos não conformes às especificações de diferentes processos industriais utilizando redes neurais artificiais e regressão por vetores suporte e a comparação com modelos obtidos através de técnicas estatísticas convencionais.

## REFERÊNCIAS

- [1]. ACOSTA, Simone M.; SANT'ANNA, Ângelo M. O.; CANGIOLIERI Osiris J. Forecasting modeling for energetic efficiency in an industrial process. *Chemical Engineering Transactions*, v. 52, p. 1081-1086, 2016.
- [2]. ACOSTA, Simone M.; AMOROSO, Anderson L.; SANT'ANNA, Ângelo M. O. Modelagem de um processo produtivo utilizando regressão por vetores suporte. In: Encontro Nacional de Engenharia de Produção, 2016, João Pessoa. Encontro Nacional de Engenharia de Produção. Rio de Janeiro: ABEPRO, 2016.
- [3]. BAKAR, Suraya H. A. et al. A. Flexible and operable heat exchanger networks. *Chemical Engineering Transactions*, v. 32, p. 1297-1302, 2013.
- [4]. BRAGA, Antônio de P.; CARVALHO, André P. de L. F.; LUDERMIR, Teresa B. *Redes neurais artificiais: teoria e aplicações*. 2. ed. Rio de Janeiro: LTC, 2012.
- [5]. CHARNET, Reinaldo et al. *Análise de modelos de regressão linear com aplicações*. Campinas: Editora da Unicamp, 2008.
- [6]. COSTA, Antônio F. B.; EPPRECHT, Eugênio K.; CARPINETTI, Luiz C. R. *Controle estatístico de qualidade*. 2. ed. São Paulo: Atlas, 2010.
- [7]. DE CASTRO, Leandro N. *Fundamentals of natural computing: basic concepts, algorithms, and applications*. London: Chapman & Hall/CRC, 2006.
- [8]. DU, Ke-Lin; SWAMY, M. N. S. *Neural networks and statistical learning*. Heidelberg: Springer, 2013.
- [9]. FACELI, Katti et al. *Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro: LTC, 2011.
- [10]. HAYKIN, Simon. *Neural networks and learning machines*. 3rd ed. New York: Prentice Hall, 2009.
- [11]. MONTGOMERY, Douglas C. *Introdução ao controle estatístico da qualidade*. Rio de Janeiro: LTC, 2004.
- [12]. OSUOLALE, F. N., ZHANG, J. Energy efficient control and optimization of distillation column using artificial neural network. *Chemical Engineering Transactions*, n. 39, p. 37-42, 2014.
- [13]. RIBEIRO, José L. D.; CATEN, Carla S. *Controle estatístico do processo*. Série

Monográfica Qualidade. Porto Alegre: FEENG/UFRGS, 2012.

[14]. SANT'ANNA, Ângelo M. O.; CATEN, Carla S. Modelagem da fração de não-conformes em processos industriais. Pesquisa Operacional, v. 30, n. 1, p. 53-72, janeiro/abril, 2010.

[15]. SCARCELLI, Franco, TSOI, Ah C. Universal approximation using feed-forward neural networks: a survey of some existing methods and some new results. Neural Networks, v. 11, p. 15-37, 1998.

[16]. SILVA, Ivan N. da; SPATTI, Danilo H.; FLAUZINO, Rogério A. Redes neurais artificiais: para engenharia e ciências aplicadas. São Paulo: Artliber: 2010.

[17]. VAPNIK, Vladimir N.; GOLOWICH, Steven; SMOLA, Alex J. Support vector [21].

method for function approximation, regression estimation and signal processing. Advances in Neural Information Processing Systems, v. 9, p. 281-287, 1996.

[18]. VAPNIK, Vladimir N. Statistical learning theory. New York: John Wiley & Sons, 1998.

[19]. YONG, J. Y. et al. Heat exchanger network modification for waste heat utilization under varying feed conditions. Chemical Engineering Transactions, n. 43, p. 1279-1284, 2015.

[20]. YU, L., WANG, S., LAI, K.K. An integrated data preparation scheme for neural network data analysis. IEE Transactions on Knowledge and Data Engineering, v. 18, p. 217-230, 2006.



# Capítulo 18

## *DESENVOLVIMENTO DO MENOR CAMINHO PARA A MELHORIA DAS LINHAS DE ÔNIBUS EM UM BAIRRO NO MUNICÍPIO DE ARACAJU - SE*

*Tayane Magalhaes Alvaia*

*Hellen Mariany Santos*

*Marcos Wandir Nery Lobao*

*Jose Ricardo Menezes Oliveira*

*Glaucia Regina de Oliveira Almeida*

**Resumo:** A população da cidade de Aracaju sofre diariamente com o transporte público, gerando reclamações durante muitos anos. Com a existência deste problema, foi decidido formular um projeto para solucioná-lo e melhorar a qualidade de vida da população. O projeto consiste na solução do problema da má formulação das linhas de ônibus na cidade de Aracaju, mais especificamente no bairro Jabotiana, utilizando a Pesquisa Operacional, com os itinerários das linhas presentes em tal bairro sendo analisadas. A partir daí, os dados relativos à distância entre os pontos de ônibus serão coletados e analisados, formulando então as funções-objetivo referentes a cada linha, com suas respectivas restrições. Com isso, utilizando o método do caminho mínimo, uma nova rota será elaborada, com mais ou menos pontos de parada e menor quilometragem, objetivando diminuir o tempo do trajeto ao ser comparado com a rota anterior, beneficiando os usuários do transporte coletivo.

\* Artigo originalmente apresentado no Encontro Nacional de Engenharia de Produção (ENEGEP, 2017)

## 1. INTRODUÇÃO

As grandes cidades brasileiras enfrentam diariamente problemas de transporte, principalmente referente ao setor público, visto que é o principal meio usado pelas pessoas no deslocamento diário. Na capital de Sergipe, Aracaju, o ônibus é o principal veículo utilizado nesse sistema. A cidade hoje encontra-se com diversos problemas na gestão dos serviços de transporte. Frequentemente, passageiros reclamam em meios de comunicação sobre o tempo de espera pelos ônibus, a escassez de linhas ou a qualidade na conservação, sendo prejudicial para todos os usuários de transporte público. Com isso, tornou-se necessário um plano de rotas para o melhor atendimento aos usuários.

Utilizou-se da pesquisa operacional como alternativa à solução do problema, através da formulação de uma proposta de rota para os ônibus de Aracaju, utilizando o método do menor caminho, com o intuito de reduzir o tempo de espera entre os ônibus, como melhorar a qualidade de vida dos seus usuários.

Em particular, para os moradores do bairro Jabotiana as queixas não são diferentes, ao contrário, são intensificadas pelo fato de existir apenas duas vias de acesso a esse bairro. A partir de dados do IBGE (2010), a Jabotiana é um bairro da zona sul de Aracaju, e conta com aproximadamente 20.000 habitantes. O bairro passa por uma explosão imobiliária. No que diz respeito ao transporte coletivo possui uma linha radial que o liga ao Centro: 706-Santa Lúcia/Centro. Existem ainda quatro linhas alimentadoras que atravessam o bairro e fazem ligação ao Terminal do Distrito Industrial de Aracaju (DIA): 402.1-Santa Lúcia/DIA 01, 402.2-Santa Lúcia/DIA 02 e 402.3-Santa Lúcia/DIA 03.

## 2. PESQUISA OPERACIONAL

A Pesquisa Operacional (PO) refere-se à modelagem matemática de fenômenos dinâmicos ou estáticos. Prado (1998), diz a respeito da pesquisa operacional: "... permite estabelecer a 'mistura' ótima de diversas variáveis segundo uma função linear de efetividade e satisfazendo a um conjunto de restrições lineares para estas variáveis", sendo esta uma ciência que permite atingir os melhores resultados possíveis.

Existe uma ordem a seguir para a sintetização de um modelo de Pesquisa Operacional (HILLIER; LIEBERMAN, 2006):

Definir o problema a ser utilizado e coletar os dados necessários;

Formular um modelo matemático para representar o problema;

Desenvolver um algoritmo a partir do modelo para derivar soluções a fim de solucionar o problema;

Testar o modelo e aperfeiçoá-lo caso seja necessário;

Organizar uma aplicação contínua do modelo a partir do que foi requisitado pela gerência;

Implementar.

Além disso, as técnicas de pesquisa operacional são métodos de análise matemática que fornecem uma base quantitativa para decisões de gerenciamento. Ela pode minimizar o custo a fim de alcançar um melhor sistema de transporte, por exemplo, maximizando outros fatores necessários para esta melhoria. Comprovando que a Pesquisa Operacional é algo crucial para a gestão (AGARANA et al., 2017).

A Pesquisa Operacional possui diversas vertentes, sendo uma destas a programação linear.

### 2.1. PROGRAMAÇÃO LINEAR

A programação linear utiliza um modelo matemático para descrever um determinado problema. As funções matemáticas utilizadas nesse modelo são lineares, criando então um modelo que é empregado como um planejamento, não como programação envolvendo computação. Com isso, a programação linear envolve este planejamento de atividades a fim de obter um resultado ótimo, ou seja, um resultado que a partir do modelo matemático atinja o melhor objetivo especificado entre as alternativas viáveis (HILLIER; LIEBERMAN, 2006).

Uma função linear é do tipo  $f: \mathbb{R} \rightarrow \mathbb{R}$ , onde  $y = ax + b$ , sendo  $a$  e  $b \in \mathbb{R}$ . Caso seja representado em um gráfico, a representação será uma reta. Na programação linear, esta função é maximizada ou minimizada, caracterizando uma função-objetivo, sem infringir um sistema linear de igualdades ou

desigualdades que são denominadas restrições do modelo.

De acordo com Nossa e Chagas (1997), restrição é "... aquilo que impede um melhor desempenho de um sistema e representa normalmente limitações de recursos disponíveis ou exigências e condições que devem ser cumpridas no problema". Tais restrições demarcam uma região de um plano na qual o conjunto das soluções viáveis é inserido. A melhor dessas soluções é apontada como solução ótima. A partir daí é possível ver o objetivo da programação linear, que é determinar a solução ótima de cada função-objetivo para solucionar um problema.

Para atingir o objetivo de solucionar um problema, dois procedimentos devem ser realizados:

Determinar o problema, identificando o objetivo, as variáveis que estão ligadas a ele (denominadas variáveis de decisão) e as relações e limitações a que elas estão sujeitas;

Codificar o modelo em uma linguagem simbólica utilizando equações e/ou inequações matemáticas lineares.

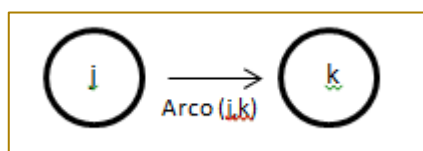
Com isso, o ambiente em que o problema está localizado é representado por um modelo matemático, o que libera a aplicação de métodos e técnicas matemáticas que facilitam ou até atingem a solução do problema (BODANESE et al., 2005).

Assim como a Pesquisa Operacional, a programação linear apresenta suas vertentes, sendo uma delas a otimização de redes, a qual leva para o caminho mais curto.

### 2.1.1. OTIMIZAÇÃO DE REDES: MÉTODO DO CAMINHO MÍNIMO

Segundo Winston e Goldberg (2004), uma rede é definida por dois elementos: Nós e arcos. Os nós são um conjunto de pontos, ou vértices  $V$ . Já um arco  $A$  é um par ordenado de vértices, representando uma possível direção do movimento que pode ocorrer entre vértices. Caso uma rede contenha um arco  $(j,k)$ , então o movimento é possível do nó  $j$  para o nó  $k$ , como mostra a Figura 1.

Figura 2 - Esquema de um arco entre os nós  $j$  e  $k$ .



Fonte: Autor próprio (2017)

Uma sequência de arcos em que cada um tem exatamente um vértice em comum com o arco anterior é chamado de caminho. Já um ciclo é um caminho onde o nó de destino de cada arco é idêntico ao nó de origem do próximo arco (WINSTON; GOLDBERG, 2004).

Se um fluxo representado por um arco é permitido em apenas uma direção, é chamado de arco direcionado. Caso esse fluxo possa percorrer ambas as direções, o arco é dito arco não-direcionado. Da mesma forma, se a rede possuir apenas arcos direcionados é chamada de rede direcionada, assim como se possuir apenas arcos não-direcionados ela passa a ser uma rede não-direcionada. Porém, uma rede com os dois tipos de arcos pode ser convertida para uma rede direcionada. (HILLIER; LIEBERMAN, 2006).

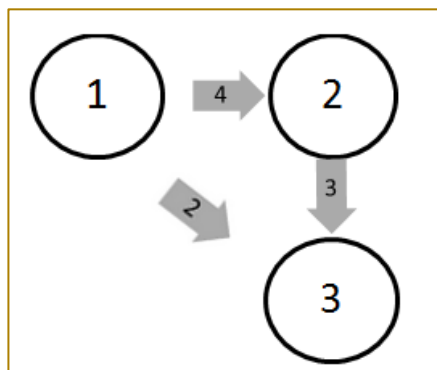
Ao elaborar o modelo matemático, as variáveis de decisão são consideradas como os arcos, sendo representado por  $x_{ij}$ , com  $i$  sendo o nó de origem e  $j$  o nó de destino. Também é importante a noção do balanço do nó, onde a soma dos arcos que saem é equivalente aos arcos que chegam. Caso  $x_{ij}$  seja considerado como parte do caminho mínimo, é considerado como um fluxo de 1. Se não for,  $x_{ij} = 0$ .

O problema do caminho mínimo é um dos métodos disponíveis para interagir com a rede. Como o próprio nome já diz, e serve para encontrar um menor caminho entre o nó de origem e o nó de destino em observância das restrições existentes (LACHTERMACHER, 2009).

O seguinte exemplo, na Figura 2, pode ser realizado para melhor compreensão do

método do caminho mínimo:

Figura 3- Esquema do exemplo do caminho mínimo



Fonte: Autor próprio (2017)

Variáveis de decisão:

$x_{12}$ : Caminho do nó 1 para o nó 2;

$x_{13}$ : Caminho do nó 1 para o nó 3;

$x_{23}$ : Caminho do nó 2 para o nó 3;

Balanços de nó (Restrições):

Nó 1:  $x_{12} + x_{13} = 1$   $x_{ij} \geq 0, \forall i \text{ e } \forall j$

Nó 2:  $-x_{12} + x_{23} = 0$

Nó 3:  $-x_{13} - x_{23} = -1$

Função Objetivo:  $\text{Min } Z = 4x_{12} + 2x_{13} + 3x_{23}$

$Z = 2$   $x_{13} = 1, x_{12} \text{ e } x_{23} = 0$

Menor caminho:  $1 \rightarrow 3$

Várias áreas podem ser beneficiadas utilizando o método do caminho mínimo. Planejar capacidades, desenhar rotas de transporte, coordenar fluxos de caixa e planejar rotas de transporte são algumas das áreas abrangidas pelo método do caminho mínimo.

### 3. METODOLOGIA

Foram analisados os itinerários das linhas de ônibus percorrem o bairro Jabotiana através do banco de dados da Superintendência Municipal de Transportes e Trânsito de Aracaju - SMTT (2017). Em seguida, a localização dos pontos foi demarcada após seguir o percurso como passageiros dentro do ônibus de cada linha. Após demarcar todos, as distâncias entre os pontos foram traçadas a partir do aplicativo Google Maps

(2017), possibilitando a criação dos trajetos que foram utilizados como base para a criação das variáveis de decisão, função objetivo e restrições. Vale ressaltar que enquanto os trajetos eram criados, novas opções de rotas foram traçadas, tendo como base as reclamações dos passageiros e caminhos alternativos, para a possibilidade de reduzir o tempo modificando o caminho. Além disso, o Google Maps fornece apenas dados arredondados, existindo uma margem de erro nas rotas otimizadas.

Com as funções necessárias obtidas, os dados foram inseridos no Microsoft Excel®, seguido da aplicação do Solver® utilizando a metodologia proposta por Hillier e Lieberman (2013) para a obtenção do caminho mínimo seguindo os seguintes passos:

Listaram-se os arcos utilizando duas colunas, uma "De" que indica em qual nó inicia o arco e outra "Para" informando em qual nó o arco chega;

Ao longo dos arcos, na coluna seguinte, listou-se a distância em metros entre os nós;

Uma coluna denominada "Na Rota" foi criada subsequente à coluna das distancias, esta coluna representa as variáveis de definição e o Solver, quando aplicado, indica o valor dessas variáveis;

Em outra coluna foram indicados os "Nós";

Subsequente à coluna dos nós foi criada a coluna de "Fluxo", que é equacionado no Excel com a função SOMASE;

Uma coluna de “Oferta/Demanda” foi preenchida com os valores 1 para origem, -1 para destino e 0 para demais;

Uma célula equacionada no Excel com a função SOMARPRODUTO selecionando as colunas “Na Rota” e “Distância” é inserida e representa a função objetivo no Solver;

O Solver foi inicializado e as células correspondentes à função objetivo foram selecionadas, foi determinada a célula que será Min Z, pois o objetivo será a menor resposta, assim como as células onde serão inseridos os valores das variáveis;

As células correspondentes aos valores dos coeficientes envolvidos nas restrições foram selecionadas e igualadas entre si;

Os valores entregues pelo programa em cada célula pré-determinada foram analisadas, sendo esse o caminho mínimo a ser utilizado.

Consequentemente foram calculadas as diferenças entre as rotas originais e as obtidas pelo Solver (em metros), e o quanto elas podem influenciar durante um dia de itinerário a partir do tempo percorrido pelas rotas originais, utilizando proporção. A quantidade de vezes que a rota percorre o itinerário também é definida pela SMTT (2017).

#### 4. RESULTADOS E DISCUSSÕES

Nesta seção são apresentados os resultados da aplicação do método do menor caminho nas linhas de ônibus do bairro Jabotiana. Todas as linhas partem do Terminal de

Variáveis de decisão:

$x_{ij}$ : Caminho de  $i$  até  $j$ ,  $\forall i$  e  $\forall j$ .

$$x_{ij} = \begin{cases} 0 & \text{se não for caminho} \\ 1 & \text{se for caminho} \end{cases}$$

Função objetivo:

$$Z_{min} = \sum_{(i,j) \in A} c_{ij} x_{ij}$$

Sendo os balanços de nó:

$$\sum_{\{j:(i,j) \in A\}} x_{ij} - \sum_{\{j:(j,i) \in A\}} x_{ji} = \begin{cases} n-1 & \text{para } i = s \\ -1 & \text{para todo } i \in N - \{s\} \end{cases}$$

$$x_{ij} \geq 0, \forall (i,j) \in A$$

Esse modelo também foi abordado por Méndez e Guardia (2008).

Integração do Distrito Industrial de Aracaju (D.I.A.) sentido bairro Jabotiana, e, ao realizarem o percurso do bairro, voltam para o mesmo Terminal de Integração. Diante deste fato, todas as linhas estudadas foram divididas em duas rotas, a primeira denominada “IDA” realiza o trajeto com ponto de partida na Jabotiana, passando pelos pontos obrigatórios, e segue em direção ao Terminal de Integração do D.I.A., a segunda rota, denominada “VOLTA”, realiza o caminho inverso, ou seja, inicia no Terminal de Integração do D.I.A. e tem o seu fim no ultimo ponto de ônibus do bairro.

Os resultados da aplicação do menor caminho são apresentados no modelo de diagrama de redes. Cada nó simboliza um ponto de ônibus e cada arco representa a direção e a distância entre dois pontos. Para melhor entender o diagrama de rede deve-se saber que:

Os nós de cor branca são os pontos de ônibus já existentes enquanto os de cor bege são novos pontos sugeridos;

Os arcos no formato uniforme são os caminhos já existentes, enquanto os tracejados são os caminhos sugeridos;

Por fim, a rota na cor vermelha é a rota otimizada encontrada através do método do caminho mínimo a partir do *Solver®*.

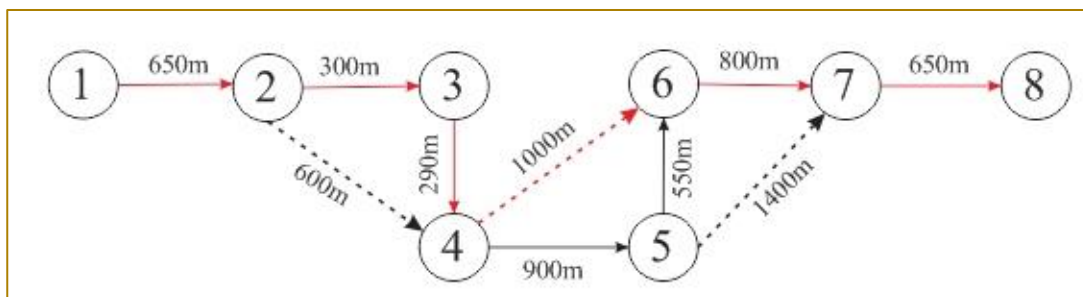
Como explanado no tópico 2.1.1 deste artigo, existe uma modelagem matemática contendo variáveis de decisão, função objetivo e restrição (ou balanço de nó). Para as rotas estudadas a modelagem segue o seguinte padrão de equações:

As linhas de ônibus estudadas, como mencionadas anteriormente, foram as

seguintes: 402.1-Santa Lúcia/DIA 01, 402.2-Santa Lúcia/DIA 02, 402.3-Santa Lúcia/DIA 03 e 706-Santa Lúcia/Centro. Os resultados e discussões serão apresentados a seguir.

Para a linha 402.1-Santa Lúcia/DIA 01 da rota "IDA" obteve-se diagrama de redes exibido na Figura 3:

Figura 4 - Diagrama de redes da linha 402-1 da rota "IDA"



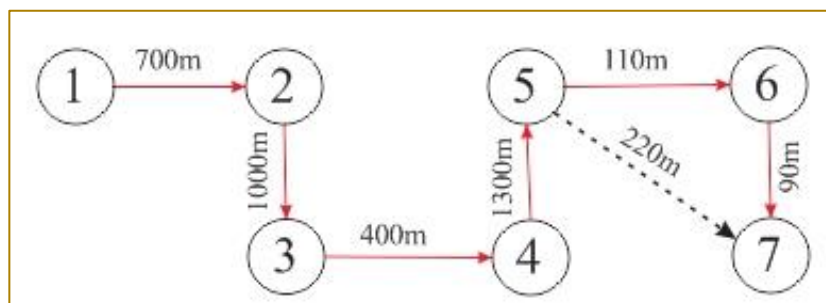
Fonte: Autor próprio (2017)

A rota atual percorre um total de 4140m em 15 min enquanto a rota otimizada percorre 3690m em 13,2 min obtendo assim uma redução de 450m e 1,8min por percurso. Esse trajeto é realizado 38 vezes por dia,

alcançando uma redução total de 17,1Km e 68,4 min por dia.

Para a linha 402.1-Santa Lúcia/DIA 01 da rota "VOLTA" obteve-se diagrama de redes exibido na Figura 4:

Figura 5 - Diagrama de redes da linha 402-1 da rota "VOLTA"

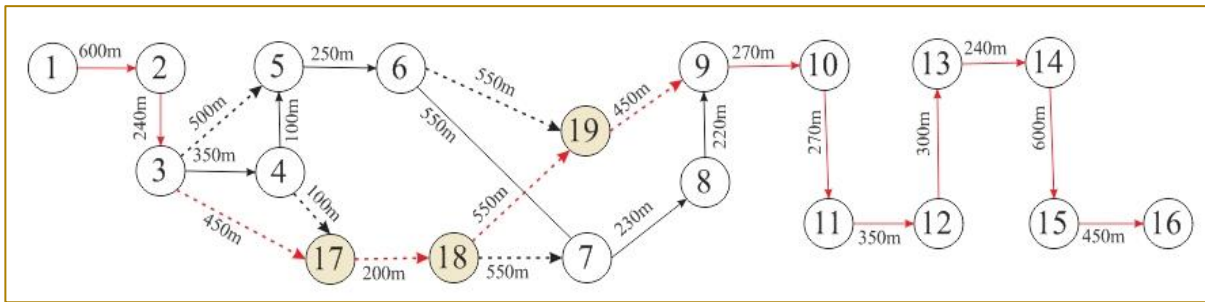


Fonte: Autor próprio (2017)

A rota atual, para esta linha, é a própria rota otimizada indicada pelo Solver, não havendo, portanto, uma redução de quilometragem ou tempo de trajeto.

Para a linha 402.2-Santa Lúcia/DIA 02 da rota "IDA" obteve-se diagrama de redes exibido na Figura 5:

Figura 6 - Diagrama de redes da linha 402-2 da rota "IDA"



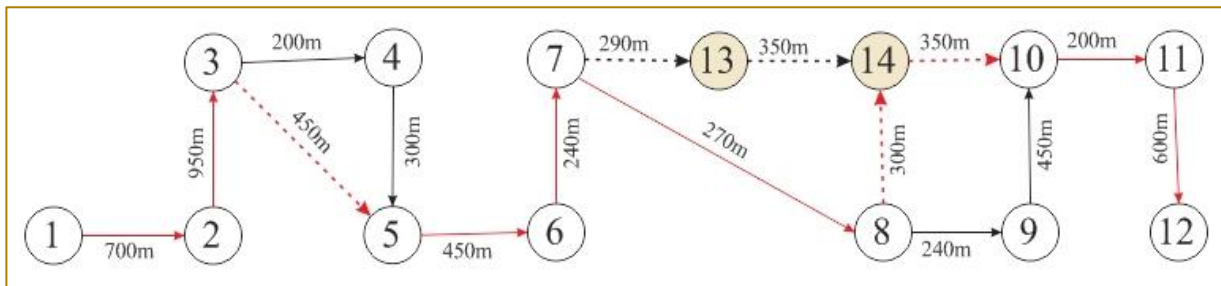
Fonte: Autor próprio (2017)

A rota atual percorre um total de 5020m em 15 min enquanto a rota otimizada percorre 4970m em 14,51 min obtendo assim uma redução de 50m e 0,49min por percurso. Esse trajeto é realizado 61 vezes por dia,

alcançando uma redução total de 3050m e 29,89 min por dia.

Para a linha 402.2-Santa Lúcia/DIA 02 da rota "VOLTA" obteve-se diagrama de redes exibido na Figura 6:

Figura 7 - Diagrama de redes da linha 402-2 da rota "VOLTA"



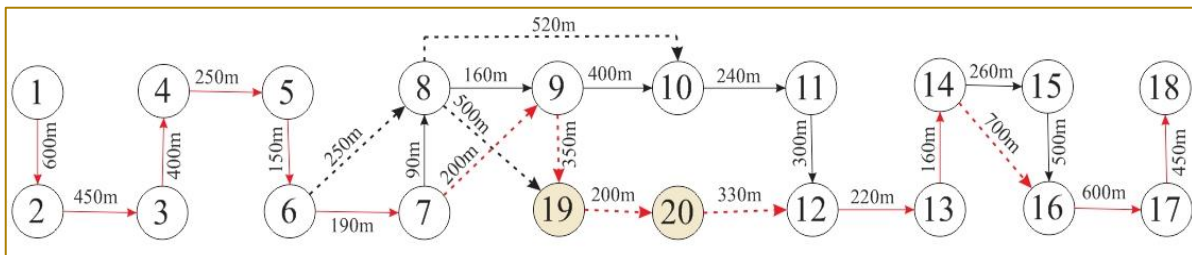
Fonte: Autor próprio (2017)

A rota atual percorre um total de 4600m em 15 min enquanto a rota otimizada percorre 4510m em 14,42 min obtendo assim uma redução de 90m e 0,58min por percurso. Esse trajeto é realizado 58 vezes por dia,

alcançando uma redução total de 5220m e 33,64 min por dia.

Para a linha 402.3-Santa Lúcia/DIA 03 da rota "IDA" obteve-se diagrama de redes exibido na Figura 7:

Figura 8 - Diagrama de redes da linha 402-3 da rota "IDA"



Fonte: Autor próprio (2017)

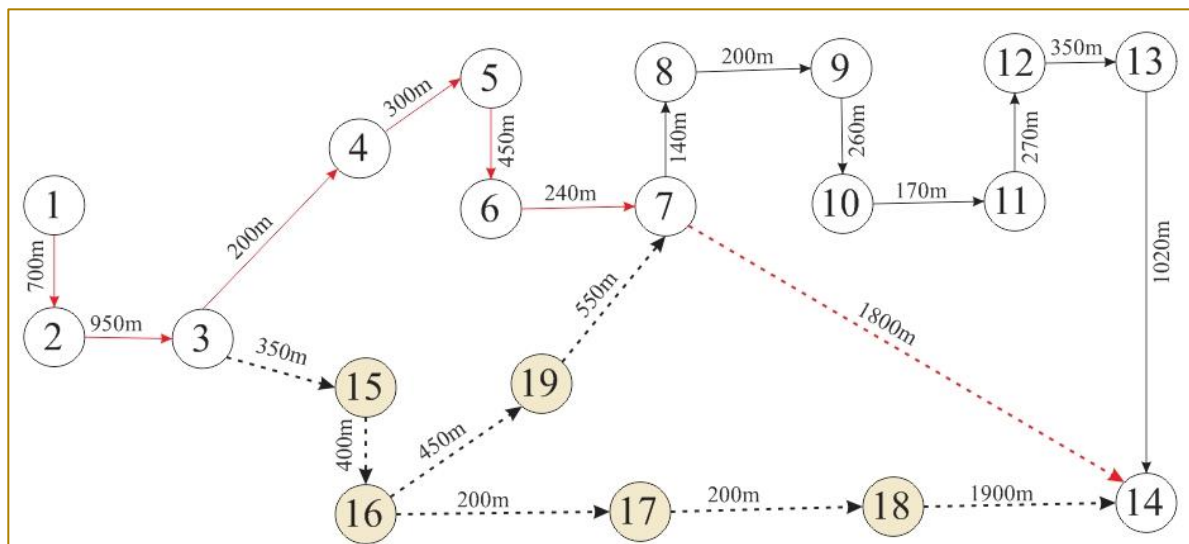
A rota atual percorre um total de 5420m em 20 min enquanto a rota otimizada percorre

4250m em 19,22 min obtendo assim uma redução de 170m e 0,78min por percurso.

Esse trajeto é realizado 18 vezes por dia, alcançando uma redução total de 3060m e 14,04 min por dia.

Para a linha 402.3-Santa Lúcia/DIA 03 da rota "VOLTA" obteve-se diagrama de redes exibido na Figura 8:

Figura 8 - Diagrama de redes da linha 402-3 da rota "VOLTA"



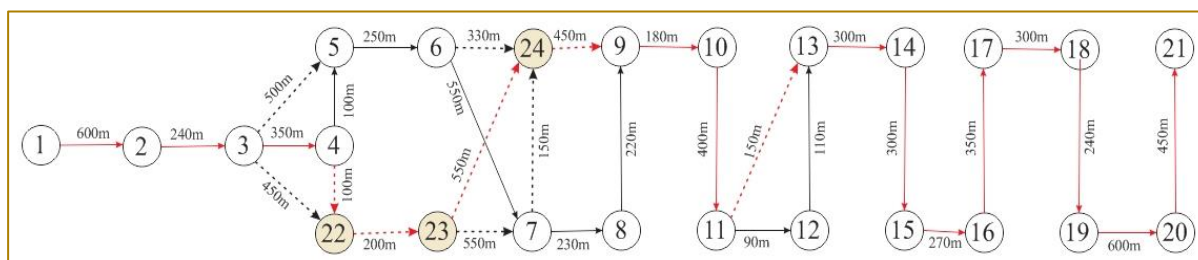
Fonte: Autor próprio (2017)

A rota atual percorre um total de 5250m em 20 min enquanto a rota otimizada percorre 4640m em 17,40 min obtendo assim uma redução de 610m e 2,6min por percurso. Esse trajeto é realizado 18 vezes por dia,

alcançando uma redução total de 10980m e 46,8 min por dia.

Para a linha 706-Santa Lúcia/Centro da rota "IDA" obteve-se diagrama de redes exibido na Figura 9:

Figura 9 - Diagrama de redes da linha 706 da rota "IDA"



Fonte: Autor próprio (2017)

A rota atual percorre um total de 6130m em 20 min enquanto a rota otimizada percorre 6030m em 19,40 min obtendo assim uma redução de 100m e 0,6min por percurso. Esse trajeto é realizado 16 vezes por dia, alcançando uma redução total de 1600m e 9,6min por dia.

maior quantidade de trajetos realizados por dia, promovendo, assim, uma redução do tempo de espera do passageiro no ponto de ônibus, como também, a redução do tempo do passageiro dentro do ônibus.

Ao observar o melhor trajeto a ser percorrido pelo ônibus notou-se que houve uma redução de quilômetros percorridos e, conseqüentemente, uma redução no tempo de trajeto. Essa quantidade de metros reduzidos pode ser convertida para uma



## 5. CONCLUSÃO

O artigo demonstra que o caminho mínimo é um método capaz de proporcionar melhorias em diversos setores, inclusive no transporte. Além disso, o Solver® também confirma ser um programa eficiente para problemas relacionados ao método do caminho mínimo. A existência de pontos de ônibus criados sem estudo adequado por meio dos órgãos públicos é a principal causa de trajetos mais longos encontrados em determinadas linhas.

As novas rotas não podem ser colocadas em prática devido à legislação, que conforme a lei 12.587, que institui as diretrizes da política nacional de mobilidade urbana, ou seja, uma série de procedimentos é necessária para a

aplicação das rotas do bairro Jabotiana, além da aprovação dos poderes citados acima.

A existência de rotas com pouca ou até sem alterações é normal, pois alguns trajetos foram elaborados de acordo com a necessidade dos passageiros ao longo dos anos, tendo atingido a sua otimalidade naturalmente.

Portanto, ao somar as reduções de percurso e de tempos, por dia, de todas as linhas estudadas, 41 km e 202,37min respectivamente, nota-se que foi atingido o objetivo esperado, pois a diminuição do percurso e do tempo de espera do passageiro reflete diretamente na qualidade de vida dos cidadãos.

## REFERÊNCIAS

- [1]. AGARANA, Michael C.; ADELEKE, Olawale J.; OWOLOKO, Enahoro A. Optimizing Public Transport Systems in Sub-Saharan Africa Using Operational Research Technique: A Focus on Nigeria. *Procedia Manufacturing*, África do Sul, 7.ed, p. 590-595, Jan. 2017.
- [2]. BRASIL. Lei nº 12.587, de 03 de janeiro de 2012.
- [3]. BODANESE, Ronaldo E. ; OLIVEIRA, José A.; SCALABRIN, Idionir.; MORES, Claudionor J. Teoria das restrições, pesquisa operacional e programação linear, estudo de caso com utilização do solver. IX Congresso Internacional de Custos, Florianópolis, Nov. 2005.
- [4]. CENSO DEMOGRÁFICO 2010. Sinopse por setores: IBGE, 2011. Disponível em: <<http://www.censo2010.ibge.gov.br/sinopseporsetores/?nivel=st>>. Acesso em: 12 abr. 2017.
- [5]. GOOGLE. Google Earth. Versão: 2017. Bairro Jabotiana. Disponível em: <<https://www.google.com.br/maps/search/bairro+jabotiana/@-10.9410718,-37.0901767,17z/data=!3m1!4b1>>. Acesso em: 15 abr. 2017.
- [6]. HILLIER, Frederick S.; LIEBERMAN, Gerald J. *Introdução à Pesquisa Operacional*. [13]. Tradução Ariovaldo Griesi; Revisão técnica Pierre J. Ehrlich. 9.ed. Porto Alegre: AMGH, p.8-10; p.25-30; p.360-368, 2013.
- [7]. LACHTERMARCHER, Gerson. *Pesquisa Operacional na tomada de decisões*. 4.ed..São Paulo: Pearson Prentice Hall, p.142, 2009.
- [8]. MÉNDEZ, Yasmin S.; GUARDIA, Luis E.T. *Problema do caminho mais curto: Algoritmo de Dijkstra*. SPOLM, Rio de Janeiro, Ago. 2008.
- [9]. NOSSA, Valcemiro; CHAGAS, José Ferreira. Usando programação linear na contabilidade decisória. *Revista Brasileira de Contabilidade (RBC)*, nº 107, Vitória, Set/Out. 1997.
- [10]. PRADO, Darci. *Programação Linear*. Belo Horizonte: Falconi, p.15,1998.
- [11]. SUPERINTENDÊNCIA MUNICIPAL DE TRANSPORTES E TRÂNSITO DE ARACAJU, Itinerário e Horário dos Ônibus. Disponível em: <<http://www.smttaju.com.br/smtt/transporte/itinerario-e-horario-dos-onibus>>. Acesso em: 15 abr. 2017.
- [12]. WINSTON, Wayne L.; GOLDBERG, Jeffrey B. *Operations research: applications and algorithms*. 4ed. Boston: Duxbury press, p. 413-414, 2004.

# Capítulo 19

## OTIMIZAÇÃO DE ESCALAS DE TRIPULAÇÃO DE TRANSPORTE PÚBLICO URBANO COM A META HEURÍSTICA VARIABLE NEIGHBORHOOD SEARCH

*Eduardo Luis Festa*

*Christiane Wenck Nogueira Fernandes*

*Silvia Lopes de Sena Tagliarenha*

**Resumo:** O presente trabalho aborda o Problema de Programação de Tripulação de ônibus urbano (PPT) o qual consiste em determinar jornadas diárias de trabalho para tripulantes que conduzirão uma frota em operação com o menor custo possível e atendendo restrições operacionais e trabalhistas. Devido à complexidade computacional para se obter uma solução utilizando métodos exatos, neste trabalho considera-se um método pertencente a uma classe alternativa de métodos de solução, denominada meta heurística, mais especificamente, a meta heurística busca em vizinhança variável (variable neighborhood search – VNS), que para escapar de ótimos locais realiza trocas sistemáticas de estruturas de vizinhança no espaço de busca. A implementação proposta é testada com os dados reais de uma empresa de médio porte e os resultados são comparados com o cenário atual da empresa.

**Palavras-chave:** Crew scheduling, Programação de Tripulação, Variable Neighborhood Search

\* Artigo originalmente apresentado no Encontro Nacional de Engenharia de Produção (ENEGEP, 2017)

## 1. INTRODUÇÃO

Em muitas cidades o ônibus é o principal, senão o único meio de transporte público de passageiros, portanto, é necessário que as empresas do ramo tenham suas atividades muito bem planejadas para atender à demanda dos usuários do sistema de maneira econômica. Para tanto, devem ser utilizadas ferramentas computacionais que auxiliam o processo de decisão. Nesse contexto é que se inserem os métodos de otimização, os quais podem levar à redução dos custos. Entre os custos de maiores pesos se destacam os operacionais, com os veículos, motoristas e cobradores (BOUZADA, 2003).

Segundo Silva e Reis (2014), o planejamento de um sistema de transporte público é decomposto em quatro etapas a fim de tornar o problema tratável do ponto de vista computacional.

A etapa inicial é constituída pela elaboração das rotas e dos quadros de horários, de maneira que todas as regiões da cidade sejam atendidas e que o tempo de espera nos pontos e terminais atenda a um dado nível de qualidade estipulado no processo licitatório. Na segunda etapa trata do problema da alocação de veículos que realizarão as viagens das tabelas de horários. Nessa etapa é determinada a quantidade necessária de veículos para atender as viagens de cada rota, assim como especificar as viagens a serem realizadas por cada veículo durante a operação. A terceira etapa compreende o problema da alocação das tripulações, composta pelo motorista e o cobrador, tomando como base as viagens alocadas a cada veículo da frota em operação. Nessa etapa deve ser determinada a quantidade necessária de tripulantes para conduzir os veículos e quais as viagens que cada tripulação realizará. Dessa maneira são definidas as jornadas diárias de trabalho que deverão ser executadas pelas tripulações da empresa.

Na etapa final do planejamento da operação do sistema diz respeito ao problema do rodízio das tripulações é conhecida na literatura especializada como Problema da Programação de Tripulações (PPT) (crew scheduling problem), são geradas as escalas mensais das tripulações com o menor custo possível de maneira a cobrir todas as jornadas diárias, respeitando as regras trabalhistas para o período. Essas etapas geralmente são realizadas na ordem que

foram expostas e a solução do problema de uma etapa é usada como entrada da etapa subsequente.

Para definir as jornadas diárias, devem ser respeitadas todas as leis trabalhistas e regras operacionais impostas pela empresa. Entre essas, destacam-se a carga horária máxima de trabalho, o tempo de descanso entre as jornadas, o número máximo de tripulações que podem trabalhar em um determinado tipo de jornada, entre outras.

Essas jornadas existem para atender as demandas mais acentuadas que ocorrem especificamente nos horários de pico. Durante uma jornada de trabalho, os tripulantes podem realizar uma troca de veículos, que pode ser uma forma de reduzir o tempo ocioso das tripulações. A empresa restringe o número de trocas realizadas pelas tripulações normalmente para facilitar o controle sobre o desgaste dos veículos em operação. Uma jornada deve ser definida de tal forma que o tempo de descanso entre o seu final e seu início no dia seguinte seja maior ou igual ao tempo mínimo de descanso estipulado pela legislação. Devido às restrições impostas, o problema é classificado como um problema NP-Difícil, ou seja, não existem algoritmos polinomiais que obtenham as soluções ótimas em tempo polinomial da entrada de dados (FISCHETTI et al., 1987). Portanto, é necessária a utilização de métodos heurísticos para resolver de forma satisfatória problemas de grande porte.

O objetivo principal deste trabalho é aplicar um método heurístico para resolver o problema de programação de escalas de motoristas de ônibus urbano em uma empresa real em uma cidade no Norte Catarinense.

## 2. REFERENCIAL TEÓRICO

O principal objetivo do PPT é atribuir o conjunto de tarefas requeridas aos tripulantes para operação dos veículos responsáveis pelo transporte público, respeitando as regulamentações trabalhistas, acordos sindicais, regras de segurança e políticas das empresas, de tal forma que a minimização do custo total da tripulação seja combinada com a satisfação dos tripulantes e dos usuários.

Vários métodos de resolução do PPT são demonstrados na literatura por diferentes autores, com resultados comprovadamente satisfatórios em relação à redução de custos operacionais quando métodos heurísticos são

aplicados em casos reais (SILVA; GUALDA, 2000; SILVA; GUALDA, 2001; WREN, 2004; SMITH; WREN, 1988). As abordagens mais utilizadas para lidar com este modelo são problemas de conjunto de cobertura ou de particionamento.

A estratégia de geração de colunas é em grande parte utilizada para resolver o problema, como pode ser visto nas obras de Fores et al. (1999), Barnhart (1998), Desrochers e Soumis (1989), Smith e Wren (1988), bem como as técnicas Branch-and-Bound (FORES et al., 1999; SMITH; WREN, 1988), o Branch-and-Cut (FRIBERG; HAASE, 1999) e o Branch-and-Price (BARNHART et al., 1998; DESROCHERS; SOUMIS, 1989). No entanto, modelos exatos são limitados em aplicações práticas, uma vez que são incapazes de resolver problemas muito grandes. Assim, é necessário utilizar métodos heurísticos para resolver problemas que aparecem na vida real, que são amplos e demandam alto custo computacional, obtendo-se resultados muito interessantes, embora a otimalidade das soluções não seja garantida.

Um dos grupos pioneiros nesta área, chamado de Scheduling and Constraint Management Group of Leeds University, realizou um conjunto de implementações heurísticas usando Algoritmos Genéticos (LI; KWAN, 2003) e Busca Tabu (SHEN; KWAN, 2001). Os modelos desenvolvidos por este grupo são extremamente utilizados no Reino Unido para alocar escalas de tripulação, bem como a programação operacional da frota (WREN, 2004).

Embora o PPT tenha sido amplamente estudado e aplicado em mais países na Europa e na América do Norte, as suas técnicas de resolução não são divulgadas e raramente são aplicadas à realidade brasileira. Parcialmente, isso acontece porque as empresas não possuem os dados necessários de entrada, e também porque faltam modelos e sistemas comerciais que representem a realidade operacional brasileira.

Entre os estudos sobre o PPT aplicados na realidade brasileira, podemos destacar modelos meta heurísticos que usam Simulated Annealing, Busca Tabu, Greedy Randomized Adaptive Search Procedure (GRASP) e Variable Neighborhood Search (VNS) (REIS; SILVA, 2011; SOUZA et al., 2004). Esses modelos têm sido testados com dados de

empresas que operam no sistema de transporte público do Brasil e os resultados mostram que há grandes possibilidades de reduzir custos em relação às soluções adotadas pelas empresas.

Outros modos de transportes apresentam problemas relacionados à alocação de funcionários, como consta na obra de Ernst et al. (2004), onde é mencionado que a complexidade dos problemas depende do tipo de transporte, da categoria da tripulação, dos tipos de frota, das regras e regulamentações, da regularidade das viagens e dos custos a serem considerados. Ainda segundo o autor, uma das abordagens mais utilizadas para resolver o PPT decompõe-se nas seguintes etapas: geração das jornadas, otimização das jornadas e definição do rodízio da tripulação. Assim, a complexidade do problema é reduzida, viabilizando de uma melhor forma a sua solução.

Um caso interessante do PPT é considerado por Fischetti (2001), onde o problema é modelado com programação inteira e não como um problema de recobrimento. A modelagem inteira resultante tem uma relaxação linear muito livre, mas pode ser apertada por meio de novos cortes, baseada em um algoritmo branch-and-cut. Testes numéricos mostram como o método é competitivo em alguns tipos de problemas.

Em Silva e Cunha (2010) o PPT é resolvido pela combinação da meta heurística GRASP (Greedy Randomized Adaptive Search Procedure) com o método de busca em vizinhança de grande porte, o VLNS (Very Large-Scale Neighborhood Search).

### 3. MÉTODO PROPOSTO

A Busca em Vizinhança Variável, VNS (Variable Neighborhood Search), é uma meta heurística que foi proposta em meados da década de 90, por (MLADENOVIC, 1995; MLADENOVIC; HANSEN, 1997). É baseada num princípio simples: explorar o espaço de soluções através de trocas sistemáticas de estruturas de vizinhança durante o processo de busca (HANSEN; MLADENOVIC, 1997; HANSEN; MLADENOVIC, 2001c). Ou seja, a meta heurística de busca em vizinhança variável é uma extensão de um algoritmo de busca local que utiliza a estratégia de mudança de tamanho da vizinhança para sair de soluções ótimas locais. Seu

desenvolvimento foi rápido, com inúmeras publicações (HANSEN; MLADENOVIC; PÉREZ, 2001; HANSEN; MLADENOVIC, 2001a; HANSEN; MLADENOVIC, 2001b; HANSEN; MLADENOVIC, 2002; HANSEN; MLADENOVIC, 2003a; HANSEN; MLADENOVIC, 2003b).

A troca de estruturas representa a substituição do componente analisado por outro mais eficiente encontrado pelo processo de busca VNS ou VND (Método de Descida em Vizinhança Variável). A essência do método VND consiste em buscar a melhora

da solução atual utilizando uma estrutura de vizinhança principal. O método troca a vizinhança corrente pela próxima quando a melhora não é mais viável e quando uma melhor solução é encontrada ele retorna para a vizinhança principal. O fim do algoritmo é dado quando não há mais a possibilidade de melhorar a solução atual após o uso de todas as estruturas de vizinhança.

O VNS, ao partir de uma solução inicial, realiza busca local em sua vizinhança igualmente ao VND. Na figura 1 apresenta-se o pseudocódigo do algoritmo VNS.

Figura 1: Procedimento VNS

```

Procedimento VNS
  S := S0;
  K := 1;
  Nk := 1;
  While (critério de parada) do
  Begin
    While (K < Kmax) do
    begin
      Aplicar_busca_local S' := Nk(S);
      If f(S') < f(S) then
      begin
        S := S';
        K := 1;
      end
      else
        k := k + 1;
      End;
    End;
  End;
End;

```

Fonte: Elaborada pelo autor e baseada em (MLADENOVIC; HANSEN, 1997)

Nota-se que com a inicialização de uma solução inicial  $S_0$  e após cada iteração é selecionado um vizinho  $S_1$  da vizinhança  $N$ . A busca local é efetuada pesquisando seus vizinhos e visando alcançar uma solução ótima local  $S_2$ . Se  $S_2$  for melhor que a solução corrente, então  $S_2$  passa a ser a nova melhor solução  $S$  e a busca é reiniciada. Se a solução alcançada não for melhor do que a corrente, o sistema percorrerá o caminho de busca através da próxima estrutura de vizinhança. Quando o critério de parada for

atingido, bem como o número máximo de iterações possíveis, a rotina é encerrada.

A VNS, ao contrário de muitas meta heurísticas baseadas em busca local, não permite degradação na função objetivo para realizar um movimento, mas explora, a partir de uma solução inicial, uma sequência crescente de vizinhanças, e só realiza o

### 3.1 A EMPRESA EM ESTUDO

Nesta seção descreve-se a empresa em estudo e apontam-se suas principais características em relação à metodologia empregada atualmente na escala dos motoristas, bem como as consequências geradas, parâmetros de qualidade e valor da tarifa.

A empresa em estudo possui cerca de 850 colaboradores, frota de 260 veículos que percorrem aproximadamente 1.300.000 km/mês e mais de 4500 horários de ônibus distribuídos em 114 itinerários e os dados utilizados são referentes à movimentação e um único terminal.

O planejamento das escalas é elaborado respeitando normas que favorecem a longevidade da frota, que passam por revisão a cada 20.000km rodados. Todos os ônibus rodam juntos, em média, 43.333km por dia.

Atualmente, as escalas são realizadas por profissionais que estão empregados na empresa há bastante tempo, tendo assim um suficiente conhecimento da maior parte dos motoristas e da frota, a ponto de elaborar os projetos dos mesmos manualmente. Essa metodologia, quando aplicada em uma larga escala de dados, pode se tornar imprecisa, gerando uma solução muito distante da ótima.

Dessa forma, o tempo ocioso entre tarefas por motorista se torna crescente, gerando gastos excessivos para a empresa, o que mostra que manter um profissional interno especializado nesse tipo de planejamento pode ser interessante.

### 3.2 COLETA E TRATAMENTO DE DADOS

A empresa em estudo foi procurada e em reunião com os responsáveis pela sua operação, foi explicado o interesse em estudar com esse problema. Solícita, foram cedidas todas as informações necessárias para o desenvolvimento deste trabalho, ou seja, um banco de dados (no formato pdf) de um planejamento diário contendo informações relacionadas aos motoristas e outro contendo informações relacionadas aos veículos, ambos referentes ao Terminal Nova Brasília, na cidade de Joinville, SC.

No primeiro banco de dados constavam os 27 projetos elaborados, para cada um dos 27 motoristas, denominado "Projeto Escala". No segundo, por sua vez, continha o serviço alocado para cada um dos 13 veículos em seus respectivos projetos, chamados de "Projeto Veículo". Nas figuras 2 e 3 apresenta-se um exemplo de Projeto Escala e Projeto Veículo, respectivamente, juntamente com suas legendas.

Figura 2: Projeto Escala

| Projeto | Motorista | Veículo | Escala | Tarefa | Hora Saída | Hora Chegada | Tempo Ocioso | Ponto Inicial    | Ponto Final      | Linha Inicial | Linha Final |
|---------|-----------|---------|--------|--------|------------|--------------|--------------|------------------|------------------|---------------|-------------|
| 0650-2  | 5139      | 11333   | 8001   | 1      | 03:55      | 04:08        | 00:00        | A                | B                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 2      | 04:09      | 04:22        | 00:01        | B                | A                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 3      | 04:22      | 04:33        | 00:00        | A                | B                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 4      | 04:33      | 04:40        | 00:00        | B                | C                | 1507          | 1507        |
|         | 5139      | 11333   | 8001   | 5      | 04:40      | 05:07        | 00:00        | C                | B                | 1502          | 1502        |
|         | 5139      | 11333   | 8001   | 6      | 05:15      | 05:29        | 00:08        | B                | A                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 7      | 05:29      | 05:42        | 00:00        | A                | B                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 8      | 05:42      | 06:02        | 00:00        | B                | Terminal Central | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 9      | 06:02      | 06:22        | 00:00        | Terminal Central | B                | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 10     | 06:30      | 06:44        | 00:08        | B                | A                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 11     | 06:44      | 07:00        | 00:00        | A                | B                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 12     | 07:00      | 07:28        | 00:00        | B                | Terminal Central | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 13     | 07:28      | 07:42        | 00:00        | Terminal Central | B                | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 14     | 07:42      | 08:02        | 00:00        | B                | Terminal Central | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 15     | 08:04      | 08:22        | 00:02        | Terminal Central | B                | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 16     | 10:07      | 10:21        | 01:45        | B                | A                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 17     | 10:21      | 10:35        | 00:00        | A                | B                | 1505          | 1505        |
|         | 5139      | 11333   | 8001   | 18     | 10:42      | 11:01        | 00:07        | B                | Terminal Central | 0650          | 0650        |
|         | 5139      | 11333   | 8001   | 19     | 11:01      | 11:21        | 00:00        | Terminal Central | B                | 0650          | 0650        |

Fonte: Elaborado pelo autor

Projeto: identificação, conforme organização da empresa;

Motorista: identificador do motorista;

Veículo: identificador do veículo;

Escala: identificador da escala do respectivo motorista;

Tarefa: Organizadas em ordem cardinal;

Hora saída: Horário do início da tarefa;

Hora chegada: Horário do fim da tarefa;

Tempo ocioso: Diferença de tempo entre o fim da última tarefa e o começo da atual;

Ponto inicial: Local de saída da tarefa;

Ponto final: Local de chegada da tarefa;

Linha inicial: Identificador do caminho de ida percorrido pela tarefa;

Linha final: Identificador do caminho de volta percorrido pela tarefa;

Vinculado: Número identificador do conjunto de tarefas do respectivo veículo;

Tempo viagem: Duração da viagem;

Tempo parado: Diferença de tempo entre a Hora Chegada da última tarefa e a Hora Saída da atual.

Figura 3: Projeto Veículo

| Veículo | Vinculado | Linha Inicial | Linha Final | Tarefa | Hora Saída | Hora Chegada | Tempo Viagem | Tempo Parado | Ponto Inicial    | Ponto Final      |
|---------|-----------|---------------|-------------|--------|------------|--------------|--------------|--------------|------------------|------------------|
| 0651-4  | 11202     | 1505          | 1505        | 1      | 06:22      | 06:36        | 00:14        | 00:00        | B                | A                |
| 0651-4  | 11202     | 1505          | 1505        | 2      | 06:36      | 06:50        | 00:14        | 00:00        | A                | B                |
| 0651-4  | 11202     | 0651          | 0651        | 3      | 06:53      | 07:13        | 00:20        | 00:03        | B                | Terminal Central |
| 0651-4  | 11202     | 0651          | 0651        | 4      | 07:13      | 07:34        | 00:21        | 00:00        | Terminal Central | B                |
| 0651-4  | 11202     | 0651          | 0651        | 5      | 07:41      | 08:00        | 00:19        | 00:07        | B                | Terminal Central |
| 0651-4  | 11202     | 0651          | 0651        | 6      | 08:05      | 08:24        | 00:19        | 00:05        | Terminal Central | B                |
| 0651-4  | 11202     | 1505          | 1505        | 7      | 12:25      | 12:39        | 00:14        | 04:01        | B                | A                |
| 0651-4  | 11202     | 1505          | 1505        | 8      | 12:39      | 12:53        | 00:14        | 00:00        | A                | B                |
| 0651-4  | 11202     | 0650          | 0650        | 9      | 13:39      | 14:02        | 00:23        | 00:46        | B                | Terminal Central |
| 0651-4  | 11202     | 0650          | 0650        | 10     | 14:04      | 14:26        | 00:22        | 00:02        | Terminal Central | B                |
| 0651-4  | 11202     | 0650          | 0650        | 11     | 16:12      | 16:32        | 00:20        | 01:46        | B                | Terminal Central |
| 0651-4  | 11202     | 0650          | 0650        | 12     | 16:38      | 17:00        | 00:22        | 00:06        | Terminal Central | B                |
| 0651-4  | 11202     | 0650          | 0650        | 13     | 17:12      | 17:33        | 00:21        | 00:12        | B                | Terminal Central |
| 0651-4  | 11202     | 0650          | 0650        | 14     | 17:39      | 18:02        | 00:23        | 00:06        | Terminal Central | B                |
| 0651-4  | 11202     | 0651          | 0651        | 15     | 18:06      | 18:22        | 00:16        | 00:04        | B                | Terminal Central |
| 0651-4  | 11202     | 0651          | 0651        | 16     | 18:33      | 18:52        | 00:19        | 00:11        | Terminal Central | B                |

Fonte: Elaborado pelo autor

### 3.3 FORMULAÇÃO MATEMÁTICA

Como o principal objetivo do problema é reduzir o tempo ocioso dos motoristas, onde se deve respeitar o momento de intervalo por dia de trabalho, tem-se a função objetivo na

$$\text{Soma\_Tempo\_Ocioso} = \sum_{i,j=2}^N (hi_i - hf_{j-1}) - 27 \quad (1)$$

O valor 27 na equação (1) consiste no tempo de intervalo de descanso/refeição, em horas,

em que  $hi_i$  é o horário de início (hora de saída), em horas, da tarefa  $i$ , e  $hf_{j-1}$  o horário final (horário de chegada), também em horas, da tarefa anterior.

que todos os motoristas tem direito. O somatório inicia em 2 pelo fato da primeira



tarefa não possuir tempo ocioso e percorre até a última tarefa  $N$ .

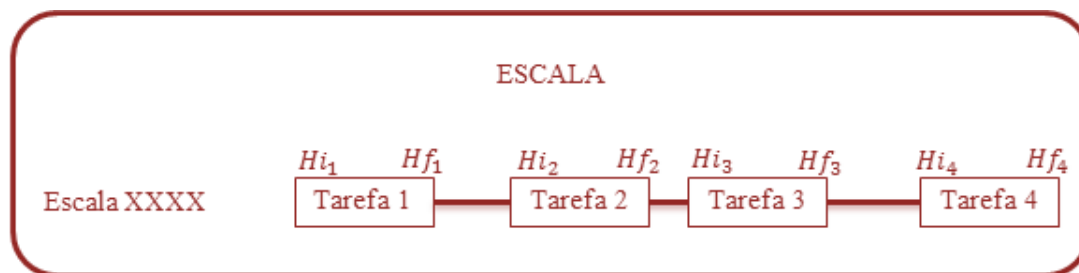
Os critérios e restrições estabelecidos para as trocas ocorrerem com factibilidade são: haver espaço no projeto para alocação de uma nova tarefa; a tarefa candidata a ser alocada para um novo projeto deve ter o ponto inicial no mesmo local onde o motorista do projeto candidato a recebê-la se encontra; respeito às leis trabalhistas: jornada máxima diária de 8 horas, com um acréscimo de no máximo 2 horas extras, sendo permitido semanalmente no máximo 44 horas de trabalho; 30 minutos de intervalo a cada 4 horas trabalhadas; 1 hora de intervalo a cada 8 horas trabalhadas; 11 horas de intervalo entre o fim de uma jornada e o início da próxima; o tempo ocioso

gerado com a tarefa realocada deve ser menor do que com ela no seu projeto de origem.

### 3.4 SOLUÇÃO INICIAL E REPRESENTAÇÃO DA SOLUÇÃO

Considerou-se como solução inicial a escala empregada atualmente pela empresa, ou seja, os projetos de escalas. Para facilitar o entendimento do método aplicado, uma escala será representada como na figura 10, em que XXXX representa o código de identificação da escala,  $H_{i_i}$  é o horário inicial e  $H_{f_i}$  o horário final da tarefa  $i$ . A linha contínua entre as tarefas representa a ociosidade entre elas.

Figura 3: Representação de uma escala



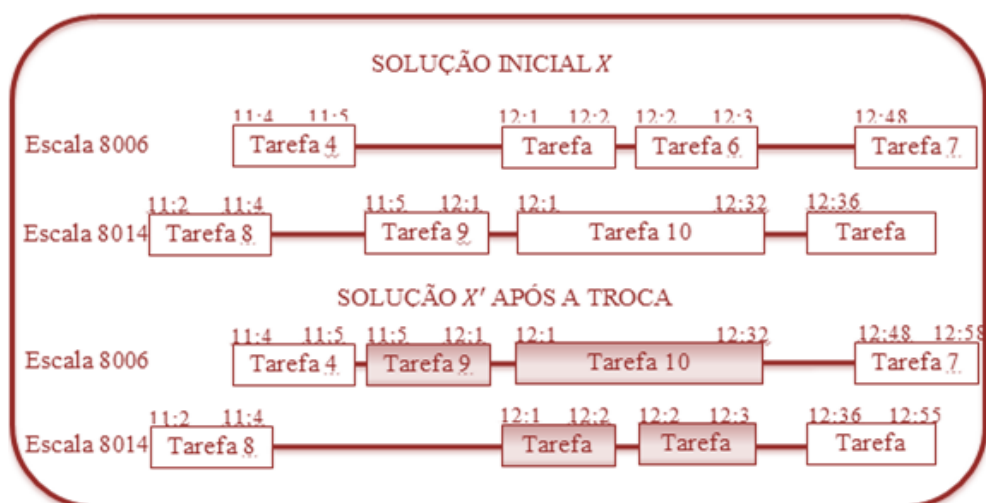
Fonte: Elaborado pelo autor

### 3.5 DEFINIÇÃO DAS ESTRUTURAS DE VIZINHANÇA

Considerou-se três estruturas de vizinhança definidas a seguir.

**Vizinhança  $N_1$ :** Dada a solução de escala  $X$ , uma solução  $X'$  é dita pertencer à vizinhança  $N_1$  se  $X'$  é melhorada partir de  $X$  realizando-se trocas de tarefas entre escalas (figura 4).

Figura 410: Operação de troca

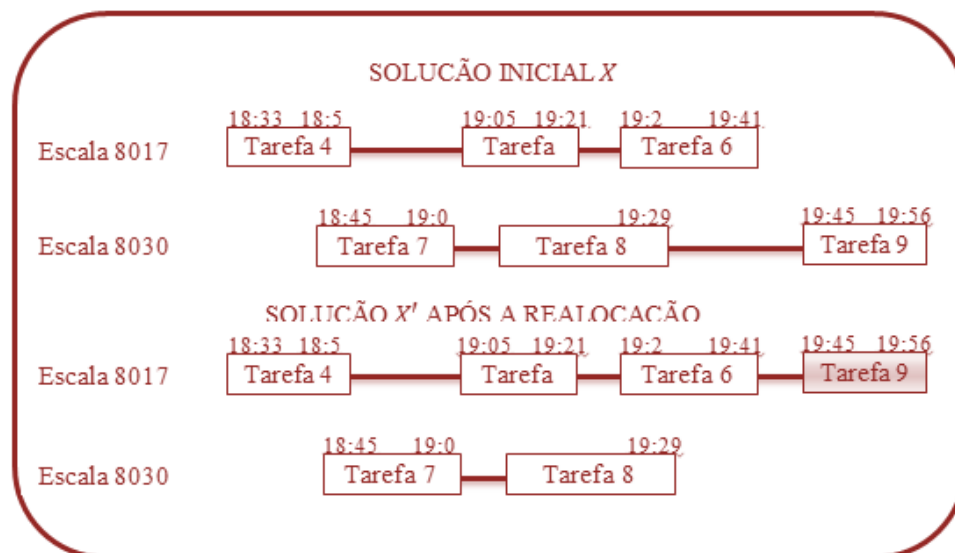


Fonte: Elaborado pelo autor

**Vizinhança  $N_2$ :** Uma solução  $X'$  pertence à vizinhança  $N_2$  se  $X'$  é obtida a partir de  $X$

realizando-se realocações de tarefas entre escalas (figura 5).

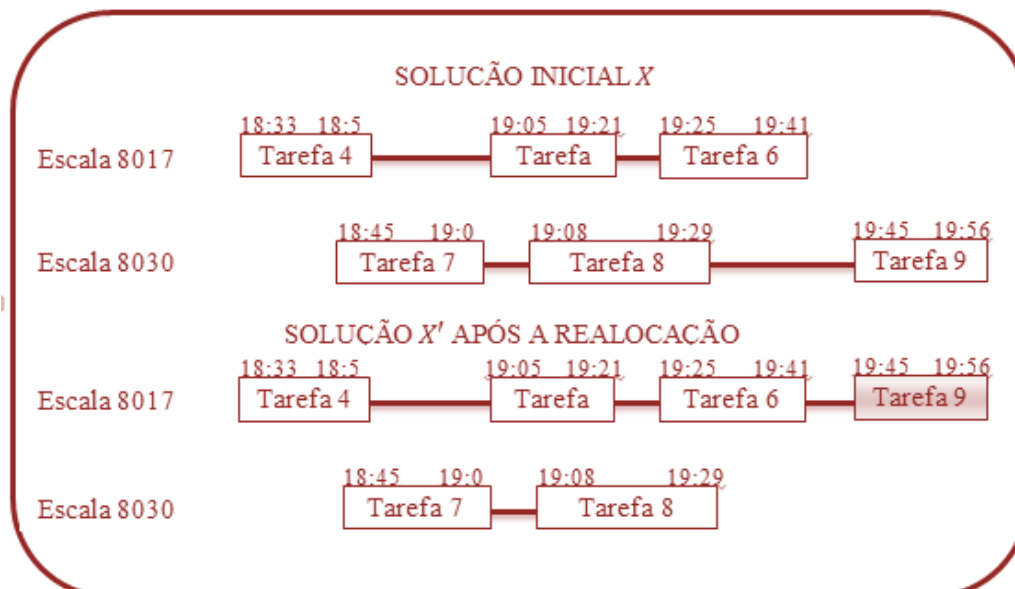
Figura 5: Operação de realocação



Fonte: Elaborado pelo autor

**Vizinhança  $N_3$ :** Da mesma forma, a solução  $X'$  pertence à vizinhança  $N_3$  se  $X'$  é obtida a partir de  $X$  realizando-se trocas e realocações de tarefas entre escalas(Figura 6).

Figura 6: Operação de realocação



Fonte: Elaborado pelo autor

#### 4. IMPLEMENTAÇÃO COMPUTACIONAL

O algoritmo foi desenvolvido na linguagem C++ através da ferramenta CodeBlocks. Para a busca de descida nas três estruturas de

vizinhança, foram desenvolvidas as rotinas RN1, RN2 e RN3, as quais realizam combinações dos movimentos definidos nas figuras 11, 12 e 13. A rotina RN1 executa trocas de acordo com os tempos ociosos

entre tarefas dos motoristas. A RN2 executa realocações de acordo com o tempo de intervalo excedente sobre o determinado pela lei e a RN3 troca e realoca tarefas de acordo com o espaço de tempo a ser preenchido até o limite de horas diárias trabalhadas determinado em lei. A lógica utilizada para a busca local nas rotinas é detalhada a seguir.

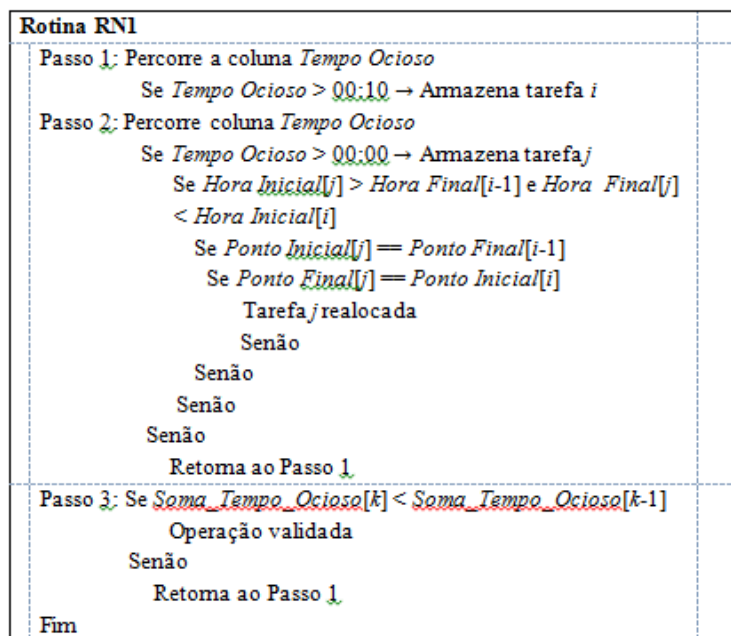
**Rotina RN1** - Considerando  $i$  a tarefa candidata a receber um vizinho,  $j$  a tarefa candidata a passar por uma troca na  $k$ -ésima iteração, realiza-se os passos na Figura 7.

Em RN1 nunca se altera o tempo total da jornada do motorista candidato a receber uma

nova tarefa, porém esse tempo pode se alterar se a jornada receber a primeira ou a última tarefa de um projeto. No entanto, esse tempo pode ser compensado em RN3, que realoca tarefas no intervalo do tempo máximo de jornada diária.

No Passo 1 foi considerado 00:10 como tempo ocioso mínimo pois segundo a empresa esse é o menor tempo de ida e volta de uma tarefa, e assim, o motorista candidato não compromete o cumprimento da sua próxima tarefa, caso assuma outra.

Figura 7: Rotina RN1

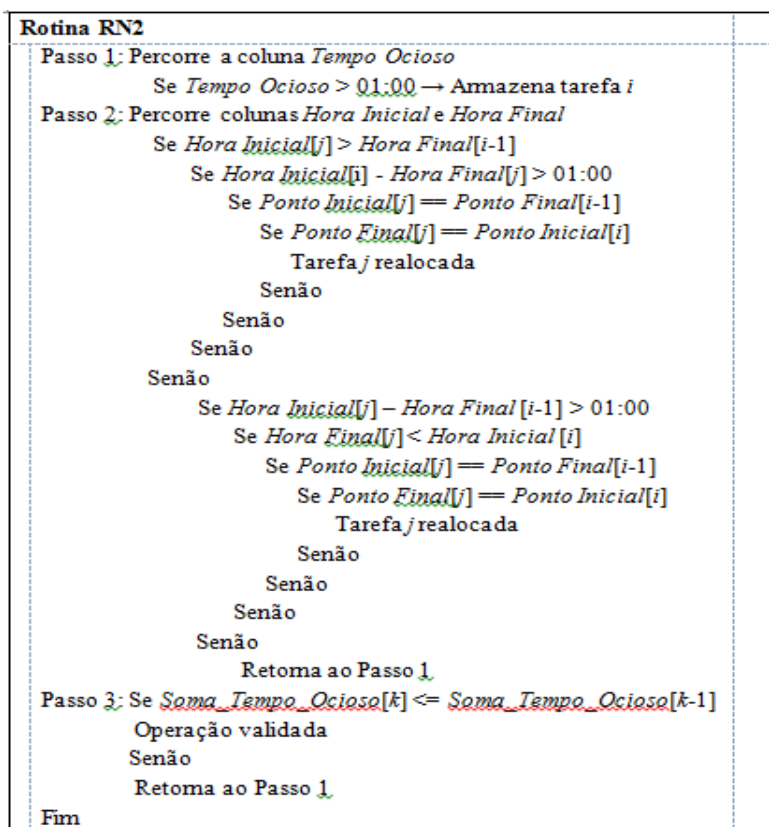


Fonte: Elaborado pelo autor

**Rotina RN2** - Considerando  $i$  a tarefa candidata a receber um vizinho,  $j$  a tarefa candidata a passar por uma realocação na  $k$ -

$ésima$  iteração, realiza-se os passos na figura 8.

Figura 8: Rotina RN2



Fonte: Elaborado pelo autor

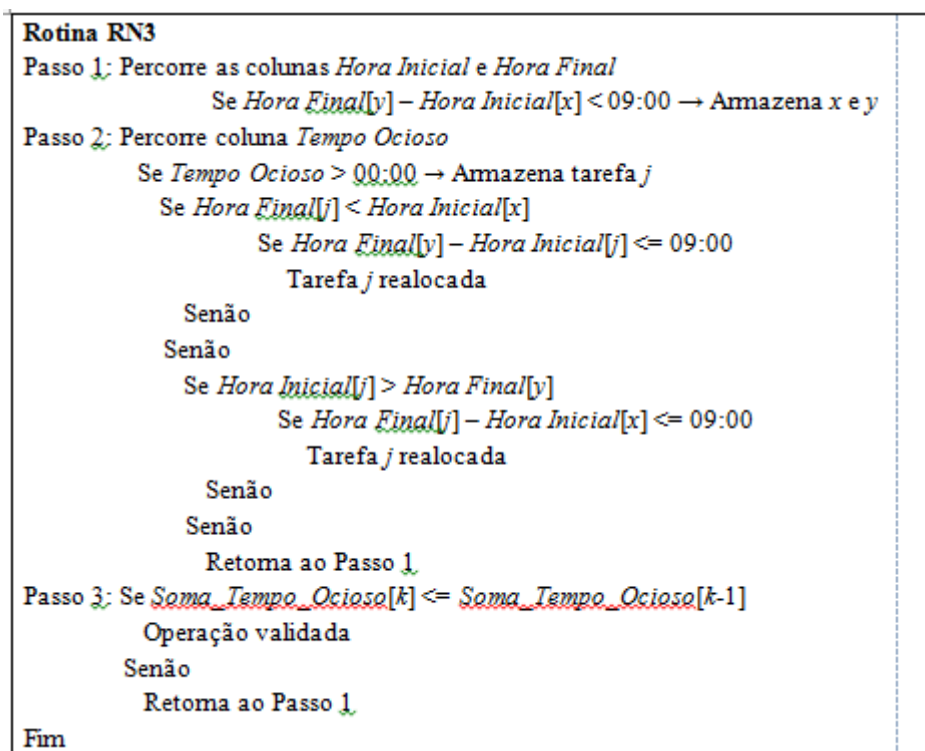
Na rotina RN2 a comparação da *Soma\_Tempo\_Ocioso* com a iteração anterior pode ser menor ou igual, pois sendo igual ainda assim pode-se deixar um espaço livre no projeto da tarefa que foi alocada a ser preenchido por eventuais operações que surgirem à frente, eliminando o tempo excedente de intervalo do projeto em questão.

**Rotina RN3** - Seja *i* a tarefa candidata a receber um vizinho, *j* a tarefa candidata a passar por uma troca/realocação na *k*-ésima

iteração. Considerando um projeto cuja primeira tarefa é *x* e a última tarefa é *y*, realizam-se os passos na figura 16.

Foi considerado um tempo máximo de 09:00 horas para o Projeto Escala devido ao fato de 08:00 horas serem determinadas pela legislação, mais 01:00 hora extra. Mesmo a legislação permitindo 02:00 horas extras, foi considerado apenas 01:00 hora, pois está dentro do que é praticado pela empresa e é razoável num eventual cálculo dentro do horizonte semanal.

Figura 9: Rotina RN3



Fonte: Elaborado pelo autor

## 5 RESULTADOS OBTIDOS

O tempo ocioso total de todos os 27 projetos antes da aplicação do método era de 53 horas e 45 minutos no planejamento diário de escala. Após a execução do método heurístico, o tempo ocioso foi reduzido para 32 horas e 40 minutos. Entretanto, três projetos foram modificados a ponto de reduzir o dia de trabalho do motorista para apenas meia jornada (4 horas de trabalho), alterando assim o seu tempo de intervalo para 30 minutos. Outro projeto foi alterado deixando o motorista com apenas duas tarefas e duração

total de trabalho de 1 hora e 31 minutos. Para esse caso específico, não foi considerado tempo de intervalo. Caso fosse possível realocar essas duas tarefas a outro(s) motorista(s), um motorista seria eliminado do quadro de escalas, gerando redução de custos para a empresa. Portanto, considerando a alteração em relação o tempo de intervalo total, que de 27 horas passou para 24 horas e 30 minutos, tem-se um novo tempo ocioso de 35 horas e 10 minutos.

## REFERÊNCIAS

- [1]. BARNHART, C. et al. Branch-and-price: column generation for solving huge integer programs. *Operations Research*. Atlanta, v. 46, p. 316-329. 1998.
- [2]. BOUZADA, C. F. *Custo do transporte coletivo por ônibus*. Belo Horizonte: C/Arte, 2003.
- [3]. CUNHA, C. B. *Uso da técnica de busca em vizinhança de grande porte para a programação de escala de motoristas de ônibus urbano*. Ouro Preto: UFOP, 2010.
- [4]. DESROCHERS, M.; SOUMIS, F. A Column Generation approach to the urban transit crew scheduling problem. Montreal: École des Hautes Études commerciales, 1989.
- [5]. ERNST, A. T.; JIANG, H.; KRISHNAMOORTHY, M.; SIER, D. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*. Melbourne, v. 153, p. 3-27. 2004
- [6]. FISCHETTI, M.; MARTELLO, S.; TOTH, P. The fixed job schedule problem with spread-time constraints. *Operations Research*. Pádua, v. 35, n. 6, p. 849-858. 1987.

- [7]. FISCHETTI, M. et al. A polyhedral approach to simplified crew scheduling and vehicle scheduling problems. *Management Science*. Pádua, v. 6, n. 47, p. 833-850. 2001.
- [8]. FORES, S.; PROLL, L.; WREN, A. An improved ILP system for driver scheduling. In: Wilson, N. H. M. (Ed.) *Computer-Aided Transit Scheduling*. p. 43-61. Montreal: Springer, 1999.
- [9]. FRIBERG, C.; HAASE, K. An exact branch and cut algorithm for the vehicle and crew scheduling problem. In: Wilson, N. H. M. (Ed.) *Computer-Aided Transit Scheduling*. p. 63-80. Montreal: Springer, 1999.
- [10]. HANSEN, P.; MLADENOVIC, N. Developments of variable neighborhood search. *Les Cahiers du GERAD*, Montreal, v. 22, n. G-01-24, 2001a.
- [11]. HANSEN, P.; MLADENOVIC, N. Industrial applications of the variable neighborhood search metaheuristics. In: ZACCOUR, G. (Ed.) *Decisions and Control in Management Science*. Boston: Kluwer, p. 261-274, 2001b.
- [12]. HANSEN, P.; MLADENOVIC, N. Variable neighborhood search: Principles and applications. *European Journal of Operational Research*, Amsterdam, n. 130, p. 449-467, 2001c.
- [13]. HANSEN, P.; MLADENOVIC, N.; PÉREZ, J. A. M. Variable neighborhood decomposition search. *Journal of Heuristics*, Boston, v. 4, n. 7, p. 335-350, 2001.
- [14]. HANSEN, P.; MLADENOVIC, N. Variable neighborhood search. In: PARDALOS, M. P.;
- [15]. RESENDE, M. G. C. (Ed.) *Handbook of Applied Optimization*. Boston: Kluwer, 2002.
- [16]. HANSEN, P.; MLADENOVIC, N. A tutorial on variable neighborhood search. *Les Cahiers du GERAD*, Montreal, v. 2003-46, n. G, 2003a.
- [17]. HANSEN, P.; MLADENOVIC, N. Variable neighborhood search. In: GLOVER, F.; KOCHENBERGER, G. A. (Ed.) *Handbook of Metaheuristics*. Dordrecht: Kluwer, p. 145-184, 2003b.
- [18]. LOURENÇO, H. R.; PAIXÃO, J. M. P.; PORTUGAL, R. Multiobjective metaheuristics for the bus driver scheduling problem. *Transportation Science*. Barcelona, v. 35, n. 3, p. 331-343. 2001.
- [19]. MLADENOVIC, N. A variable neighborhood algorithm - a new metaheuristic for combinatorial optimization. *Abstracts of Papers Presented at Optimization Days*, 12, Montreal, 1995.
- [20]. MLADENOVIC, N.; HANSEN, P. Variable neighborhood search. *Computers and Operations Research*. Belgrado, v. 24, p. 1097-1100, 1997.
- [21]. REIS, A. F. da S.; SILVA, G. P. Um estudo de diferentes métodos de busca e a meta heurística VNS para otimizar a escala de motoristas de ônibus urbano. In: *Transporte em Transformação XVI – Trabalhos Vencedores do Prêmio CNT – Produção Acadêmica 2011*. *Transporte em Transformação XVI. Anais...* p. 47-64. Ouro Preto: Positiva, 2012.
- [22]. SHEN, Y.; KWAN, R. S. K. Tabu search for driver scheduling. In: S. Voss e J. R. Daduna (Ed.), *Computer-Aided Scheduling of Public Transport*. Montreal: Springer, 2001. p. 121-135.
- [23]. SILVA, G. P. Uma metodologia baseada na técnica de geração de arcos para o problema de programação de veículos. São Paulo: Escola Politécnica USP, 2001.
- [24]. SILVA, G. P.; CUNHA, C. B. Uso da Técnica de Busca em Vizinhança de Grande Porte para a Programação da Escala de Motoristas de Ônibus Urbano. *Transportes*. Ouro Preto, v. 18, p. 64-75, 2010.
- [25]. SILVA, G. P.; GUALDA, N. D. F. O Método Arcgen para Programação de Veículos: Um Estudo de Caso da Cidade de Belo Horizonte. In: *Congresso de Pesquisa e Ensino em Transportes XV. Anais: Panorama Nacional da Pesquisa em Transportes*. v. 1, p. 129-137. Campinas: ANPET, 2001.
- [26]. SILVA, G. P.; GUALDA, N. D. F. Um Algoritmo de Geração de Arcos para o Problema de Programação de Veículos. *Transportes*. Rio de Janeiro, v. 8, n. 1, p. 35-55, 2000.
- [27]. SILVA, G. P.; REIS, A. F. da S. A study of different metaheuristics to solve the urban transit crew scheduling problem. *Journal of Transport Literature*. Ouro Preto, v. 8, n. 4, p. 227-251. 2014.
- [28]. SMITH, B. M.; WREN, A. A bus crew scheduling system using a set covering formulation. *Transportation Research*. Amsterdam, v. 22A, p. 97-108. 1988.
- [29]. WREN, A. Scheduling Vehicles and Their Drivers – Forty Years Experience. In: *9<sup>th</sup> International Conference. Computer-Aided Scheduling of Public Transport*. p. 27-40, San Diego: CASPT, 2004.
- [30]. WREN, A.; SMITH, B. M.; MILLER, A. J. *Complementary approaches to crew scheduling*. Amsterdam: Elsevier Science, 1988.

# Capítulo 20

## MODELOS MATEMÁTICOS PARA OTIMIZAÇÃO DO PROCESSO DE MONTAGEM DA CABINE DE PINTURA

*Aline Maki Igari*

*Diogo Kenji Nomura*

*Jessica Ingrid Sales*

*Michele Martins da Silva*

*Priscila Harumi Souza*

**Resumo:** O presente projeto aborda os conceitos de modelos matemáticos através da utilização do programa Microsoft excel e a sua aplicação na tomada de decisão. Um estudo quantitativo entre duas modelagens é realizado, na qual, ambas relacionam os dados dos 4 fornecedores, custos e tempo de fabricação das peças da cabine de pintura, sendo que, a primeira modelagem é de uma única linha de produção em série para cada fornecedor, e a segunda modelagem é de 3 linhas de produção em paralelo para cada fornecedor. Nesse estudo, é analisado e comparado os resultados obtidos de ambas as modelagens, a fim de otimizar o custo e tempo de fabricação das peças da cabine de pintura.

Palavras-chave: Modelagem. Linha de Produção. Estudo. Otimizar.

\* Artigo originalmente apresentado no Encontro Nacional de Engenharia de Produção (ENEGEP, 2017)

## 1. INTRODUÇÃO

Para este projeto, as disciplinas Gerenciamento de Projetos e Pesquisa Operacional I Modelos Determinísticos, são de extrema importância e essenciais na composição do mesmo. As junções das duas proporcionam o aprofundamento do estudo e base para uma análise precisa.

Em vista do mercado atual brasileiro, as empresas e indústrias reforçam a diminuição dos seus custos e tempo de fabricação. A ação fazer mais com menos, e duas vezes mais, é totalmente indispensável. Porém, na prática é outra coisa, colocar em trilhos um plano de gerenciamento para otimizar os processos de uma empresa e/ou indústria, reduzir custos, tempo e maximizar os lucros é um grande desafio. No entanto, mesmo com o cenário atual do país, o crescimento não para e independente dessa instabilidade a produção continua.

## 2. OBJETIVO

O enfoque do presente projeto é analisar os custos e tempos de fabricação das peças que compõem a cabine de pintura, através de modelos matemáticos, com o propósito de obter uma tomada de decisão para otimização do processo.

## 3. REFERENCIAL TEÓRICO

### 3.1. MANAGEMENT SCIENCES

Denominamos Management Sciences (MS) a área de estudos que utiliza computadores, estatística e matemática para resolver problemas de negócios. Ela pode ser considerada uma subárea da pesquisa operacional (PO), por se tratar de uma modelagem matemática aplicada à área de negócios (LACHTERMACHER; GERSON, 2009).

Lachtermacher (2009) diz que: Em relação aos tipos de problemas que a MS-PO (Ciência da Gestão na Pesquisa Operacional) pode ser utilizada para ajudar no processo de decisão encontram-se:

- Problemas de otimização de recursos;
- Problemas de localização;
- Problemas de roteirização;
- Problemas de carteiras de investimento;

- Problemas de alocação de pessoas;
- Problemas de previsão e planejamento;
- Problemas de alocação de verbas de mídia.

### 3.2. O PROCESSO DE MODELAGEM

Quando os gerentes se veem diante de uma situação na qual uma decisão deve ser tomada entre uma série de alternativas conflitantes e concorrentes, duas opções básicas se apresentam: (1) usar apenas a intuição gerencial; e (2) realizar um processo de modelagem da situação e exaustivas simulações dos mais diversos cenários de maneira a estudar mais profundamente o problema. Até recentemente, a primeira opção se constituía na única alternativa viável, visto que não existiam dados nem informações sobre os problemas, nem poder computacional para resolvê-los. Com o advento dos microcomputadores e com o aprimoramento da tecnologia de bancos de dados, ela deixou de ser a única opção para os tomadores de decisão. Um número cada vez maior de empresas e tomadores de decisão passou a optar pela segunda alternativa, isto é, pela elaboração de modelos para auxiliar esse processo (LACHTERMACHER; GERSON, 2009).

### 3.3. TIPOS DE MODELOS

Basicamente, podem-se ter três tipos de modelos: os modelos físicos, análogos e matemáticos ou simbólicos. Os modelos mais utilizados na modelagem de situação gerenciais são os modelos simbólicos ou matemáticos, em que as grandezas são representadas por variáveis de decisão, e as relações entre essas variáveis, por expressões matemáticas. Por tais características, esses modelos necessitam de informações quantificáveis (LACHTERMACHER; GERSON, 2009).

Para Lachtermacher (2009): Um modelo simbólico contém um conjunto suficiente de detalhes de maneira que:

Os resultados atinjam suas necessidades;

O modelo seja consistente com os dados;

O modelo possa ser analisado no tempo disponível à sua concepção.



Os modelos simbólicos em que uma das variáveis representa uma decisão gerencial a ser tomada denominam-se modelos de decisão. Geralmente, as decisões são tomadas para que um objetivo seja atingido. Portanto, nos modelos de decisão, em geral é adicionada uma variável que represente uma medida de performance dos objetivos (LACHTERMACHER; GERSON, 2009). De acordo com Lachtermacher (2009): Duas características dos modelos simbólicos devem ser ressaltadas:

O modelo sempre será uma simplificação da realidade;

Os detalhes devem ser incorporados ao modelo de forma cuidadosa, para que:

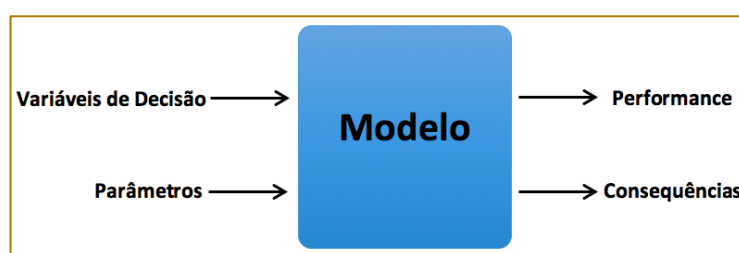
Os resultados atinjam suas necessidades;

Seja consistente com as informações disponíveis;

Seja desenvolvido e analisado no tempo disponível para tal.

A maneira mais simples de representar um modelo simbólico é denominada modelo caixa-preta, conforme figura 1.

Figura 1 - Caixa-preta



Fonte: LACHTERMACHER (2009)

Dentre as ferramentas que vêm ganhando cada vez mais adeptos, as planilhas eletrônicas são as preferidas, pois, além da facilidade de utilização, estão presentes em praticamente todas as empresas modernas. Dentre essas planilhas, as mais utilizadas são o Excel da Microsoft, o Lotus da Lotus/IBM e o Quattro-Pro da Corel. Todas elas dispõem basicamente das mesmas ferramentas, diferindo apenas na forma do comando empregado (LACHTERMACHER; GERSON, 2009).

### 3.4. NORMAS REGULAMENTADORAS PARA A MONTAGEM DAS CABINES DE PINTURA

Na questão da segurança e controle de riscos, há normas indispensáveis que as empresas devem cumprir, neste caso, na montagem da Cabine de Pintura são indispensáveis às normas regulamentadoras 9 e 10.

NR 9 - Programa de Prevenção de Riscos Ambientais: Esta norma tem como finalidade determinar ações para garantir a preservação da saúde e integridade dos trabalhadores, realizando uma identificação dos riscos existentes em seu ambiente de trabalho. O mesmo melhora a produtividade e as condições de trabalho do colaborador. Riscos

Ambientais são elementos como agentes biológicos, físicos ou químicos que, presentes no ambiente de trabalho podem ocasionar danos à saúde do trabalhador por causa de sua natureza, concentração, intensidade ou tempo de exposição;

NR 10 - Segurança em Instalações e Serviços em Eletricidade: É uma norma que tem o objetivo de regulamentar todos os serviços que abrangem a eletricidade e seus riscos, além de garantir a saúde e segurança dos que estejam envolvidos direta e indiretamente nestas atividades e serviços.

## 4. METODOLOGIA

### 4.1. VISITA TÉCNICA À MONTADORA DE CAMINHÕES

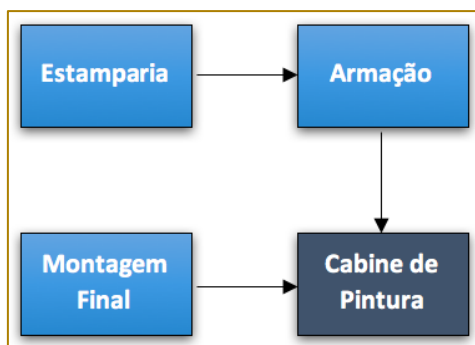
O projeto tem seu início com a visita técnica à montadora de caminhões de São Bernardo do Campo - SP realizada no dia 09/03/2016, com o propósito de conhecer os processos de produção para o estudo deste projeto. A visita foi guiada pelo Engenheiro de pintura e o Gerente de operações. A empresa opera com modelo de produção puxada (modelo Toyota), com concentração na Europa e América Latina (Argentina e Brasil). A linha de produção da fábrica (top coat) funciona da seguinte forma:

- Estamparia;
- Body (armação);
- Cabine de pintura;

- Montagem final.

O fluxo simplificado da linha de produção do sistema está demonstrado na figura 2.

Figura 2 - Linha de produção simplificada da indústria em estudo



Fonte: Indústria do ramo automobilístico (2016)

Estamparia: Onde a chapa/aço tem o design definido;

Armação: Junção de todas as peças da carroceria com uso de tecnologia robótica;

Cabine de Pintura: Estética da cor da carroceria, por meio de processos químicos utilizando-se de propriedades anticorrosivas. Considera-se estruturas para diferentes tipos de cabine;

Montagem Final: São colocados os motores, pneus, vidros, sistema elétrico e outros.

#### 4.2. VISITA TÉCNICA À EMPRESA DE CONSTRUÇÃO DE MÁQUINAS E EQUIPAMENTOS

No dia 15/03/2016 foi realizada outra visita técnica à empresa de construção de

máquinas e equipamentos localizada em São Paulo - SP, guiada pelo Gerente de Operações da empresa, a fim de complementar o estudo do projeto. De acordo com os dados fornecidos, a empresa é centrada na fabricação dos skids. Os skids possuem sistema de memória e sua função é a de transportar as carrocerias na linha de produção. A empresa também produz as cabines de pinturas que compõem a linha de produção da montadora. Os tempos de fabricação e montagem na obra de uma cabine de pintura com as dimensões 6000 x 70000 mm são:

Tempo de fabricação: 88 dias (12 semanas);

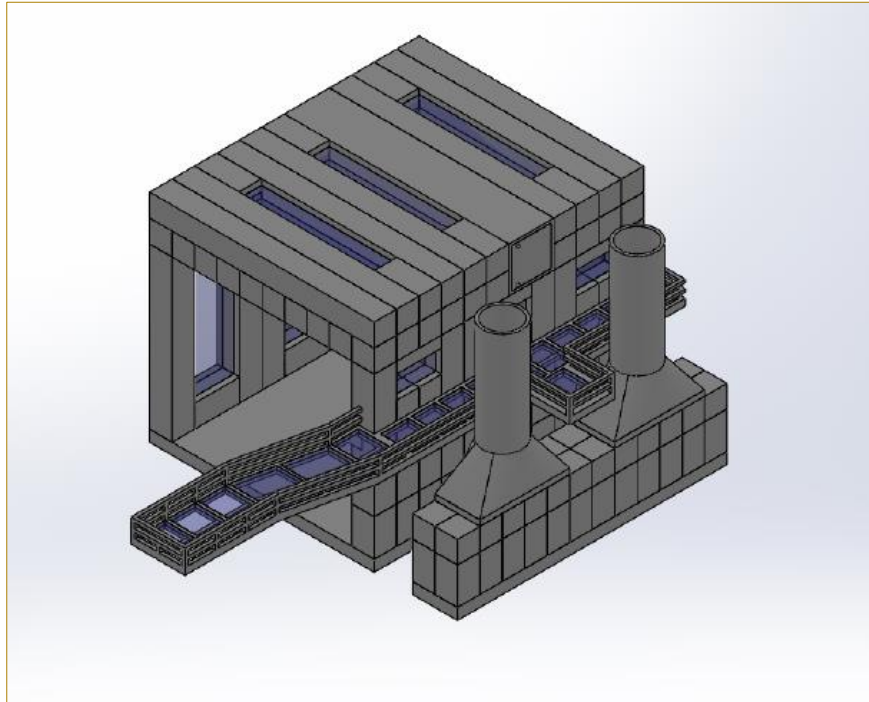
Tempo de montagem na obra (mecânica e elétrica): 47 dias (6 semanas).

Figura 3 - Cabine de Pintura



Fonte: Empresa em estudo (2016)

Figura 4 - Cabine de pintura simplificada em SolidWorks



Fonte: Arquivo Pessoal (2016)

#### 4.3. FERRAMENTAS EMPREGADAS NO PROJETO

Atualmente, tem-se à disposição programas e softwares com o propósito de aperfeiçoar a gestão dos processos, e para este projeto utilizou-se:

Microsoft Excel 2010 - Ferramenta Solver - Modelagens;

Microsoft Project 2010 - Cronograma - gestão do tempo;

Microsoft Visio 2010 (Fluxo 1) - Processo desde o desenho até a montagem final da cabine de pintura;

SolidWorks 2010 - Protótipo 3D - Cabine de Pintura.

#### 4.4. PESQUISAS COM OS FORNECEDORES

Iniciam-se as pesquisas com os fornecedores para a obtenção das cotações das peças e

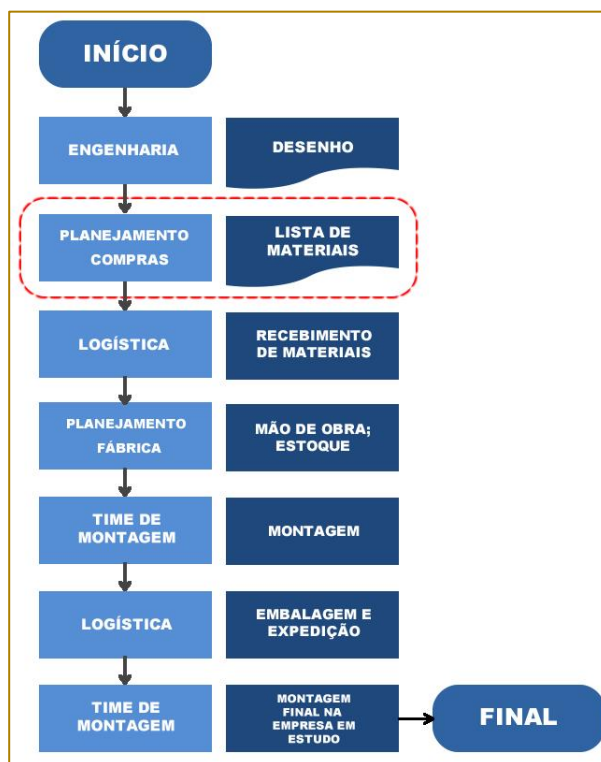
tempo em semanas de fabricação. Dessas pesquisas, utiliza-se uma amostra aleatória de 4 fornecedores nacionais que fabricam as 11 peças da cabine de pintura (tabela 1). Para este estudo, é necessário 2 modelagens na qual, a primeira é de apenas 1 linha de produção em série e a segunda é de 3 linhas de produção em paralelas.

Nesta etapa, o Programa Microsoft Excel é de fundamental importância na posterior análise a ser feita. Os passos a seguir são necessários para o prosseguimento com as modelagens 1 e 2 na ferramenta Solver.

#### 4.5. MAPEAMENTO DO PROCESSO COM FLUXOGRAMA

É desenvolvido o fluxograma (Fluxo 1) do processo de fabricação da cabine de pintura, iniciando-se no desenho até a montagem final, com o foco no processo destacado: Planejamento Compras - Lista de Materiais.

Fluxo 1 - Fluxograma do Processo



Fonte: Arquivo Pessoal (2016)

#### 4.6. MODELAGEM 1 - 1 LINHA DE PRODUÇÃO EM SÉRIE

A modelagem 1 é de uma única linha de produção em série com 4 fornecedores e seus respectivos tempo de fabricação das

peças. Na tabela abaixo é apresentada a relação dos custos das 11 peças de cada fornecedor.

Tabela 1 - Custos das peças de cada fornecedor

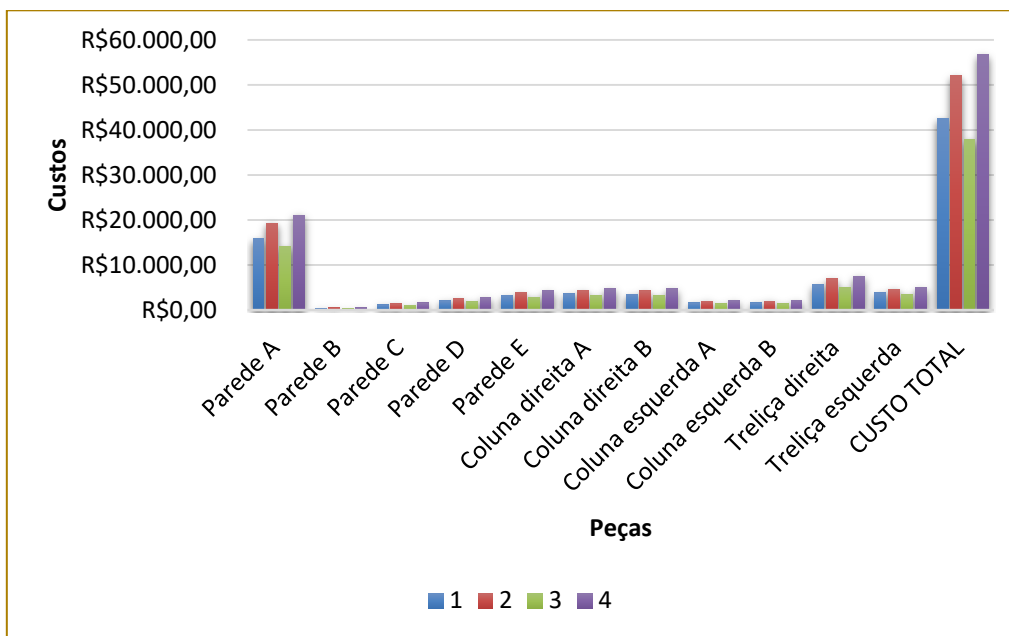
| Fornecedores      | 1             | 2             | 3             | 4             |
|-------------------|---------------|---------------|---------------|---------------|
| Peças             |               |               |               |               |
| Parede A          | R\$ 15.820,33 | R\$ 19.335,96 | R\$ 14.062,51 | R\$ 21.093,77 |
| Parede B          | R\$ 446,72    | R\$ 546,00    | R\$ 397,09    | R\$ 595,63    |
| Parede C          | R\$ 1.194,78  | R\$ 1.460,28  | R\$ 1.062,02  | R\$ 1.593,03  |
| Parede D          | R\$ 2.094,48  | R\$ 2.559,92  | R\$ 1.861,76  | R\$ 2.792,63  |
| Parede E          | R\$ 3.261,50  | R\$ 3.986,28  | R\$ 2.899,11  | R\$ 4.348,67  |
| Coluna direita A  | R\$ 3.568,72  | R\$ 4.361,77  | R\$ 3.172,19  | R\$ 4.758,29  |
| Coluna direita B  | R\$ 3.513,65  | R\$ 4.294,46  | R\$ 3.123,24  | R\$ 4.684,87  |
| Coluna esquerda A | R\$ 1.631,88  | R\$ 1.994,52  | R\$ 1.450,56  | R\$ 2.175,84  |
| Coluna esquerda B | R\$ 1.631,88  | R\$ 1.994,52  | R\$ 1.450,56  | R\$ 2.175,84  |
| Treliça direita   | R\$ 5.660,91  | R\$ 6.918,89  | R\$ 5.031,92  | R\$ 7.547,88  |
| Treliça esquerda  | R\$ 3.796,46  | R\$ 4.640,11  | R\$ 3.374,63  | R\$ 5.061,94  |
| CUSTO TOTAL       | R\$ 42.622,31 | R\$ 52.094,71 | R\$ 37.888,60 | R\$ 56.832,41 |

Fonte: Arquivo Pessoal (2016)

É destacado o fornecedor 4, pois o seu custo total das peças é maior do que os dos demais fornecedores. Em vista dos custos, o fornecedor 4 não é o mais indicado, porém, tem a questão do tempo (tabela 2).

No gráfico 1 é evidenciado que o fornecedor 4 possui os maiores custos por peças e conseqüentemente o maior custo total.

Gráfico 1 - Custos das peças de cada fornecedor



Fonte: Arquivo Pessoal (2016)

A tabela 2 contém a relação do tempo que cada fornecedor leva para fabricar as peças.

Tabela 2 - Tempo de fabricação das peças de cada fornecedor

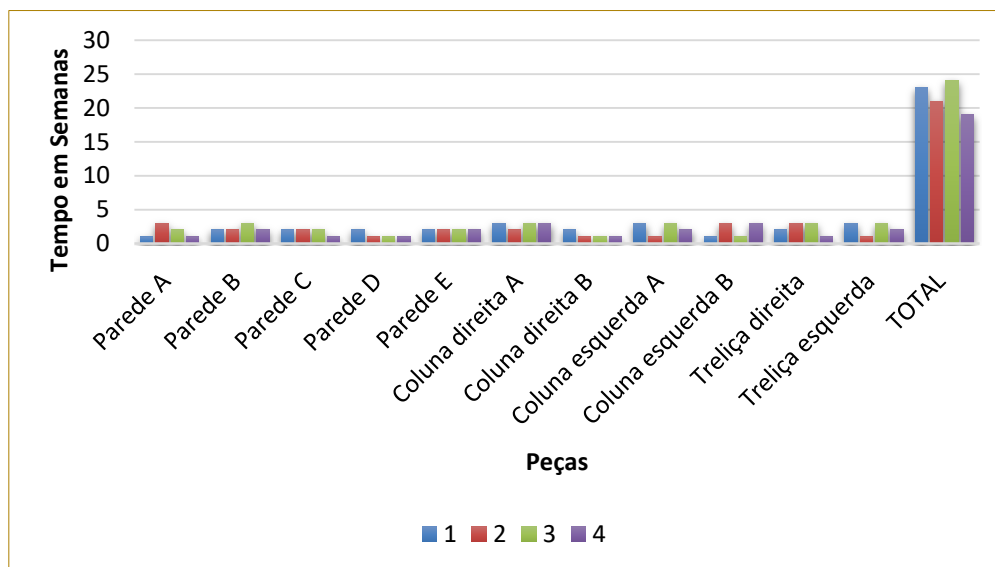
| Fornecedores      | 1                | 2  | 3  | 4  |
|-------------------|------------------|----|----|----|
| Peças             | Tempo em Semanas |    |    |    |
| Parede A          | 1                | 3  | 2  | 1  |
| Parede B          | 2                | 2  | 3  | 2  |
| Parede C          | 2                | 2  | 2  | 1  |
| Parede D          | 2                | 1  | 1  | 1  |
| Parede E          | 2                | 2  | 2  | 2  |
| Coluna direita A  | 3                | 2  | 3  | 3  |
| Coluna direita B  | 2                | 1  | 1  | 1  |
| Coluna esquerda A | 3                | 1  | 3  | 2  |
| Coluna esquerda B | 1                | 3  | 1  | 3  |
| Treliça direita   | 2                | 3  | 3  | 1  |
| Treliça esquerda  | 3                | 1  | 3  | 2  |
| TOTAL             | 23               | 21 | 24 | 19 |

Fonte: Arquivo Pessoal (2016)

Os dados informam que a peça Parede A leva 1 semana para ser fabricada pelo fornecedor 1 e 3 semanas para ser fabricada pelo fornecedor 2 e assim por diante.

O gráfico 2 mostra a relação do tempo de fabricação de cada fornecedor.

Gráfico 2 - Tempo de fabricação das peças de cada fornecedor



Fonte: Arquivo Pessoal (2016)

As restrições são:

O pedido de uma determinada peça só pode ser de um único fornecedor, o que convém a condição do binário (tabela 3);

O tempo de fabricação das peças deve ser menor ou igual a 9 semanas (tabela 4 e 5);

É necessárias quantidades determinadas de peças para a montagem da cabine de pintura (tabela 6).

### Restrição 1:

Tabela 3 - Condição do Binário

| Peças             | Condição do Binário                            |
|-------------------|--|
| Parede A          | $A_1 + A_2 + A_3 + A_4 \leq 1$                 |
| Parede B          | $B_1 + B_2 + B_3 + B_4 \leq 1$                 |
| Parede C          | $C_1 + C_2 + C_3 + C_4 \leq 1$                 |
| Parede D          | $D_1 + D_2 + D_3 + D_4 \leq 1$                 |
| Parede E          | $E_1 + E_2 + E_3 + E_4 \leq 1$                 |
| Coluna direita A  | $Cd_{A1} + Cd_{A2} + Cd_{A3} + Cd_{A4} \leq 1$ |
| Coluna direita B  | $Cd_{B1} + Cd_{B2} + Cd_{B3} + Cd_{B4} \leq 1$ |
| Coluna esquerda A | $Ce_{A1} + Ce_{A2} + Ce_{A3} + Ce_{A4} \leq 1$ |
| Coluna esquerda B | $Ce_{B1} + Ce_{B2} + Ce_{B3} + Ce_{B4} \leq 1$ |
| Treliça direita   | $Td_1 + Td_2 + Td_3 + Td_4 \leq 1$             |
| Treliça esquerda  | $Te_1 + Te_2 + Te_3 + Te_4 \leq 1$             |

Fonte: Arquivo Pessoal (2016)

**Restrição 2:**

$T_{ij}$  = Tempo de fabricação da peça  $i$  do fornecedor  $j$ .

Em relação aos fornecedores:

Tabela 4 - Soma de todas as peças x Tempo de fabricação

| Fornecedores | Soma de todas as peças x Tempo de fabricação   |
|--------------|--|
| 1            | $TA_{1+}TB_{1+}TC_{1+}TD_{1+}TE_{1+}TCd_{A1+}TCd_{B1+}TCe_{A1+}TCe_{B1+}TTd_{1+}TTe_{1+} \leq 9$ |
| 2            | $TA_{2+}TB_{2+}TC_{2+}TD_{2+}TE_{2+}TCd_{A2+}TCd_{B2+}TCe_{A2+}TCe_{B2+}TTd_{2+}TTe_{2+} \leq 9$ |
| 3            | $TA_{3+}TB_{3+}TC_{3+}TD_{3+}TE_{3+}TCd_{A3+}TCd_{B3+}TCe_{A3+}TCe_{B3+}TTd_{3+}TTe_{3+} \leq 9$ |
| 4            | $TA_{4+}TB_{4+}TC_{4+}TD_{4+}TE_{4+}TCd_{A4+}TCd_{B4+}TCe_{A4+}TCe_{B4+}TTd_{4+}TTe_{4+} \leq 9$ |

Fonte: Arquivo Pessoal (2016)

$T_{ij}$  = Tempo de fabricação da peça  $i$  do fornecedor  $j$ .

Em relação às peças:

Tabela 5 - Tempo de fabricação

| Peças             | Tempo de fabricação                               |
|-------------------|---|
| Parede A          | $TQA_{1+}TQA_{2+}TQA_{3+}TQA_{4+} \leq 9$         |
| Parede B          | $TQB_{1+}TQB_{2+}TQB_{3+}TQB_{4+} \leq 9$         |
| Parede C          | $TQC_{1+}TQC_{2+}TQC_{3+}TQC_{4+} \leq 9$         |
| Parede D          | $TQD_{1+}TQD_{2+}TQD_{3+}TQD_{4+} \leq 9$         |
| Parede E          | $TQE_{1+}TQE_{2+}TQE_{3+}TQE_{4+} \leq 9$         |
| Coluna direita A  | $TQCd_{A1+}TQCd_{A2+}TQCd_{A3+}TQCd_{A4+} \leq 9$ |
| Coluna direita B  | $TQCd_{B1+}TQCd_{B2+}TQCd_{B3+}TQCd_{B4+} \leq 9$ |
| Coluna esquerda A | $TQCe_{A1+}TQCe_{A2+}TQCe_{A3+}TQCe_{A4+} \leq 9$ |
| Coluna esquerda B | $TQCe_{B1+}TQCe_{B2+}TQCe_{B3+}TQCe_{B4+} \leq 9$ |
| Treliça direita   | $TQTd_{1+}TQTd_{2+}TQTd_{3+}TQTd_{4+} \leq 9$     |
| Treliça esquerda  | $TQTe_{1+}TQTe_{2+}TQTe_{3+}TQTe_{4+} \leq 9$     |

Fonte: Arquivo Pessoal (2016)

**Restrição 3:**

Tabela 6 - Quantidade de cada Peça

| Peças             | Quantidade de cada Peça                |
|-------------------|--|
| Parede A          | $A_{1+}A_{2+}A_{3+}A_{4+} = 1$         |
| Parede B          | $B_{1+}B_{2+}B_{3+}B_{4+} = 4$         |
| Parede C          | $C_{1+}C_{2+}C_{3+}C_{4+} = 4$         |
| Parede D          | $D_{1+}D_{2+}D_{3+}D_{4+} = 6$         |
| Parede E          | $E_{1+}E_{2+}E_{3+}E_{4+} = 4$         |
| Coluna direita A  | $Cd_{A1+}Cd_{A2+}Cd_{A3+}Cd_{A4+} = 1$ |
| Coluna direita B  | $Cd_{B1+}Cd_{B2+}Cd_{B3+}Cd_{B4+} = 1$ |
| Coluna esquerda A | $Ce_{A1+}Ce_{A2+}Ce_{A3+}Ce_{A4+} = 1$ |
| Coluna esquerda B | $Ce_{B1+}Ce_{B2+}Ce_{B3+}Ce_{B4+} = 1$ |
| Treliça direita   | $Td_{1+}Td_{2+}Td_{3+}Td_{4+} = 1$     |
| Treliça esquerda  | $Te_{1+}Te_{2+}Te_{3+}Te_{4+} = 1$     |

Fonte: Arquivo Pessoal (2016)

#### 4.7. MODELAGEM 2 - 3 LINHAS DE PRODUÇÃO EM PARALELO

A modelagem 2 segue o mesmo padrão da modelagem 1, repetindo os valores das tabelas e dos gráficos 1 e 2. As restrições são:

O pedido de uma determinada peça só pode ser de um único fornecedor, o que convém a condição do binário;

O tempo de fabricação das peças deve ser menor ou igual a 9 semanas;

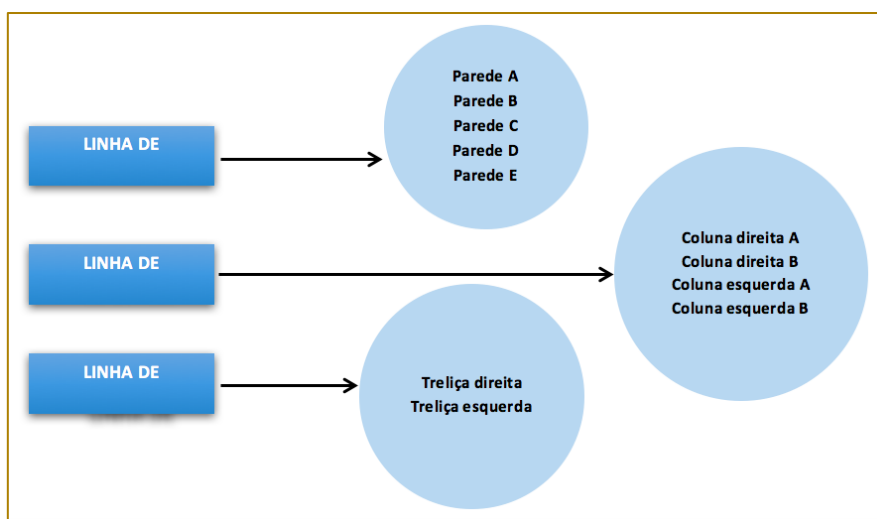
É necessárias determinadas quantidades de peças para a montagem da cabine de pintura.

Restrição 1: Idem da modelagem 1.

Restrição 2:  $T_{ij}$  = Tempo de fabricação da peça  $i$  do fornecedor  $j$ .

Em relação aos fornecedores e para cada linha de produção: A figura 4 mostra a base da modelagem 2 no sistema de 3 linhas de produção, cada linha para um tipo de peça com funcionamento em paralelo. A funções para as linhas de produção 1, 2 e 3 em relação aos fornecedores estão representadas nas tabelas 7, 8 e 9 e em relação as peças na tabela 10.

Figura 5 - Linhas de Produção



Fonte: Arquivo Pessoal (2016)

Tabela 7 - Linha de Produção 1

| Fornecedores | Tempo de fabricação   |
|--------------|---|
| 1            | $TQA_{1+} + TQB_{1+} + TQC_{1+} + TQD_{1+} + TQE_{1+} \leq 9$ |
| 2            | $TQA_{2+} + TQB_{2+} + TQC_{2+} + TQD_{2+} + TQE_{2+} \leq 9$ |
| 3            | $TQA_{3+} + TQB_{3+} + TQC_{3+} + TQD_{3+} + TQE_{3+} \leq 9$ |
| 4            | $TQA_{4+} + TQB_{4+} + TQC_{4+} + TQD_{4+} + TQE_{4+} \leq 9$ |

Fonte: Arquivo Pessoal (2016)

Tabela 8 - Linha de Produção 2

| Fornecedores | Tempo de fabricação  |
|--------------|--|
| 1            | $TQCd_{A1+} + TQCd_{B1+} + TQCe_{A1+} + TQCe_{B1+} \leq 9$ |
| 2            | $TQCd_{A2+} + TQCd_{B2+} + TQCe_{A2+} + TQCe_{B2+} \leq 9$ |
| 3            | $TQCd_{A3+} + TQCd_{B3+} + TQCe_{A3+} + TQCe_{B3+} \leq 9$ |
| 4            | $TQCd_{A4+} + TQCd_{B4+} + TQCe_{A4+} + TQCe_{B4+} \leq 9$ |

Fonte: Arquivo Pessoal (2016)



Tabela 9 - Linha de Produção 3

| Fornecedores | Tempo de fabricação      |
|--------------|--------------------------|
| 1            | $TQTd_{1+}TQTe_1 \leq 9$ |
| 2            | $TQTd_{2+}TQTe_2 \leq 9$ |
| 3            | $TQTd_{3+}TQTe_3 \leq 9$ |
| 4            | $TQTd_{4+}TQTe_4 \leq 9$ |

Fonte: Arquivo Pessoal (2016)

$T_{ij}$  = Tempo de fabricação da peça  $i$  do fornecedor  $j$ .

Em relação às peças:

Tabela 10 - Tempo em relação as peças

| Peças             | Tempo de fabricação                               |
|-------------------|---|
| Parede A          | $TQA_{1+}TQA_{2+}TQA_{3+}TQA_{4+} \leq 9$         |
| Parede B          | $TQB_{1+}TQB_{2+}TQB_{3+}TQB_{4+} \leq 9$         |
| Parede C          | $TQC_{1+}TQC_{2+}TQC_{3+}TQC_{4+} \leq 9$         |
| Parede D          | $TQD_{1+}TQD_{2+}TQD_{3+}TQD_{4+} \leq 9$         |
| Parede E          | $TQE_{1+}TQE_{2+}TQE_{3+}TQE_{4+} \leq 9$         |
| Coluna direita A  | $TQCd_{A1+}TQCd_{A2+}TQCd_{A3+}TQCd_{A4+} \leq 9$ |
| Coluna direita B  | $TQCd_{B1+}TQCd_{B2+}TQCd_{B3+}TQCd_{B4+} \leq 9$ |
| Coluna esquerda A | $TQCe_{A1+}TQCe_{A2+}TQCe_{A3+}TQCe_{A4+} \leq 9$ |
| Coluna esquerda B | $TQCe_{B1+}TQCe_{B2+}TQCe_{B3+}TQCe_{B4+} \leq 9$ |
| Treliça direita   | $TQTd_{1+}TQTd_{2+}TQTd_{3+}TQTd_{4+} \leq 9$     |
| Treliça esquerda  | $TQTe_{1+}TQTe_{2+}TQTe_{3+}TQTe_{4+} \leq 9$     |

Fonte: Arquivo Pessoal (2016)

Restrição 3: Idem da modelagem 1.

## 5. RESULTADOS E DISCUSSÕES

A função objetivo é minimizar o custo em base das restrições de redução do tempo de

fabricação das peças da cabine de pintura conforme apresentado na equação 1.

Equação 1: Função minimizar custo

$$\text{Min} = F = P_A^1 \cdot Q_A^1$$

$$\sum P_{ij} \cdot Q_{ij}$$

**Legenda:**

P = Preço da peça  $i$  do fornecedor  $j$ .

Q = Quantidade da peça  $i$  do fornecedor  $j$ .

Nas tabelas 11 e 12 apresentamos os resultados encontrados pela ferramenta Solver do Excel.

Tabela 11 - Valores obtidos na Modelagem 1

| Peças             | Fornecedores |   |   |   | Condição | Peças             | Fornecedores |   |   |   |
|-------------------|--------------|---|---|---|----------|-------------------|--------------|---|---|---|
|                   | 1            | 2 | 3 | 4 |          |                   | 1            | 2 | 3 | 4 |
| Parede A          | 1            | 0 | 0 | 0 | 1        | Parede A          | 1            | 0 | 0 | 0 |
| Parede B          | 1            | 0 | 0 | 0 | 1        | Parede B          | 4            | 0 | 0 | 0 |
| Parede C          | 0            | 0 | 0 | 1 | 1        | Parede C          | 0            | 0 | 0 | 4 |
| Parede D          | 0            | 1 | 0 | 0 | 1        | Parede D          | 0            | 6 | 0 | 0 |
| Parede E          | 0            | 0 | 1 | 0 | 1        | Parede E          | 0            | 0 | 4 | 0 |
| Coluna direita A  | 0            | 0 | 0 | 1 | 1        | Coluna direita A  | 0            | 0 | 0 | 1 |
| Coluna direita B  | 0            | 1 | 0 | 0 | 1        | Coluna direita B  | 0            | 1 | 0 | 0 |
| Coluna esquerda A | 0            | 1 | 0 | 0 | 1        | Coluna esquerda A | 0            | 1 | 0 | 0 |
| Coluna esquerda B | 0            | 0 | 1 | 0 | 1        | Coluna esquerda B | 0            | 0 | 1 | 0 |
| Treliça direita   | 0            | 0 | 0 | 1 | 1        | Treliça direita   | 0            | 0 | 0 | 1 |
| Treliça esquerda  | 0            | 1 | 0 | 0 | 1        | Treliça esquerda  | 0            | 1 | 0 | 0 |

| Tempo | Fornecedores |   |   |   |
|-------|--------------|---|---|---|
|       | 1            | 2 | 3 | 4 |
|       | 9            | 9 | 9 | 8 |

| Peças             | Tempo |
|-------------------|-------|
| Parede A          | 1     |
| Parede B          | 8     |
| Parede C          | 4     |
| Parede D          | 6     |
| Parede E          | 8     |
| Coluna direita A  | 3     |
| Coluna direita B  | 1     |
| Coluna esquerda A | 1     |
| Coluna esquerda B | 1     |
| Treliça direita   | 1     |
| Treliça esquerda  | 1     |

| Custos | Fornecedores  |               |               |               |
|--------|---------------|---------------|---------------|---------------|
|        | 1             | 2             | 3             | 4             |
|        | R\$ 17.607,23 | R\$ 26.288,59 | R\$ 13.047,01 | R\$ 18.678,31 |

|              |               |
|--------------|---------------|
| <b>Custo</b> | R\$ 75.621,13 |
|--------------|---------------|

Fonte: Arquivo Pessoal (2016)

Tabela 12 - Valores obtidos na Modelagem 2

| Peças             | Fornecedores |   |   |   | Condição | Peças             | Fornecedores |   |   |   |
|-------------------|--------------|---|---|---|----------|-------------------|--------------|---|---|---|
|                   | 1            | 2 | 3 | 4 |          |                   | 1            | 2 | 3 | 4 |
| Parede A          | 0            | 0 | 1 | 0 | 1        | Parede A          | 0            | 0 | 1 | 0 |
| Parede B          | 0            | 0 | 0 | 1 | 1        | Parede B          | 0            | 0 | 0 | 4 |
| Parede C          | 0            | 1 | 0 | 0 | 1        | Parede C          | 0            | 4 | 0 | 0 |
| Parede D          | 0            | 0 | 1 | 0 | 1        | Parede D          | 0            | 0 | 6 | 0 |
| Parede E          | 1            | 0 | 0 | 0 | 1        | Parede E          | 4            | 0 | 0 | 0 |
| Coluna direita A  | 0            | 0 | 1 | 0 | 1        | Coluna direita A  | 0            | 0 | 1 | 0 |
| Coluna direita B  | 0            | 0 | 1 | 0 | 1        | Coluna direita B  | 0            | 0 | 1 | 0 |
| Coluna esquerda A | 0            | 0 | 1 | 0 | 1        | Coluna esquerda A | 0            | 0 | 1 | 0 |
| Coluna esquerda B | 0            | 0 | 1 | 0 | 1        | Coluna esquerda B | 0            | 0 | 1 | 0 |
| Treliça direita   | 0            | 0 | 1 | 0 | 1        | Treliça direita   | 0            | 0 | 1 | 0 |
| Treliça esquerda  | 0            | 0 | 1 | 0 | 1        | Treliça esquerda  | 0            | 0 | 1 | 0 |

| Tempo | Fornecedores - L1 |   |   |   |
|-------|-------------------|---|---|---|
|       | 1                 | 2 | 3 | 4 |
|       | 8                 | 8 | 8 | 8 |

| Tempo | Fornecedores - L2 |   |   |   |
|-------|-------------------|---|---|---|
|       | 1                 | 2 | 3 | 4 |
|       | 8                 | 0 | 8 | 0 |

| Tempo | Fornecedores - L3 |   |   |   |
|-------|-------------------|---|---|---|
|       | 1                 | 2 | 3 | 4 |
|       | 0                 | 0 | 6 | 0 |

| Custos | Fornecedores  |              |               |              |
|--------|---------------|--------------|---------------|--------------|
|        | 1             | 2            | 3             | 4            |
|        | R\$ 13.046,01 | R\$ 5.841,13 | R\$ 42.836,16 | R\$ 2.382,53 |

|              |               |
|--------------|---------------|
| <b>Custo</b> | R\$ 64.105,83 |
|--------------|---------------|

| Peças             | Tempo |
|-------------------|-------|
| Parede A          | 2     |
| Parede B          | 8     |
| Parede C          | 8     |
| Parede D          | 6     |
| Parede E          | 8     |
| Coluna direita A  | 3     |
| Coluna direita B  | 1     |
| Coluna esquerda A | 3     |
| Coluna esquerda B | 1     |
| Treliça direita   | 3     |
| Treliça esquerda  | 3     |

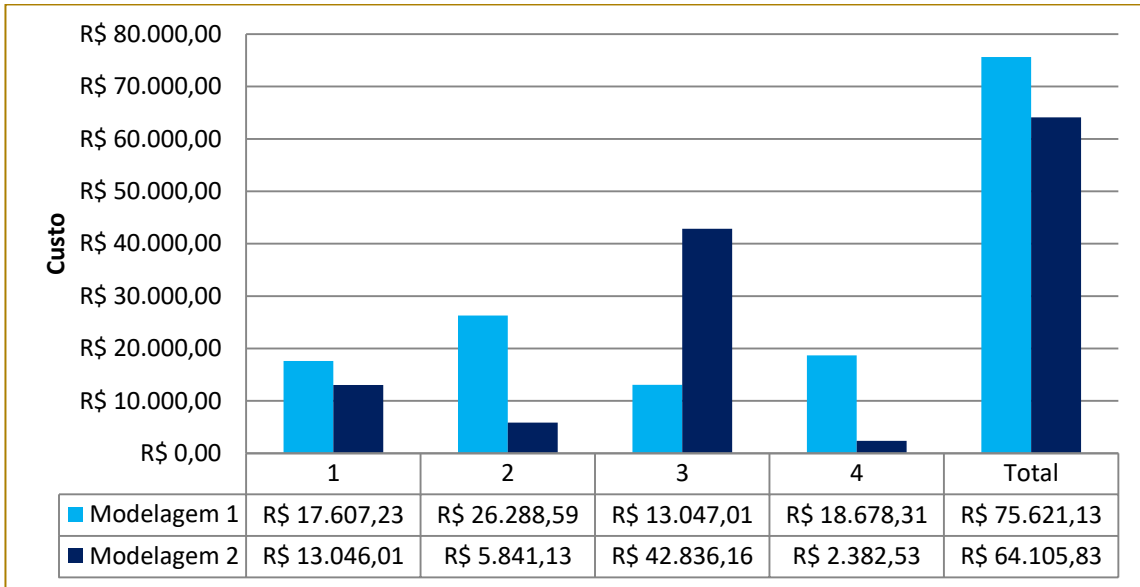
Fonte: Arquivo Pessoal (2016)

A tabela 12 demonstra que a modelagem 2 tem os custos mais relativos em relação a modelagem 1, e conseqüentemente custo total menor.

Visto que o tempo de fabricação das peças informado pela é de 88 dias, equivalente a

aproximadamente 12 semanas, é possível visualizar a redução do tempo de fabricação já que nas restrições das modelagens o tempo máximo foi de até 9 semanas, equivalente a aproximadamente 67 dias.

Gráfico 3 - Comparação de custos dos fornecedores das modelagens 1 e 2



Fonte: Arquivo Pessoal (2016)

Portanto, no fluxo do processo o planejamento e compras é otimizado 75% em relação ao

tempo de fabricação das peças da cabine de pintura.

100% ----- 12 semanas  
 X ----- 9 semanas

Na figura 6 mostra o protótipo construído para representar o fluxograma do processo da construção da cabine de pintura.

Figura 6 - Protótipos e fluxo do processo



Fonte: Arquivo Pessoal (2016)

### 5.1. COMPARAÇÃO ENTRE AS MODELAGENS 1 E 2

A partir de uma análise comparativa entre as duas modelagens:

Considerou-se que o custo total de uma linha de produção em série é maior que o custo total de 3 linhas de produção em paralelo;

Todas as hipóteses possíveis são realizadas para atingir a função objetivo e satisfazer todas as restrições;

3 linhas de produção aproveitam melhor o espaço e o tempo.

Quadro 1 - Comparativo das modelagens 1 e 2



Fonte: Arquivo Pessoal (2016)

## 6. CONCLUSÃO

Portanto, conclui-se que a análise em relação aos custos e tempo de fabricação das peças componentes da cabine de pintura foi realizada considerando duas modelagens, uma com uma linha de produção em série e a outra com 3 linhas de produção em paralelo. Verificou-se que o caminho que mais satisfizesse o objetivo consistiu na modelagem seguindo diversas linhas de produção, em que foi

analisado maior eficiência em relação ao custo e tempo.

## 7. AGRADECIMENTOS

Aos nossos professores Alessandro Ranulfo Lima Nery, Ricardo Luiz Ciuccio e Sérgio Mancini pelas orientações.

Ao nosso Deus pela força.

Aos nossos familiares e amigos pelo apoio e paciência.

## REFERÊNCIAS

[1]. BRASIL. NR9, de 29 de dezembro de 1994. Programa de prevenção de riscos ambientais. São Paulo: Portaria SSST. Disponível em: <<http://www.guiatrabalhista.com.br/legislacao/nr/nr9.htm>>. Acesso em: 30 abril 2016.

[2]. BRASIL. NR10, de 07 de dezembro de 2004. Segurança em instalações e serviços

em eletricidade. São Paulo: Portaria Ministro de estado do trabalho e emprego. Disponível em:

<<http://www.guiatrabalhista.com.br/legislacao/nr/nr10.htm>>. Acesso em: 30 abril 2016.

[3]. LACHTERMACHER, Gerson. Pesquisa Operacional na tomada de decisões. 4ª Edição. Brasil: Pearson, 2009.

# Capítulo 21

## ROTEIRIZAÇÃO DE VEÍCULOS VERSUS PERFORMANCE TÉRMICA NA DISTRIBUIÇÃO DE SORVETES

*William Daiti Kobo Takeshita*

*Gustavo Henrique Moresco*

*Vanina Macowski Durski Silva*

**Resumo:** Este artigo aborda uma pesquisa, bibliográfica e estudo de caso a respeito de um tema relacionado à cadeia de suprimentos, denominado Cadeia do Frio. Na intenção de manter a qualidade dos produtos da Cadeia do Frio durante a distribuição física, concomitantemente buscando por reduções de custos, vários cuidados distintos são necessários, como o armazenamento, o tempo de transporte e controle de temperatura durante o processo. Os produtos dessa cadeia exigem monitoramento térmico, e em função da tecnologia atual, de controle térmico e de métodos de roteirização, a distribuição física dos mesmos pode ocorrer de modo otimizado. O estudo de caso aqui proposto teve como referência uma indústria de sorvetes a qual produz e distribui diversos produtos no entorno de Joinville-SC. O artigo teve como objetivo a análise do caso sob a ótica da Cadeia do Frio, mapeando o fluxograma das operações de expedição e distribuição realizado pela empresa em estudo. Propôs-se rotas de distribuição para os produtos de maneira a reduzir a distância percorrida pelos veículos utilizados pela empresa e, ao mesmo tempo, analisou-se a variação da temperatura do compartimento de carga dos veículos durante as entregas, identificando que a maior taxa de transferência de calor ocorre quando a diferença entre a temperatura externa e interna é elevada. Esta maior taxa de transferência de calor provoca um maior aumento da temperatura interna do compartimento, que reduz gradativamente conforme a temperatura interna é elevada.

**Palavras-Chave:** Cadeia do frio; distribuição física de produtos; roteirização; sorvetes; performance térmica

\* Artigo originalmente apresentado no Encontro Nacional de Engenharia de Produção (ENEGEP, 2017)

## 1. INTRODUÇÃO

A necessidade da busca de redução de custos e melhoria no controle de qualidade dos alimentos transportados nas cadeias de suprimentos, tem traduzido uma melhor eficiência no transporte dessas cargas e incentivado cada vez mais os investimentos em tecnologias. Assim surge o conceito de Cadeia do Frio (CF), onde os produtos necessitam de controle de temperatura, umidade e outras variáveis externas ambientais, tornando a busca por melhorias mais difíceis. Desde a produção ou colheita, até o momento do consumo, são necessários cuidados especiais com o produto, que são específicos de cada um deles. Essa grande importância no setor se deve à globalização e ao livre comércio de alimentos (Likar e Jevsnik, 2003).

Segundo Pereira *et al.* (2010), a Cadeia do Frio consiste em todo o processo de armazenamento, conservação, distribuição, transporte e manipulação dos produtos, como a manutenção da temperatura. Todos esses processos devem ser rigorosamente efetuados com a manutenção da temperatura, desse modo os produtos transportados sofrem as menores perdas possíveis na sua qualidade. A CF precisa funcionar de maneira segura e garantir que não haja diferença significativa nas temperaturas entre transporte, armazenamento e até mesmo, na conservação doméstica pelo consumidor final (Pereira, 2011).

Heap (2006), afirma que quanto maior a distância de deslocamento, maiores devem ser os cuidados com o controle da temperatura. Porém, o autor ainda afirma que o inverso dessa informação não pode ser tido como verdadeiro, mas é o que acontece em inúmeros casos, tornando os deslocamentos curtos os pontos mais críticos da cadeia, devido à falta de cuidados com a carga.

Conforme Torassi (2009), os microrganismos proliferam-se muito rápido, gerando danos irreparáveis nos alimentos como mudanças no sabor, odor e cor, além de liquefação, podridão, entre outros. E para que esse processo se desenvolva causando os menores danos possíveis ao alimento, pode-se empregar alguns métodos de Pesquisa Operacional (PO), a fim de auxiliar no monitoramento e controle de temperatura.

Sendo assim, o trabalho pretende apresentar a cadeia logística de produtos congelados e refrigerados e realizar um estudo de caso em

uma empresa do ramo de sorvetes situada na cidade de Joinville - SC, analisando dados fornecidos como locais de entrega, temperatura do caminhão, e também, realizando visitas técnicas para uma melhor análise dos processos, desde os pedidos até a entrega do produto. Além dos conceitos da CF, será utilizada o conceito de PO, aplicando uma heurística com o intuito de otimizar a rota de entrega empregada pela empresa. Juntamente, será analisado o desempenho térmico em cada um desses pontos, com uma adaptação das análises realizadas previamente por Novaes *et al.* (2005), respeitando as restrições encontradas no problema.

## 2. REFERENCIAL TEÓRICO

### 2.1. ARMAZENAGEM E TRANSPORTE NA CADEIA DO FRIO

A CF é composta por inúmeros processos e para garantir a qualidade dos produtos, que necessitam de um rigoroso controle de temperatura, é necessário respeitar as especificações da carga nesses processos. A temperatura é uma importante variável, mas outros parâmetros são necessários no monitoramento, como umidade, atmosfera modificada, pressão etc. (CARVALHO, NOVAES e LIMA JUNIOR, 2012).

De certa forma a "Cadeia do Frio é uma vertente da cadeia de suprimentos e da refrigeração, pois agrega, de um lado, a gestão logística e suas funções de controle e monitoramento, e de outros, os recursos técnicos da engenharia de refrigeração" (SILVA, 2010, p. 7). Sendo assim há grande necessidade de avanços tecnológicos e de gestão nessas duas áreas para uma boa gestão da CF.

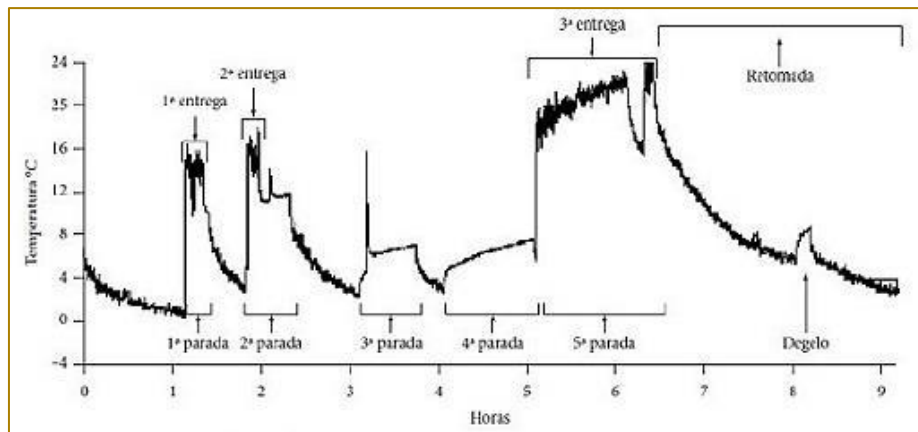
A etapa do transporte representa um dos pontos mais importantes da CF devido aos cuidados especiais para manter a qualidade do produto. Segundo Carvalho (2013) o transporte é o maior responsável pela perda da qualidade dos produtos devido à falta de controle do ambiente.

Conforme Baptista (2006), o dever da refrigeração durante o transporte é somente zelar pela manutenção da temperatura, e não tem como objetivo resfriar a carga. Para que isso ocorra, é necessário o pré-resfriamento do compartimento de carga, prática bastante utilizada durante essa etapa da cadeia (Silva, 2010).

Inserido na atividade do transporte, o momento de entrega dos produtos requer uma maior atenção, pois são realizadas inúmeras aberturas de portas e paradas do caminhão, fazendo com que a temperatura interna do compartimento de carga aumente e

podendo causar prejuízos aos produtos. A figura 1 representa uma situação prática onde ocorrem 3 entregas e 5 paradas do motor, que fornece energia para o refrigerador, bem como as variações de temperatura ao longo do tempo.

Figura 1 – Histórico da variação de temperatura em uma entrega.



Fonte: Pereira *et al.* (2010).

O transporte representa, para a maioria das empresas, o elemento de maior valor logístico (Ballou, 2001). Contudo, as perdas de produtos nas CF estão relacionadas diretamente com uma má gestão dos processos logísticos. Então para isso, são necessárias medidas que diminuam os custos em transportes e investimentos em processos de controle da temperatura. Uma boa prática é a roteirização dos veículos de entrega para diminuir a distância total percorrida.

### 3. ROTEIRIZAÇÃO

As necessidades logísticas de uma empresa começam pela redução de custos, a fim de reduzir os preços finais dos produtos, tornando a organização mais competitiva. Um campo de redução de custos, como citado anteriormente, é o ramo dos transportes. Conforme Novaes (2001), tradicionalmente, as empresas tendem a focar seus custos dentro do domínio estrito das suas atividades.

Visto que as despesas que ocorrem externas à empresa são de difícil quantificação, esse fator é desconsiderado por muitas organizações. Porém deslocamentos excessivos e viagens desnecessárias fazem com que os custos relacionados ao transporte aumentem, conseqüentemente, aumentando o preço final do produto. Uma ferramenta para otimizar os deslocamentos efetuados por uma

empresa pode ser obtida através de conceitos empregados pela Pesquisa Operacional, intitulado roteirização. Para maiores informações, ver Goldberg e Luna (2000).

Os métodos de roteirização de veículos se dividem em duas categorias: os exatos e os heurísticos. Conforme Cechin (2006, apud BEZERRA, 2005) as abordagens exatas devem ser utilizadas para pequenas instâncias do problema, pois para problemas maiores, elas demandariam muito tempo de processamento. Portanto, neste trabalho, como o cenário analisado possui mais que 15 nós, resultando em mais de 200 arcos de ligação entre os nós, optou-se por utilizar métodos heurísticos para a resolução. Segundo Hillier e Lieberman (2013), um método heurístico é um procedimento que busca encontrar uma solução viável, não necessariamente uma solução ótima. Porém, a solução encontrada pelo método heurístico é capaz de fornecer uma solução próxima a ideal.

#### 3.1. PROBLEMA DO VIZINHO MAIS PRÓXIMO

Para resolução do problema de roteirização utilizou-se o *software* MATLAB, a fim de minimizar a distância total percorrida pelo veículo de entrega sob a ótica do problema

do vizinho mais próximo, aplicado por Moresco e Silva (2016) em um estudo realizado previamente. O algoritmo baseia-se em um ponto de origem pré-definido, local da empresa em estudo, e a partir deste, busca o ponto mais próximo e o adiciona a rota. Ao término da execução não restaram mais pontos de entrega para o algoritmo selecionar, então definiu-se que ele retorne até o local da empresa para terminar o percurso. Após isso, a rota gerada define que o veículo deve sair da empresa, passar em todos os pontos de entrega somente uma vez e retornar para o mesmo local de origem.

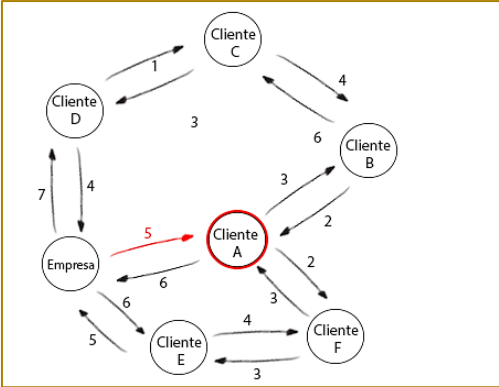
A fim de obter diferentes respostas, foram desenvolvidos Moresco e Silva (2017) quatro

métodos que possuem funcionamento semelhante, distinguindo-se apenas pelo ponto inicial de cada problema. Isso deve-se ao fato de diferentes problemas de roteirização possuírem diferentes características, portanto nem sempre o mesmo método retornará a menor distância para problemas distintos.

### 3.1.1. MÉTODO 1

O primeiro método a ser analisado parte do princípio que o ponto de saída do algoritmo é o mesmo ponto inicial da rota, ou seja, o local da empresa em estudo. A figura 2 ilustra o funcionamento do método.

Figura 2: Funcionamento do método 1.



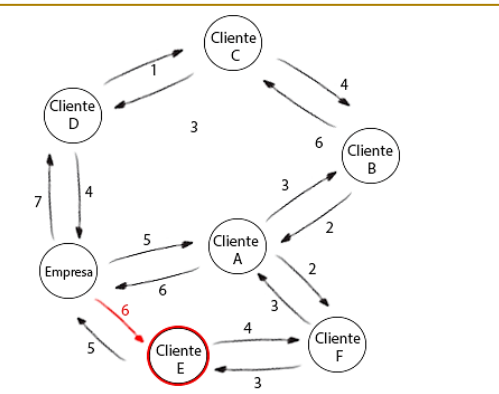
Fonte: Autoria própria.

Como mostra a figura 2, o ponto escolhido é o caminho de ida mais próximo do ponto inicial. O prosseguimento do método é realizado da mesma maneira até que o ponto final seja a empresa.

### 3.1.2. MÉTODO 2

O método dois possui grande semelhança ao primeiro, sendo somente a escolha do vizinho do ponto inicial do problema diferente, como mostra a figura 3.

Figura 3: Funcionamento do método 2.



Fonte: Autoria própria.

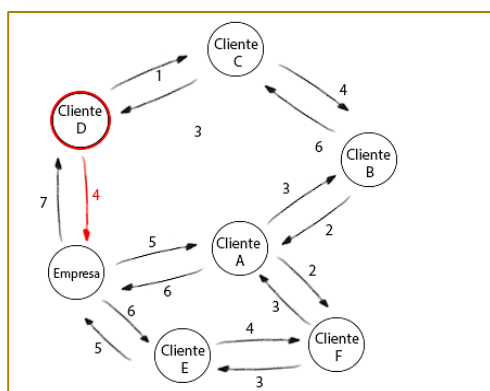


Como é possível perceber na figura 3, a escolha do próximo ponto da rota, nesse método, se dá pelo segundo vizinho mais próximo ao ponto inicial do problema. O restante da rota é realizado da mesma maneira que no método um, ou seja, incluindo na rota o vizinho mais próximo ao ponto atual que ainda não tenha sido incluído na rota.

### 3.1.3. MÉTODO 3

O terceiro método possui uma análise inversa do realizado no método um. O algoritmo realiza uma busca por qual vizinho teria a menor distância para chegar ao ponto final, ou seja, ao contrário dos métodos anteriores que buscavam o vizinho mais próximo saindo do ponto atual, este tem o objetivo de buscar o vizinho pelo qual seria mais próximo chegar ao ponto atual. A figura 4 ilustra o funcionamento do método.

Figura 4: Funcionamento do método 3.



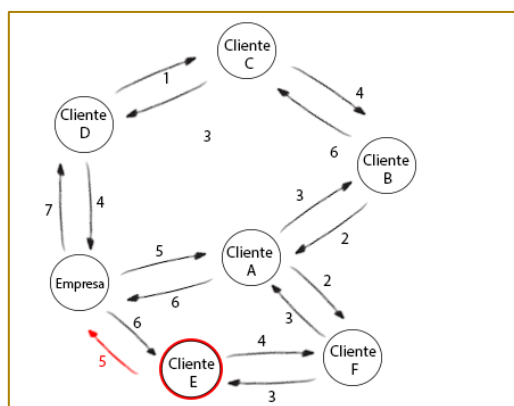
Fonte: Autoria própria.

Como é possível perceber na figura 4, o ponto escolhido se refere a origem mais próxima tendo como destino o ponto atual. Desta maneira a rota será preenchida de maneira inversa ao que será realizado durante as entregas.

### 3.1.4. MÉTODO 4

O quarto e último método possui o mesmo mecanismo de análise que o terceiro, porém partindo do princípio que a primeira escolha será do segundo vizinho mais próximo ao ponto atual. A figura 5 mostra o funcionamento do método.

Figura 5: Funcionamento do método 4.



Fonte: Autoria própria.

Como ilustra a figura 5, o método quatro possui mesma relação ao terceiro método que o método um possui com o segundo. Alterna-se somente a maneira de escolha do primeiro vizinho. O restante do quarto método é realizado da mesma maneira que o terceiro.

#### 4. COMPORTAMENTO TÉRMICO

Após a criação da melhor rota pelo algoritmo, surge a necessidade de avaliar e acompanhar o desempenho térmico em cada um dos

$$\theta_{iB} = \alpha_0 (\theta_{iA})^{\alpha_1} \times (\theta_{ext})^{\alpha_2} \times \exp[\gamma(t_B - t_A)] \quad (1)$$

Onde  $\theta_{iB}$  representa a temperatura no valor futuro, sendo  $A$  o ponto inicial do instante analisado e  $B$  o ponto final.  $\theta_{iA}$  representa a temperatura no ponto  $A$  e  $\theta_{ext}$ , a temperatura externa,  $t_B - t_A$  indica o intervalo de tempo entre o instante  $A$  e  $B$ .  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$  e  $\gamma$  são constantes que foram calibradas conforme o cenário proposto, através de dados de rastreamento da temperatura de um roteiro de entrega.

pontos de entrega. A fim de realizar o acompanhamento térmico ao longo de todo o processo de entregas, utilizou-se uma formulação adaptada da proposta por Novaes *et al.* (2005). Portanto, foi necessário dividir o processo em etapas, com distintas equações que representam a temperatura. Porém, as etapas utilizam de uma equação base, diferenciando-se somente nas constantes avaliadas.

*Fase de deslocamento até os clientes:* Essa etapa representa o início e o final do processo, onde ocorre o deslocamento da empresa até o primeiro cliente e do último cliente até a volta a empresa. Nessa etapa desconsidera-se a interferência da temperatura externa e da temperatura do instante anterior, avaliando-se somente a carga refrigerada com constante sendo  $\gamma = 0,003695$ , resultando na equação:

$$\theta_{iB} = \exp[\gamma(t_B - t_A)] \quad (2)$$

*Fase de descarga:* Essa etapa refere-se ao momento de maior acréscimo de temperatura durante o processo, pois nesse instante o compartimento de carga está aberto e ocorre a troca de calor imediata entre veículo e ambiente. Nessa etapa utiliza-se a equação 1 para representar a temperatura com as constantes sendo  $\alpha_0 = 0,365086$ ,  $\alpha_1 = 0,990072$ ,  $\alpha_2 = 0,35764$  e  $\gamma = -0,002858$ .

*Fase de deslocamento entre as entregas:* Esse estágio representa o momento de deslocamento entre um cliente e outro, tempo em que o compartimento refrigerado readequa a temperatura do compartimento de carga. As constantes foram calibradas sendo  $\alpha_0 = 11,67614$ ,  $\alpha_1 = 0,89134$ ,  $\alpha_2 = -0,76252$  e  $\gamma = 0,00739$ .

#### 5. ESTUDO DE CASO

Através de visitas técnicas, acompanhamento de carregamento e de entregas além de conversas com o setor logístico e de

transporte de uma empresa que fabrica e distribui sorvetes na cidade de Joinville há 25 anos, foram realizadas pesquisas e análises para melhor entender o comportamento da temperatura, no interior dos veículos que realizam a distribuição dos sorvetes, tendo em vista a escolha de uma rota de entrega.

A empresa em estudo está presente em boa parte dos estados de Santa Catarina e Paraná, mas iniciou-se com produção de sorvetes artesanais. A empresa possui seu próprio serviço de distribuição (matriz) situado juntamente com a fábrica. Também em sua própria planta fabril está localizada a área de armazenamento, que é composta por uma câmara fria a qual possui 3 ambientes com temperaturas diferentes, variando de  $-10^{\circ}\text{C}$  até  $-40^{\circ}\text{C}$ . Bragante (2010) diz que após o sorvete já estar embalado é necessária uma etapa da produção chamada endurecimento, que deve acontecer a  $-25^{\circ}\text{C}$  para evitar formação de cristais de gelo. Na empresa em estudo, esse processo acontece na câmara fria e o produto permanece por lá aproximadamente 24 horas. Ainda na câmara fria, os produtos são separados e

armazenados conforme sua especificação. São elas: potes de 1, 2,6 e 10 litros, picolés e açai. Os sabores vão desde tradicionais até sabores especiais, assim fazendo com que exista uma gama de dimensões de embalagens.

A empresa ainda possui outros três Centros de Distribuição (CD) para auxiliar no serviço de distribuição, localizados em Florianópolis, Mafra e Curitiba.

A figura 6 mostra a variedade e diferença das embalagens.

Figura 6 – Variedades e diferença das embalagens.



Fonte: Autoria própria.

Segundo Bragante (2010) o derretimento causa uma má aparência no sorvete, o que é prejudicial ao consumidor. Portanto o derretimento deve ser evitado durante o processo de armazenagem e distribuição para evitar perdas, para isso o produto deve ser mantido o mais próximo da temperatura de  $-25^{\circ}\text{C}$ . Segundo os motoristas da empresa em estudo, o processo de derretimento e perda do produto acontece a partir de uma temperatura próxima do positivo.

A empresa possui uma frota de 33 veículos distribuídos entre os CD's, com a grande maioria concentrada na matriz. Os veículos são equipados com baús frigoríficos do tipo sorveteiro com divisórias de alumínio, que facilitam o carregamento e auxiliam em um melhor encaixe dos produtos, evitando tombamentos na hora do transporte. O sistema de refrigeração utilizado é o eutético, com alimentação elétrica independente do

motor. No processo de refrigeração do compartimento de carga é injetado um gás refrigerante impulsionado pelo motor elétrico para um circuito de serpentina das placas. Assim o gás é propagado por convecção e a solução eutética congela, podendo chegar a  $-40^{\circ}\text{C}$ , mas devido às trocas de calor com o meio e a carroceria, a estabilidade acontece em torno de  $-28^{\circ}\text{C}$ . Dentre os veículos da frota, há 3 tipos de veículos de tamanhos diferentes: um maior, que é utilizado para abastecer os CD's; um menor, que é utilizado para pequenas entregas e restrições de tráfego e, por último, o veículo que será abordado neste trabalho, um intermediário que é usado para entregas na região.

A figura 7 mostra a carroceria do tipo sorveteiro, as divisões internas e as cortinas de PVC nas portas para diminuir as trocas de calor com o ambiente externo.

Figura 7 – Caminhão, suas divisórias e cortinas de PVC.



Fonte: Autoria própria.

No processo de carregamento do caminhão, os produtos são divididos por tipo entre as

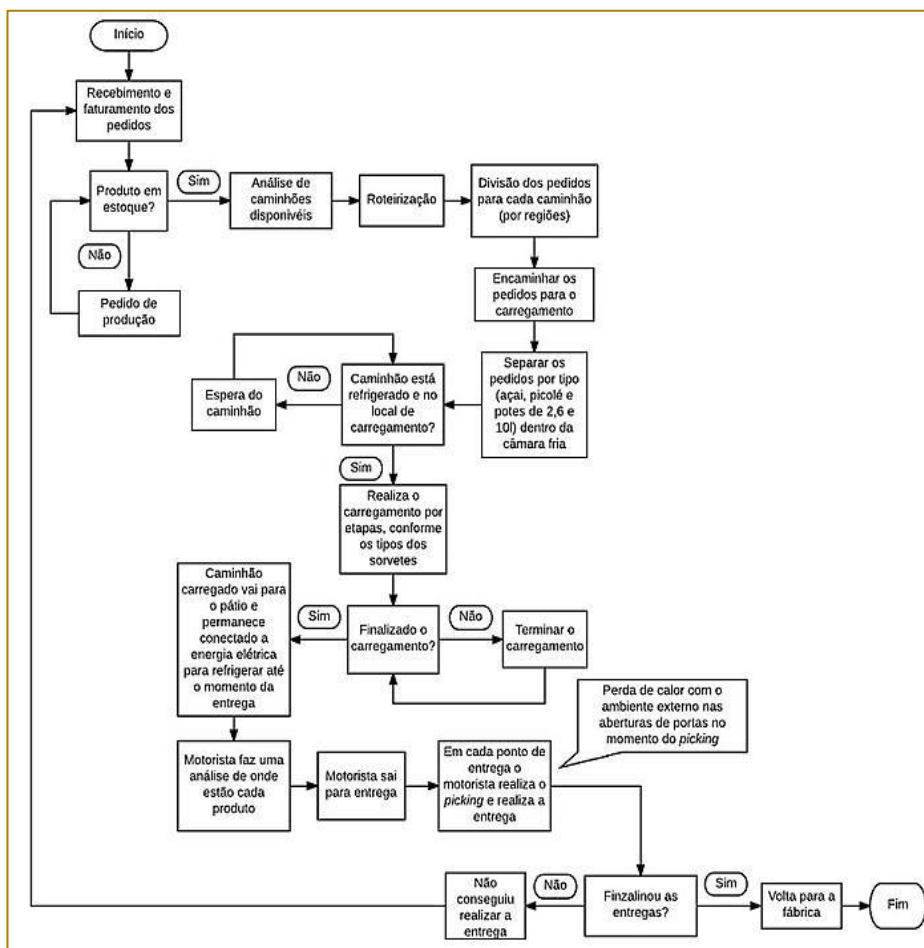
divisórias do tipo sorveteiro, sendo assim, os produtos não ficam organizados e separados

por pedidos de cada cliente e sim separados por sabores. Cada motorista tem a obrigação de identificar onde está cada produto antes de sair para a entrega, e o *picking* é feito no momento da entrega.

A figura 8 mostra um fluxograma simplificado do processo de distribuição de sorvetes,

desde o momento de recebimento do pedido até a entrega do mesmo. Os pedidos são faturados no final do dia anterior à entrega, sendo assim os processos de carregamento do caminhão são realizados no período da madrugada, e após finalizados, o caminhão permanece no pátio refrigerando até o momento de entrega.

Figura 8 - Fluxograma do processo desde o pedido até a entrega.



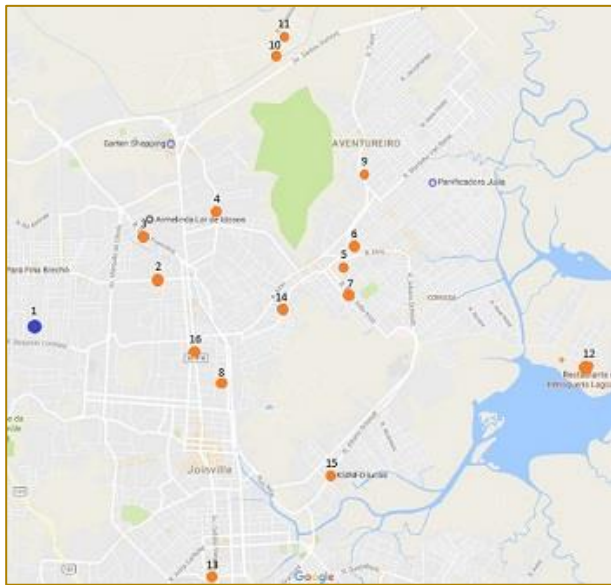
Fonte: Autoria própria.

Através de um acompanhamento de entrega realizado no final do mês de abril de 2017 na própria cidade de Joinville, foram anotados os pontos de entrega e os tempos de atendimento (TA- tempo em que a porta do compartimento refrigerado fica aberta), para que através do método descrito anteriormente fosse estimada uma melhor rota entre os métodos propostos por Moresco e Silva (2017). Foram anotados 15 pontos de entrega, todos inseridos na cidade de Joinville. A empresa realiza uma divisão geográfica para

a realização da distribuição, ou seja, um veículo deixa a fábrica com o foco de atender uma determinada região. Portanto, normalmente saem no mínimo 5 caminhões diários para as regiões de Blumenau, Itajaí, Balneário Camboriú, Curitiba e Joinville. Também, a partir da fábrica, saem os veículos direcionados a atender os CD's.

A figura 9 mostra os pontos de entregas (em laranja) e a fábrica (em azul), e estão numeradas para melhor identificação após a escolha da melhor rota.

Figura 9 – Distribuição dos pontos de entrega.



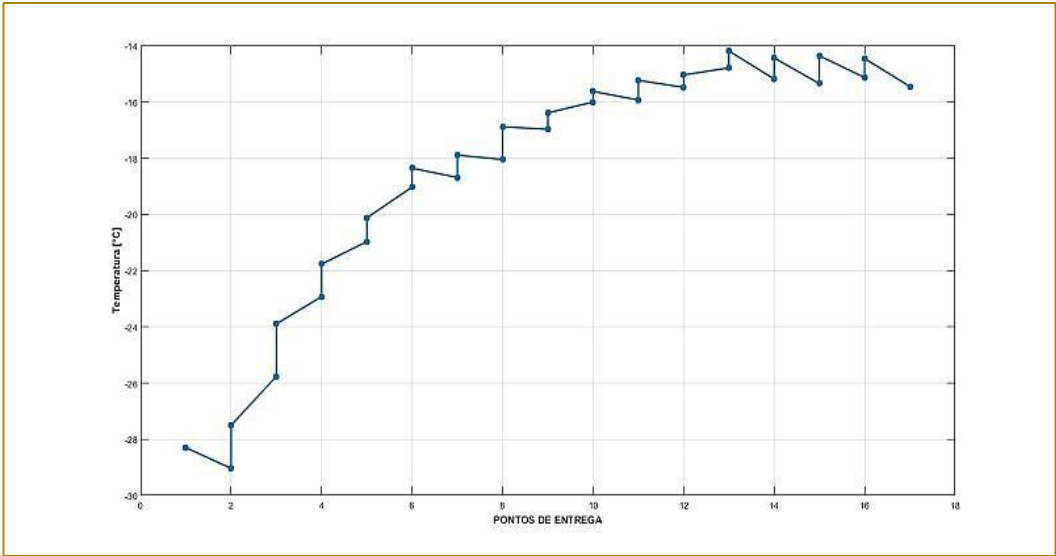
Fonte: Google Maps.

Portanto, através do *software* MATLAB e do algoritmo do problema do vizinho mais próximo aplicado por Moresco e Silva (2017), o menor caminho percorrido foi pelo método 3 descrito anteriormente, e a distância total foi de 53,95 quilômetros. A ordem dos clientes atendidos foi: 3 – 2 – 4 – 10 – 11 – 9 – 6 – 5 – 7 – 14 – 8 – 16 – 13 – 15 – 12. Assim, sabendo o sequenciamento de entregas, e os TA's, o maior TA foi no ponto 2 com 21 min e o menor

TA foi de 4,72 min do ponto 7. No início, o caminhão sai da fábrica com temperatura inicial do compartimento de carga refrigerado -28,3°C, momento em que é desconectado da energia elétrica, podendo-se presumir o comportamento da temperatura final do compartimento interno do caminhão.

A figura 10 mostra a variação da temperatura em relação aos pontos de entrega, no decorrer de sua trajetória de entrega.

Figura 10 – Comportamento da temperatura em relação aos pontos de entrega.



Fonte: Autoria própria.

A figura 10 mostra o quanto o aumento de temperatura está diretamente ligado ao TA, sendo que o ponto 2 representa a primeira entrega da rota, e assim, o ponto 3 da figura representa a segunda entrega da rota (ponto 2), com o maior tempo de atendimento e assim o maior incremento de temperatura, e conseqüentemente o ponto 10 do gráfico, representa a nona entrega e trata do ponto 7 com o menor TA e o menor aumento de temperatura.

Portanto na figura 10, a maior taxa de transferência de calor ocorre quando a diferença entre a temperatura externa e interna é elevada. Esta maior taxa de transferência de calor provoca um maior aumento da temperatura interna do compartimento, que reduz gradativamente conforme a temperatura interna é elevada.

No estudo de caso, a máxima temperatura que o caminhão atingiu foi de  $-14,2^{\circ}\text{C}$ , mas conforme Bragante (2010) o ideal é manter o produto o mais próximo de  $-25^{\circ}\text{C}$  para evitar o derretimento e a má aparência do produto. Porém os motoristas realizam a entrega e somente é dada a perda da carga se o compartimento refrigerado se aproximar de um valor de temperatura positivo. É válido comentar que não foi medida a temperatura do produto no interior do veículo e sim somente a temperatura do veículo. Para melhor análise, sugere-se investigar o comportamento da variação de temperatura dos produtos transportados durante a distribuição física, para então se avaliar a qualidade dos mesmos.

## 6. CONSIDERAÇÕES FINAIS E SUGESTÕES

O referente trabalho permitiu verificar a existência de ferramentas disponíveis e acessíveis para auxiliar na análise prévia do comportamento de temperatura de produtos da Cadeia do Frio, fator muito importante para a redução de perdas de cargas quando o assunto é relacionado a produtos perecíveis.

A simulação realizada através do método desenvolvido por Moresco e Silva (2017) leva em conta somente a distância entre os pontos de entrega, sem uma restrição de ordem de preferência. Porém conforme divulgado pela empresa, alguns clientes possuem horários agendados e pré-estabelecidos para o recebimento do produto, assim sendo necessário um acréscimo de uma restrição de ordem das entregas. Portanto, para estudos futuros, faz-se necessário a utilização de uma extensão do problema do caixeiro viajante, o VRPTW (vehicle routing problem with time windows).

Ainda para estudos futuros, sugere-se que seja feita uma análise do volume de produtos entregues para cada cliente e o TA, fazendo-se uma regressão linear para uma possível análise prévia do TA, para posterior análise do comportamento da temperatura em diferentes rotas.

Como descrito anteriormente, o *picking* dos produtos é realizado pelos motoristas dos caminhões, no momento da entrega aos clientes, portanto para estudos futuros, sugere-se também a utilização do problema da mochila (knapsack problem), para facilitar a entrega do produto e assim diminuir TA, fazendo com que a organização dos produtos no caminhão seja por pedidos, necessitando somente um momento de abertura de porta.

Assim, entende-se que o estudo realizado serve como uma ferramenta de apoio à análise prévia do comportamento da temperatura dos caminhões em relação à rota escolhida para a distribuição de seus produtos. Isto permite analisar o cenário antes mesmo de efetuar-lo, contribuindo para a redução de perdas inesperadas de produtos.

Por fim, esse trabalho agregou conhecimento sobre ferramentas de roteirização e permitiu à empresa em estudo, um entendimento melhor do comportamento da temperatura nos momentos de entrega.

## REFERÊNCIAS

- [1]. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 14701*: Transporte de produtos alimentícios refrigerados: procedimentos e critérios de temperatura. Rio de Janeiro: ABNT, 2001.
- [2]. BALLOU, Ronald H. Gerenciamento da Cadeia de Suprimentos: Planejamento, organização e logística empresarial. 4ª Edição. Porto Alegre: Bookman, 2001.
- [3]. BAPTISTA, Paulo. Higiene e segurança alimentar no transporte de produtos alimentares. Forvisão-Consultoria em Formação Integrada, Guimarães, Portugal, 2006.
- [4]. BRAGANTE, A. G. Tecnologia de Fabricação de Sorvetes. 2010. Disponível em: <[http://abgtecalim.yolasite.com/resources/Tecnologia da Fabricação de Sorvetes.pdf](http://abgtecalim.yolasite.com/resources/Tecnologia%20da%20Fabricação%20de%20Sorvetes.pdf)>. Acesso em: 27 jun. 2015.
- [5]. CECHIN, Adelmo Luis. Comparação entre Métodos Exatos e Heurísticos para tratar o Problema de Roteamento de Veículos em um Ambiente Fabril. 2006.
- [6]. HEAP, Robert. Cold chain performance issues now and in the future. Bulletin of the IIR, 4. 2006.
- [7]. HILLIER, Frederick S., LIEBERMAN, Gerald J. Introdução à pesquisa operacional. McGraw Hill – 9. Ed. – Porto Alegre: AMGH. 2013.
- [8]. LIKAR, Kristina. JEVSNIK, Mojca. Cold chain maintaining in food trade. Food Control, v. 17, n. 108, p. e113. 2006.
- [9]. LUNA, Henrique Pacca; GOLDBARG, Marco Cesar. Otimização combinatória e programação linear. Rio de Janeiro: Campus, 2000.
- [10]. MORESCO, Gustavo Henrique. SILVA, Vanina M. D. Cold chain mapping: concepts and applications. XXVII Congresso Regional de Iniciação Científica e Tecnológica em Engenharia. Joinville – Santa Catarina, 2016.
- [11]. MORESCO, Gustavo Henrique; SILVA, Vanina Macowski Durski. Cadeia do frio: Estudo de caso e aplicação de método heurístico para roteirização de entrega. Em: AHRENS, Rudy de Barros. Coletânea nacional sobre engenharia de produção 3: gestão da produção. Curitiba: Atena Editora, 2017. Cap. 14. P. 192-205.
- [12]. NOVAES, Antônio G. N. et al. Thermal performance of refrigerated vehicles in the distribution of perishable food. Pesquisa Operacional, v. 35, n. 2, p. 251-284, 2015.
- [13]. PEREIRA, Vítor F. et al. Avaliação de temperaturas em câmaras frigoríficas de transporte urbano de alimentos resfriados e congelados. Ciênc. Technol. Aliment, v. 30, n. 1, p. 158-165, 2010.
- [14]. SILVA, Gerson B. A Gestão da Cadeia do Frio: uma análise de fatores logísticos. Brasil: Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, 2010.
- [15]. SOARES, Antonio Gomes. Perdas pós-colheita de frutas e hortaliças. Fórum Agronegócios da UNICAMP. Qualidade e Segurança de Alimentos. Mesa Redonda - Qual o Tamanho do Desperdício. Universidade Estadual de Campinas. Campinas, 2009.

# Capítulo 22

## APLICAÇÃO DA ABORDAGEM ESTOCÁSTICA BASEADA EM SIMULAÇÃO DE MONTE CARLO [SMC]: UM ESTUDO DE CASO

*Paulo Henrique dos Santos*

*Solange da Silva*

*Juliano Resende Bucchianeri*

*Nilson Carvalho da Mata*

*Thiago Moura de Moraes*

**Resumo:** O artigo diz respeito à aplicação e benefícios da Simulação de Monte Carlo [SMC] proposto para análise das incertezas e riscos de resultados de determinada operação no mercado de mini contratos futuros. O *software* Crystal Ball® [CB] é um programa de simulação que ajuda a analisar as incertezas associadas a modelos. O principal objetivo do artigo é analisar as incertezas e o risco considerado nos resultados de um estudo de caso envolvendo mini contratos futuros na modalidade *day trade*. Os resultados da simulação [10.000 replicações] consistem em gráficos de frequência e relatórios estatísticos. Os resultados deste estudo podem ser utilizados na realização de uma análise de ganhos financeiros relacionados a operações *day trade* de diversas origens. Concluiu-se que a abordagem estocástica é um método poderoso para a quantificação do parâmetro de incerteza, podendo ajudar os profissionais e tomadores de decisão, na gestão de suas diversas aplicações.

**Palavras Chave:** Simulação de Monte Carlo [SMC]; Crystal Ball® [CB]; Mini contrato futuro.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)



## 1. INTRODUÇÃO

A simulação e as estatísticas têm a maior taxa de aplicação sobre todas as outras ferramentas. Os benefícios de uma abordagem de modelagem de simulação são: [I] a compreensão da probabilidade de resultados específicos, [II] a capacidade de identificar e examinar a condução de variáveis dentro de um modelo, [III] um modelo muito mais flexível e [V] tabelas resumo claro e relatórios. A simulação é particularmente útil quando os problemas exibem uma significativa incerteza, que é geralmente difícil de tratar analiticamente [Bieda, 2013].

O foco do artigo diz respeito à aplicação e benefícios da Simulação de Monte Carlo [SMC] proposto para estimar o ganho financeiro de determinada operação financeira, analisando suas incertezas e riscos. Para conhecimento, não há SMC na literatura até o momento sobre este tipo de resultado obtido de uma operação empírica semelhante em mini contratos futuros.

O principal objetivo do artigo é analisar as incertezas e o risco considerado nos resultados de um estudo de caso envolvendo mini contratos futuros na modalidade *day trade*, encontrando o percentual de risco considerado após a execução do *software* Crystal Ball® [CB].

O *software* Crystal Ball® [CB], programa de simulação que ajuda a analisar as incertezas associadas a modelos Microsoft® Excel por SMC, foi proposto para calcular os resultados das operações *day trade* em mini contratos futuros.

A Seção 2 traz conteúdos essenciais sobre a SMC. Na seção 3 haverá informações, as quais originaram os dados coletados. A seção 4 traz o método utilizado. O experimento e resultados obtidos pela computação são mostrados na Seção 5. A conclusão das observações é apresentada na seção 6..

## 2. PERMISSÕES DA SIMULAÇÃO DE MONTE CARLO [SMC]

Os modelos de previsões baseadas em SMC são capazes de lidar com a incerteza significativa nas entradas da dinâmica complicada de condições durante a tomada de decisão [CAI *et al.*, 2015], neste estudo de caso, as decisões relacionadas ao *setup*.

A SMC é uma das ferramentas estocásticas mais generalizadas que analisa incertezas do

modelo [Bieda, 2013]. Uma das aplicações de SMC é a obtenção de médias ao longo de um estatístico conjunto canônico [Bandyopadhyay, 2013]. Assumindo a definição estocástica dos dados do processo, pode-se avançar a pesquisa, que propõe o uso de técnicas estatísticas e simulação de Monte Carlo para estimar as probabilidades [ACEBES *et al.*, 2015].

Monte Carlo é obtível dentro de uma quantidade bastante razoável de esforço computacional. SMC é um método usado bem conhecido para determinar probabilidades. Resolver um problema de programação dinâmica por simulação não é uma novidade [CARDOSO, *et al.*, 2013]. No entanto, enquanto amplamente utilizado em muitos contextos, não têm sido utilizadas SMCs na resolução de problemas de risco em análise de operações *day trade* de mini contratos futuros. O alvo de avaliação é o valor esperado do resultado financeiro alcançado probabilisticamente interpretado de maneira simples [CAI *et al.*, 2015].

A ferramenta de SMC permite considerar um comportamento flutuante para a demanda, mas também vários cenários de correlação entre os dados [VALLÉE *et al.*, 2013]. A análise financeira do caso de aplicação é capaz de demonstrar que a SMC gera informações substanciais sobre o risco do projeto [Arnold; Yildiz, 2015], neste caso o *setup*.

SMC é um modelo estocástico que utiliza geração de números aleatórios para selecionar valores para os parâmetros [SOUZA *et al.*, 2015]. Assim, é uma técnica que é bem adequada para elucidar os efeitos do parâmetro de incerteza. Também poderia ser usada para explorar os efeitos do modelo e o cenário de incerteza [HUDSON; TILLEY, 2014].

Permite resolver problemas com a simulação de cenários e o posterior cálculo de um valor esperado, possibilitando a implantação de hipóteses adicionais nas previsões. Os experimentos da SMC são excelentes para destacar o conceito de amostras repetidas, a base da maior parte da inferência estatística clássica [SOUZA *et al.*, 2015].

### 2.1. APLICAÇÃO DA SMC EM ÁREAS VARIADAS

Kasperski *et al.* [2015] utilizaram a SMC para demonstrar como a morfologia dos

agrupamentos supramoleculares podem ser ajustados por meio da manipulação de parâmetros intrínsecos dos blocos de construção e modificada pela presença de átomos de metal adsorvidos. Usando a SMC, conseguiu a ativação de dois ou mais segmentos terminais, que levam à formação de nanoestruturas variando isoladas. Permitiu-se a identificação das principais diferenças estruturais, encontrando a melhor escolha para obter a estrutura alvo.

A abordagem de Acebes *et al.* [2015] usa a SMC para obter informações sobre o comportamento esperado do projeto. Explorando diversas maneiras, usando diferentes metodologias de aprendizagem estatística de forma estruturada. A abordagem permitiu detectar situações de anomalias no que diz respeito a definição do projeto, tendo em conta a possível correlação entre tempo e custo que as metodologias anteriores ignoram.

Arnold e Yildiz [2015] usaram a SMC para análise de risco com base em toda uma representação do ciclo de vida de projetos sobre tecnologias de energias renováveis descentralizadas. A SMC ofereceu consideráveis vantagens metodológicas, permitindo a variação de vários parâmetros de entrada difusos ou incertos na avaliação financeira de um projeto que reflete o fato de que as previsões estão sempre expostas a algum grau de incerteza. As probabilidades de ser menor ou maior do que um determinado valor alvo ou de estar dentro de um intervalo de confiança pode ser determinada pela SMC. Portanto, possibilitou aos investidores avaliarem o risco e a volatilidade de um investimento, projetando uma infraestrutura com mais facilidade. A SMC auxilia na otimização conceitual da concepção de um projeto de investimento em relação aos retornos de capital e segurança. Uma vez que ambas as questões são decisivas para credores e investidores.

Na busca empírica sobre informações dos mini contratos do índice Bovespa, não se obteve sucesso em publicações relevantes, portanto este artigo irá explorar com maior veemência a SMC e as análises extraídas do CB. Aproveita-se o estudo real dos dados

sobre os contratos de mini índices, mantendo o caráter inédito da pesquisa e demonstrando as vantagens de se trabalhar com a simulação antes de iniciar um investimento financeiro.

### 3. ESTUDO DE CASO SOBRE A OPERAÇÃO DOS MINI CONTRATOS FUTUROS

Segundo a BMF Bovespa®, mini contrato futuro é um indicador do desempenho médio, expressado por quantidade de pontos, das cotações dos ativos de maior negociabilidade e representatividade do mercado de ações brasileiro, busca viabilizar que os investidores, em especial, pessoas físicas e pequenas empresas, possam iniciar suas atividades no mercado de derivativos.

A operação de compra e venda dos mini contratos é realizado através do *Home Broker*. A operação que originou os dados deste artigo se chama *day trade*, ou seja, operações de compra e venda de contratos futuros no mesmo dia, possibilitando iniciar uma operação de venda antes da compra e vice-versa.

Neste estudo de caso real, o investidor iniciou suas operações baseadas em uma maneira empírica de analisar suas compras e suas vendas. Ele realiza duas compras e duas vendas ou vice-versa por dia.

A primeira decisão de compra ou venda acontece às 10:59:59 do dia, após entrar na posição neste horário, o investidor determina o fechamento da posição às 16:59:59. O investidor estipula uma regra básica em sua maneira de operar, nesta posição, que se pode chamar de P1, ele opera com uma perda máxima no dia de 500 pontos e um ganho ilimitado.

A segunda decisão de compra ou venda acontece às 17:00:00, ele abre uma nova posição, chamada de P2, com um limite de perda máximo de 100 pontos e ganho ilimitado.

Na Figura 1 é mostrado o gráfico que o investidor em questão utiliza para determinar suas compras ou vendas dos dias.

Figura 1: Gráfico de velas de mini contratos futuros.



Fonte: [BMF Bovespa®, WINFUT 60 minutos, dia 24 de Outubro de 2016]

O gráfico utilizado, é chamado pelos *traders* de “gráfico de velas”, também denominado pela corretora. São gráficos que medem a quantidade de pontos em determinado tempo, este, por exemplo, está medindo os pontos de 60 em 60 minutos, iniciando às 09:00:00 e encerrando às 18:00:00. O dia do exemplo da Figura 1 foi 24/10/2016. Nele se observa o momento da abertura de cada hora do gráfico, o valor máximo e o valor mínimo de cada hora e seu momento de fechamento. A vela vermelha representa uma queda nos pontos, por exemplo, a vela das 16:00:00 às 17:00:00 identificada na Figura 1, abriu com um valor de 65.315 pontos, chegou a 65.335 pontos, obteve um mínimo de 65.005 pontos e fechou em 65.095 pontos, como sua abertura foi maior que seu fechamento, quer dizer que das 16:00:00 até às 17:00:00 o mercado obteve um perda de 220 pontos. A vela verde, representa um aumento dos pontos, por exemplo, a vela das 10:00:00 às 11:00:00 identificada na Figura 1, abriu em 65.530 pontos, obteve uma máxima de 65.810 pontos, chegou a uma mínima de 65.500 pontos e fechou em 65.800 pontos, caracterizando em um ganho no mercado de 270 pontos.

A posição P1 é aberta em compra, se caso o mercado atingir igual ou maior que 65 pontos acima do ponto máximo da vela de 10:00:00 às 11:00:00, por exemplo, na vela da Figura 1, o ponto máximo é de 65.810 pontos, então o ponto de compra seria quando o mercado atingir 65.875 pontos. A abertura do P1 em venda acontece da mesma forma mas agora obtendo o mínimo como parâmetro, exemplo, como o mínimo foi 65.500 pontos, há uma subtração de 65 pontos, então se o mercado atingir 65.435 se abre a posição em venda. No exemplo da Figura 1 aconteceu uma venda.

Na posição P2 a operação é diferente, imediatamente às 17:00:00 em ponto, assim que a posição P1 é fechada, a posição P2 é aberta. A posição P2 observa a cor da vela das 16:00:00 às 17:00:00, se estiver em queda, imediatamente se inicia uma venda, se estiver em crescimento, imediatamente se inicia uma compra. Os limites de compra e venda de cada posição já foram especificados anteriormente.

Os resultados referentes às operações realizadas no período, foram expressados no Quadro 1, são os quocientes mensais em pontos da soma dos resultados das posições diárias de P1 e P2.

Quadro 1: Resultados em pontos, obtidos pelas operações.

|    | A      | B         | C      | D         | E      | F         | G      | H         |
|----|--------|-----------|--------|-----------|--------|-----------|--------|-----------|
| 1  | Data   | Resultado | Data   | Resultado | Data   | Resultado | Data   | Resultado |
| 2  | -      | -         | jan/13 | 2020      | jan/14 | 1590      | jan/15 | -990      |
| 3  | -      | -         | fev/13 | 1405      | fev/14 | 2445      | fev/15 | -225      |
| 4  | -      | -         | mar/13 | -920      | mar/14 | 3395      | mar/15 | 3240      |
| 5  | -      | -         | abr/13 | 1320      | abr/14 | 2885      | abr/15 | 935       |
| 6  | -      | -         | mai/13 | 460       | mai/14 | 1690      | mai/15 | 1610      |
| 7  | -      | -         | jun/13 | -1390     | jun/14 | 2510      | jun/15 | -600      |
| 8  | jul/12 | 2300      | jul/13 | -65       | jul/14 | -220      | jul/15 | 1365      |
| 9  | ago/12 | -745      | ago/13 | 1125      | ago/14 | 600       | ago/15 | 3170      |
| 10 | set/12 | 790       | set/13 | 2500      | set/14 | 1015      | set/15 | 1685      |
| 11 | out/12 | 1820      | out/13 | 2430      | out/14 | 2645      | out/15 | 1275      |
| 12 | nov/12 | 3385      | nov/13 | 320       | nov/14 | -650      | nov/15 | 385       |
| 13 | dez/12 | 905       | dez/13 | 790       | dez/14 | -165      | dez/15 | 735       |

Fonte: [Elaborado pelo autor]

O Quadro 1 mostra os ganhos e perdas das operações realizadas pelo investidor entre o período de Julho de 2012 à Dezembro de 2015. Os resultados estão em pontos e identificados respectivamente por cada mês de operação.

#### 4. MÉTODO

Na pesquisa foi aplicado o método descritivo, através da estatística descritiva dedicada a um estudo de caso empírico. A metodologia adotada neste presente estudo é fidedigna para alcançar os resultados, ajudando a determinar a decisão sobre as incertezas dos quocientes do estudo de caso de uma operação em mini contratos futuros utilizando a SMC, ou seja, o método empregado é capaz de conduzir a pesquisa para o alcance dos objetivos estabelecidos.

O estudo também mede o valor em risco referente aos resultados, analisa o comportamento do modelo e fornece uma análise de sensibilidade. Os dados foram obtidos através de um investidor, que os disponibilizou para pesquisa, incluindo os resultados de suas operações de Julho/2012 à Dezembro/2015.

O investidor faz operações *day trade* desde 2012, mas não se embasa em nenhum estudo científico para realizar suas operações. Suas decisões são tomadas baseando-se no histórico de médias de ganhos ou perdas dos rompimentos para mais ou para menos das velas de 60 minutos mostradas na Figura 1, já exemplificados neste artigo. Ele notou que os rompimentos das velas de 10:00:00 às 11:00:00 com determinadas regras, se obtém ganhos a longo prazo.

Este estudo de caso se mostra como boa oportunidade para aplicabilidade da SMC para o alcance da dimensão do risco e

incerteza da operação financeira. A análise probabilística utilizando o método de MC e a construção de modelos usando CB, consiste nos seguintes passos [Bieda, 2013]:

- Desenvolver o modelo de planilha;
- Definir o pressuposto para variáveis probabilísticas;
- Definir células de previsão, ou seja, as variáveis de interesse de saída;
- Definir um número de replicação;
- Executar a simulação.

#### 5. EXPERIMENTO, DISCUSSÃO E RESULTADOS

A princípio foi verificada a normalidade dos dados coletados do estudo de caso. Spivey *et al.* [2013] indicam que para um melhor ajuste do modelo simulado, a normalidade dos dados é importante. Em seu estudo ele indica o teste de Shapiro-Wilk, que pode ser realizado para verificar se o conjunto de dados é normalmente distribuído. No teste de Shapiro-Wilk a hipótese nula é que as amostras seguem a distribuição normal.

O teste de Shapiro-Wilk calcula uma estatística de teste que corresponde a alguns valores de p. Um valor-p é a probabilidade de obter essa estatística de teste de valor, assumindo que a hipótese nula, da distribuição normal, é verdade. Dependendo do nível de significância estatística desejada, um nível de  $\alpha$  é escolhido. E se o teste de Shapiro-Wilk der um valor de p menor do que o nível de  $\alpha$  escolhido em seguida, a hipótese nula dos dados distribuídos normalmente é rejeitada. Isto significa que o desvio da distribuição normal é significativo o suficiente para justificar a sua rejeição.

Este teste envolve *software* Action Stat juntamente com o Excel. No teste de

normalidade Shapiro-Wilk realizado com os dados deste estudo de caso, tendo como valor de  $\alpha = 5\%$ , o resultado do valor-p é dado por 0,4395 ou 43,95%. Portanto, a hipótese nula não é rejeitada e amostra dos dados é considerada normal.

### 5.1. DESENVOLVIMENTO DO MODELO DE PLANILHA

Partindo dos resultados apresentados no Quadro 1, foi construído o Quadro 2, neste quadro foram trazidas as variáveis determinísticas e probabilísticas de entrada da simulação, para tal análise.

Nas linhas Custos, todos sendo valores determinísticos, está o custo de custódia, este custo é cobrado pela empresa responsável

pelo *home broker* a cada mês pela custódia do dinheiro, também consta o custo de corretagem, o custo de corretagem representa o custo por cada abertura de posição de compra ou de venda, sendo o valor de R\$ 0,35 por contrato. O investidor tem operado com 10 contratos desde Julho/2012, gerando um custo diário igual a  $0,35$  [corretagem/contrato]  $\times$   $10$  [quantidade de contratos] =  $3,50$   $\times$   $4$  [quantidade de abertura de posições/dia] = R\$ 14,00/dia. Chegou-se a o valor de R\$ 308,00, através da multiplicação da corretagem/dia pela quantidade considerada de 22 dias úteis dentro do mês.

Quadro 2: Tabela com os dados a serem simulados.

|   | A                     | B           | C              |
|---|-----------------------|-------------|----------------|
| 1 | <b>Custos</b>         |             |                |
| 2 | Custo de custódia     | R\$ 6,90    | Determinístico |
| 3 | Custo de corretagem   | R\$ 308,00  | Determinístico |
| 4 | Custo total           | R\$ 314,90  | Determinístico |
| 5 | <b>Ganhos/Perdas</b>  |             |                |
| 6 | Ganho/Perda em ponto  | 0           | Normal         |
| 7 | Ganho/Perda por ponto | R\$ 2,00    | Determinístico |
| 8 | <b>Apuração</b>       |             |                |
| 9 | Lucro (previsão)      | -R\$ 314,90 | Previsão       |

Fonte: [Elaborado pelo autor]

Nas linhas de Ganhos/Perdas, primeiramente estão os ganhos/perdas, em valores probabilísticos, relacionados a quantidade de pontos alcançados por mês. Todavia, para este exemplo será utilizada apenas a distribuição normal, já verificada no início desta seção, sobre os dados históricos das operações do investidor. De forma determinística se traz o ganho/perda por ponto, sendo o valor de R\$ 0,20 [ganho/perda por contrato]  $\times$   $10$  [quantidade de contratos] = R\$ 2,00/ponto.

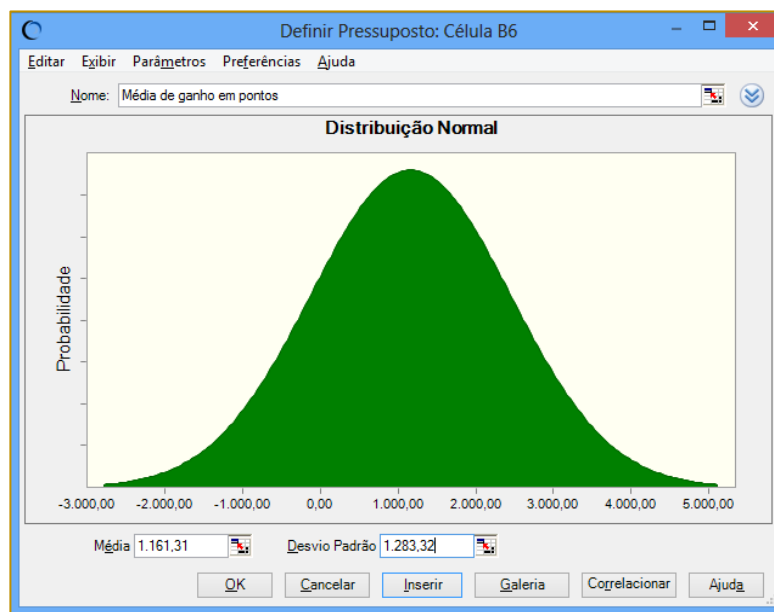
Na linha de apuração, o Lucro representa o ganho/perda por ponto  $\times$  o ganho/perda em pontos, menos o custo total. As etapas apresentadas no Quadro 2 são muito importantes, pois toda modelagem usará as variáveis identificadas. Se houver algum erro nos dados, poderá comprometer a análise e o resultado do experimento, ocasionando em decisões incertas a respeito da simulação.

### 5.2. DEFINIÇÃO DOS PRESSUPOSTOS PARA VARIÁVEIS PROBABILÍSTICAS

As incertezas neste caso estão relacionadas às quantidades de pontos ganhos ou perdidos nas operações do investidor, podendo determinar futuras perdas ou ganhos financeiros utilizando seu método empírico de operação.

A variável probabilística está localizada na coluna "B" linha "6" identificada como pressuposto da análise no Quadro 2. O pressuposto para a variável probabilística de entrada foi definido utilizando a distribuição normal. A definição do pressuposto para uma distribuição normal exige a inclusão do desvio padrão e a média dos dados, neste caso, os dados do Quadro 1, utilizando a fórmula =DESVPAD.P[B8:B13;D2:D13;F2:F13;H2:H13] no Microsoft Excel obtêm-se o valor de 1283,318828 como desvio padrão e utilizando a fórmula =MÉDIA[H2:H13;F2:F13;D2:D13;B8:B13] obteve-se o valor de 1161,309524, estas entradas no *software* CB, são mostradas na Figura 2.

Figura 2: Configuração do pressuposto para variável de entrada.



Fonte: [Crystal Ball®, 2017]

A abordagem estatística para a incerteza cada vez se torna mais semelhante aos eventos ambientais. Técnicas probabilísticas, como a SMC, toma a relevância ambiental dos pressupostos em conta, e, finalmente, fazer a interpretação clara. Entre as ferramentas probabilísticas, a análise SMC tem-se aumentado nos últimos anos [BIEDA, 2013].

### 5.3. DEFINIÇÃO DA CÉLULA DE PREVISÃO, OU SEJA, AS VARIÁVEIS DE INTERESSE DE SAÍDA

A variável de interesse de saída que o *software* dá-se o nome de “Previsão”, representa a apuração dos resultados identificada como lucro, sua fórmula no Microsoft Excel é descrita por = [B6\*B7]-B4, que seria o ganho/perda por ponto x o ganho/perda em pontos, menos o custo total, servindo para averiguar os valores da simulação. A unidade de valor de saída da simulação escolhida foi expressa em R\$.

### 5.4. DEFINIÇÃO DO NÚMERO DE REPLICAÇÕES

Souza *et al.* [2015], argumentam que com o *software* CB pode realizar cerca de 10.000 replicações de forma simples, coisa que

antigamente somente era possível com supercomputadores.

Sabbag e Costa [2015] utilizam 10.000 replicações em seu estudo, envolvendo os custos da produção de leite. Noronha *et al.* [2014] utilizam da mesma forma 10.000 replicações, onde se estuda a gestão do processo de desenvolvimento de produtos em uma indústria de autopeças. Oliveira e Neto [2012] trazem em sua pesquisa a avaliação da pertinência da utilização da técnica de SMC na mensuração das incertezas inerentes à metodologia de avaliação de empresas pelo fluxo de caixa descontado, de maneira semelhante usou-se 10.000 replicações.

Pegando como base o apoio teórico de outras pesquisas, neste artigo também se realizou 10.000 replicações da simulação. Segundo Arnold e Yildiz [2015], o número de replicações pode interferir nos resultados da simulação, pressupondo que um número maior de replicações se alcança resultados mais estáveis.

### 5.5. EXECUTANDO A SIMULAÇÃO

Na Figura 3, após a execução da simulação é possível observar os seguintes dados relacionados aos valores de previsão: a média, mediana, desvio padrão, variância, erro padrão da média, dentre outros resultados observados.

Figura 3: Estatísticas da simulação realizada.

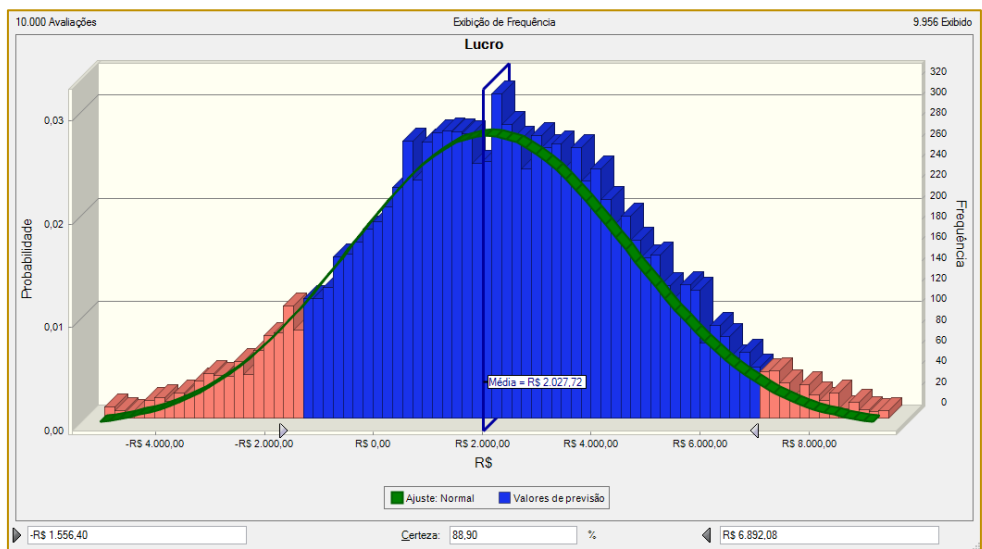
| 10.000 Avaliações        |                  |                     |
|--------------------------|------------------|---------------------|
| Estatística              | Ajuste: Normal   | Valores de previsão |
| ▶ Avaliações             | ---              | 10.000              |
| Caso Base                | ---              | -R\$ 314,90         |
| Média                    | R\$ 2.027,72     | R\$ 2.027,72        |
| Mediana                  | R\$ 2.027,72     | R\$ 2.031,18        |
| Moda                     | R\$ 2.027,72     | ---                 |
| Desvio Padrão            | R\$ 2.572,12     | R\$ 2.572,12        |
| Variância                | R\$ 6.615.808,99 | R\$ 6.615.808,99    |
| Obliquidade              | 0,00             | -0,0151             |
| Curtose                  | 3,00             | 2,98                |
| Coefficiente de Variação | 1,27             | 1,27                |
| Mínimo                   | -Infinito        | -R\$ 8.005,50       |
| Máximo                   | +Infinito        | R\$ 11.199,59       |
| Erro Padrão Média        | ---              | R\$ 25,72           |

Fonte: [Crystal Ball®, 2017]

Posteriormente a execução da simulação, observa-se na Figura 4 que a média é de R\$ 2.027,72. Como já mostrado na Figura 2, uma média de pontos de 1.161,31 e no Quadro 2 o valor de R\$ 2,00 por cada ponto ganho ou perdido, chega-se ao valor de R\$ 2.322,62, que representa a média de pontos vezes o valor por ponto, mostrando que o valor do histórico dos dados é maior do que o resultado da simulação.

Utilizando o mesmo cálculo do parágrafo anterior, pode-se comparar o valor mínimo do histórico com o valor mínimo do resultado simulado, o histórico representa R\$ - 2.780,00 e o simulado representa R\$ - 8.005,50. Já o valor máximo do histórico é de R\$ 6.790,00 e o simulado de R\$ 11.199,59. Através deste resultado, mostra-se que na simulação há uma variabilidade dos dados maior do que no histórico.

Figura 4: Exibição do gráfico de frequência de previsão de lucro.



Fonte: [Crystal Ball®, 2017]

Observando a Figura 4, conclui-se que há uma certeza de 88,90% que este tipo de operação realizada pelo investidor do estudo de caso terá seu resultado financeiro mensal entre R\$ - 1.556,40 e R\$ 6.892,08. Mostra-se

uma um grau de certeza elevado para estes resultados. O CB possibilita a definição de intervalos para a verificação do percentual de certeza, este recurso pode ser usado para verificar o

percentual de certeza interessante da operação.

## 5.6. ANÁLISE DO RISCO

Modelos de previsões baseadas em SMC são capazes de lidar com incertezas significativas. De forma que os resultados podem ser probabilisticamente interpretados de maneira simples [CAI *et al.*, 2015].

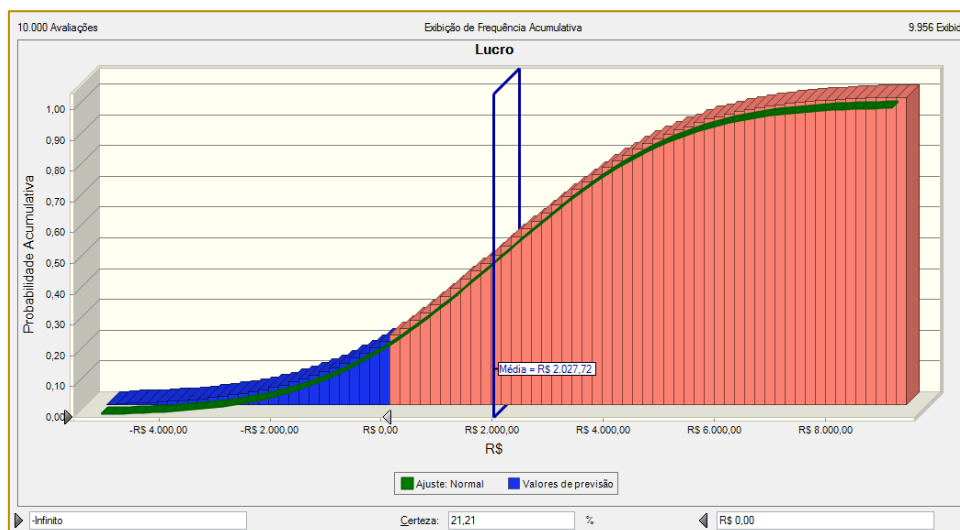
De acordo com Souza *et al.* [2015] a SMC pode ser utilizada para a avaliação dos riscos envolvidos expressos de forma simples e de fácil entendimento para auxílio nas decisões. Obtendo desta forma, indicadores estocásticos probabilísticos ao invés de indicadores determinísticos.

No estudo de Rocha *et al.* [2016], explanam-se que os resultados da análise de simulação são fortemente dependentes das premissas feitas para a variável de risco. Em sua pesquisa foram construídos três cenários: pessimista, realista e otimista.

Neste estudo de caso, considera-se o cenário realista, a análise já feita sobre os resultados da Figura 4, restando fazer a análise pessimista e otimista dos resultados.

A análise pessimista é mostrada na Figura 5, onde se mostra que há uma obtenção de certeza de 21,21% que os resultados da operação do investidor ficarão entre R\$ 0,00 e R\$ - 8.005,50. Este grau de certeza se mostra baixo para este cenário pessimista, mesma havendo o risco do prejuízo.

Figura 5: Exibição do gráfico de frequência de previsão pessimista.



Fonte: [Crystal Ball®, 2017]

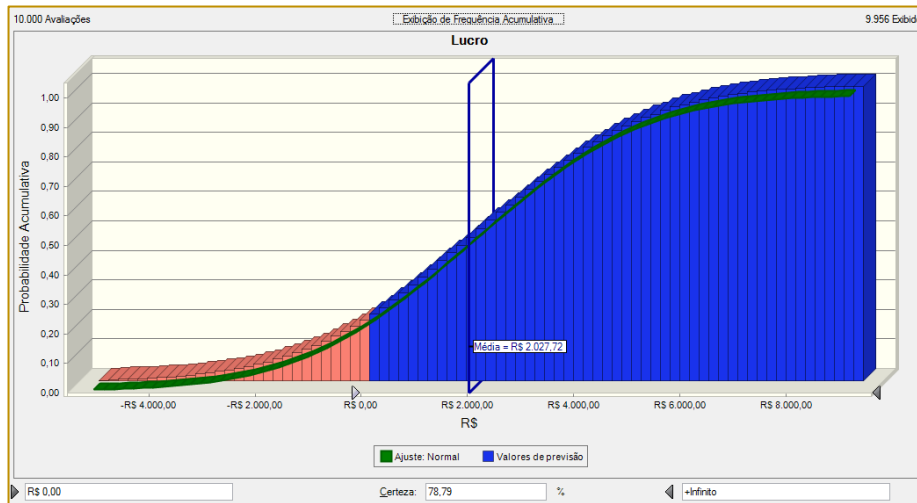
A simulação não tem a capacidade de garantir 100% de certeza, mas pode se alcançar a mitigação do risco relacionado a tal operação financeira.

A análise otimista é mostrada na Figura 6, com a obtenção de uma certeza de 78,79%

que os resultados da operação do investidor alcancem um lucro que fique entre R\$ 0,00 e R\$ 11.199,59. O cenário otimista se mostra um risco baixo e grau de certeza elevado.



Figura 6: Exibição do gráfico de frequência de previsão otimista.



Fonte: [Crystal Ball®, 2017]

Os resultados financeiros previstos neste artigo ainda carecem de outros custos, que será atribuído aos valores. Os custos relacionados são referentes às margens de garantia para a realização das operações. Estas margens variam de corretora para corretora, mas não influenciam nos resultados apresentados, pois é um dinheiro que o investidor deposita como garantia, somente perde o dinheiro se a operação for mal sucedida. O outro custo é o imposto de renda, que de forma semelhante não influencia nos resultados financeiros da previsão, mas terá que ser descontado do lucro do investidor a cada mês. No Brasil segundo o site da BMF Bovespa®, o imposto para operações *day trade*, está em 21%.

## 6. CONCLUSÃO

Neste estudo, os resultados da simulação sugerem que o método de MC, empregado para analisar as incertezas e os riscos considerados nos resultados de um estudo de caso, onde envolve mini contratos futuros na modalidade *day trade*, é uma ferramenta muito útil, pois podem ser observados

cenários e obter uma porcentagem de certeza perante os resultados previstos.

O *software* Crystal Ball® [CB] revela-se como um instrumento de instalação simples e de compreensível configuração e utilização. Comprovam-se funções no CB que sustenta as decisões de riscos de estimativas de lucro ou prejuízo em determinada operação financeira, onde há a possibilidade de aferições mais precisas.

É necessária a construção de um bom modelo para a correta utilização da simulação. A abordagem estocástica é um método poderoso para a quantificação do parâmetro de incerteza dos resultados obtidos, promovendo suporte aos profissionais e tomadores de decisões na gestão de suas diversas operações financeiras, encontrando o percentual de certeza para determinados resultados.

A simulação revela-se como um potente instrumento, porém não pode ser atribuída a mesma como a única fonte de informação para tomada de decisões. Portanto é uma ferramenta de apoio e não uma verdade plena.

## REFERÊNCIAS

- [1] Acebes, F.; Pereda, M.; Poza, D.; Pajares, J.; Galán, J. M. Stochastic earned value analysis using Monte Carlo simulation and statistical learning techniques. *International Journal of Project Management*, v. 33, n. 7, p. 1597-1609, 2015.
- [2] Arnold, U.; Yildiz, Ö. Economic risk analysis of decentralized renewable energy

infrastructures—A Monte Carlo Simulation approach. *Renewable Energy*, v. 77, p. 227-239, 2015.

- [3] Bandyopadhyay, P. Increasing the efficiency of Monte Carlo simulation with sampling from an approximate potential. *Chemical Physics Letters*, v. 556, p. 341-345, 2013.

- [4] Bieda, B. Application of stochastic approach based on Monte Carlo [MC] simulation for life cycle inventory [LCI] to the steel process

chain: Case study. *Science of the Total Environment*, v. 481, p. 649-655, 2014.

[5] Bieda, B. Stochastic approach to municipal solid waste landfill life based on the contaminant transit time modeling using the Monte Carlo [MC] simulation. *Science of the Total Environment*, v. 442, p. 489-496, 2013.

[6] Bm&F Bovespa. Índice Bovespa: Futuro Mini de Ibovespa. Disponível em: <[http://www.bmfbovespa.com.br/pt\\_br/produtos/listados-a-vista-e-derivativos/renda-variavel/futuro-mini-de-ibovespa.htm](http://www.bmfbovespa.com.br/pt_br/produtos/listados-a-vista-e-derivativos/renda-variavel/futuro-mini-de-ibovespa.htm)> Acesso em: 28 de Outubro de 2016.

[7] Bocchetti, V.; Diep, H. T. Monte Carlo simulation of melting and lattice relaxation of the [111] surface of silver. *Surface Science*, v. 614, p. 46-52, 2013.

[8] Cai, T.; Wang, S.; Xu, Q. Monte Carlo optimization for site selection of new chemical plants. *Journal of environmental management*, v. 163, p. 28-38, 2015.

[9] Cardoso, R. T. N.; Takahashi, R. H. C.; Cruz, F. R. B. Pareto optimal solutions for stochastic dynamic programming problems via Monte Carlo simulation. *Journal of Applied Mathematics*, 2013.

[10] Hudson, A.; Tilley, D. R. Assessment of uncertainty in emergy evaluations using Monte Carlo simulations. *Ecological Modelling*, v. 271, p. 52-61, 2014.

[11] Kasperski, a.; nieckarz, d.; szabelski, p. Structure formation in adsorbed overlayers comprising functional cross-shaped molecules: a monte carlo study. *Surface science*, v. 641, p. 269-277, 2015.

[12] Law, A. M. *Simulation Modeling & Analysis*. 4. ed. Boston: Ed. McGraw-Hill, 2007.

[13] Noronha, J. C.; Mello, C. H. P.; SILVA, C.; NORONHA, J. Opções reais aplicadas à gestão do processo de desenvolvimento de produto em uma

indústria de autopeças. *Gestão & Produção*, São Carlos, v. 21, n. 1, p. 77-94, 2014.

[14] Oliveira, M. R. G.; Neto, M. L. B. Simulação De Monte Carlo E Valuation: UMA Abordagem Estocástica. *REGE Revista de Gestão*, v. 19, n. 3, 2012.

[15] Oracle. Oracle Crystal Ball: Informações Inigualáveis Sobre os Fatores Críticos que Afetam o Risco. Disponível em: <<https://www.oracle.com/br/products/applications/crystalball/overview/index.html>> Acesso em: 27 de Outubro de 2016.

[16] Portal Action. Action Stat. Disponível em: <<http://www.portalaction.com.br/sobre-o-action>> Acesso em: 29 de Outubro de 2016.

[17] Rocha, C. H. M.; Sousa, R. R. C.; Campos, N. S. Uma análise da situação financeira da indústria brasileira de aviação civil. *Journal of Transport Literature*, v. 10, n. 3, p. 35-39, 2016.

[18] Sabbag, O. J.; Costa, S. M. A. L. Análise de custos da produção de leite: aplicação do método de Monte Carlo. *Extensão Rural*, v. 22, n. 1, p. 125-145, 2015.

[19] Souza, J. G. D. M.; DE Nascimento, A. A.; SOUSA, R. R. D. C.; CAMPOS, N. D. S.; MONTES, A. R. S. D. S. Incerteza da Viabilidade Econômica de Um Projeto Portuário: Uma Aplicação DA Simulação Monte Carlo. *Gestão e Saúde*, v. 1, n. 1, p. Pag. 1042-1056, 2015.

[20] Spivey, R. F.; Teki, R.; LU, T. M. Full 3D Monte Carlo simulation of pit-type defect evolution during extreme ultraviolet lithography multilayer deposition. *Thin Solid Films*, v. 540, p. 173-182, 2013.

[21] Vallée, F.; Versèle, C.; Lobry, J.; Moïny, F. Non-sequential Monte Carlo simulation tool in order to minimize gaseous pollutants emissions in presence of fluctuating wind power. *Renewable energy*, v. 50, p. 317-324, 2013.

# Capítulo 23

## ANÁLISE DE EFICIÊNCIA DAS POLICLÍNICAS DO ESTADO DO CEARÁ UTILIZANDO ANÁLISE ENVOLTÓRIA DE DADOS COM PONDERAÇÃO DE RESULTADOS

*Camila Campos Colares das Dores*

*Gerardo Valdisio Rodrigues Viana*

*Antônio Clécio Fontelles Thomaz*

**Resumo:** O presente trabalho propõe a análise de eficiência de 19 Policlínicas do Estado do Ceará utilizando-se a Análise Envoltória de Dados (*Data Envelopment Analysis - DEA*) associada ao método de Savage, o qual permite uma ponderação entre a fronteira DEA clássica (otimista) e a fronteira invertida (pessimista). Através desta associação e tomando por base os dados de funcionamento das Policlínicas entre os anos de 2011 e 2014, realiza-se uma análise de sensibilidade da eficiência dessas unidades. Conclui-se que é possível avaliar a eficiência relativa das DMUs (*Decision Making Units*) em unidades dessa natureza, demonstrando assim como é factível o emprego da metodologia DEA para proporcionar subsídios no processo decisório, abrindo campo para outras pesquisas dentro da área da saúde.

**Palavras Chave:** DEA, Policlínicas, Savage.

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1 INTRODUÇÃO

Grandes esforços têm sido dedicados ao estudo das formas de mensurar e fazer análises da eficiência de empresas privadas e de instituições públicas. Desde a década de 50, a partir dos trabalhos de [Koopmans(1951)], [Farrel(1957)] e [Debreu(1951)], técnicas não-paramétricas vêm sendo utilizadas para avaliar se um plano de operação é eficiente. Entretanto, foi a partir da década de 70 que houve grande progresso na aplicação dessas técnicas, visto que, sob o prisma gerencial, seus resultados revelaram-se mais expressivos que aqueles obtidos através da abordagem paramétrica tradicional. Assim, [Charnes et~al.(1978)Charnes, Cooper, e Rhodes] generalizaram os estudos de Farrel tanto no sentido de trabalhar com múltiplos recursos e múltiplos resultados, quanto na obtenção de um indicador que atendesse ao conceito de eficiência. Essa generalização deu origem a uma técnica de construção de fronteiras de produção e indicadores da eficiência produtiva conhecida como Análise Envoltória de Dados (Data Envelopment Analysis - DEA).

A Análise Envoltória de Dados – DEA é uma técnica não-paramétrica que emprega programação matemática para construir fronteiras por “unidades tomadoras de decisões”. Estas unidades (do inglês Decision Making Units - DMUs) empregam processos tecnológicos semelhantes para transformar múltiplos insumos em múltiplos produtos. Tais fronteiras são utilizadas para avaliar a eficiência relativa dos planos de operação executados pelas DMUs e servem também como referência para o estabelecimento de metas eficientes para cada unidade produtiva.

Baseado em [Pimenta e Soares~de Mello(2005)], este trabalho propõe uma

avaliação da eficiência de 19 destas unidades de atendimento durante os anos de 2011 a 2014, utilizando a Análise Envoltória de Dados associada ao método de Savage [Savage(1972)]. Com isto, pretende-se auxiliar os gestores de saúde no apoio à tomada de decisão, refletido na correta utilização do orçamento destinado ao funcionamento destas unidades.

## 2 A METODOLOGIA COM PONDERAÇÃO DE RESULTADOS

Nos modelos clássicos do DEA, os pesos atribuídos aos inputs e outputs são definidos pelas próprias DMUs. Dessa forma, as DMUs são favorecidas em suas melhores práticas, podendo então o modelo apontar como eficientes algumas DMUs que, ao serem analisadas através da fronteira invertida, mostram-se ineficientes. A análise de fronteira invertida é obtida através da inversão entre os inputs e outputs [Entani et~al.(2002)Entani, Maeda, e Tanaka] e permite uma observação mais sofisticada do problema pois as DMUs são avaliadas naquilo em que são ineficientes.

É possível calcular uma fronteira de eficiência composta ponderando-se os resultados obtidos através das fronteiras padrão e invertida. O resultado é alcançado através da média aritmética entre a eficiência padrão e o valor da subtração da unidade pela eficiência invertida, como mostrado na equação 1. É interessante ainda normalizarem-se esses resultados dividindo-se o valor da eficiência composta pelo maior valor entre todos os valores de eficiência composta, conforme mostrado na equação 2 [Meza et~al.(2004)Meza, Biondi~Neto, MELLO, Gomes, e Coelho].

$$\text{Eficiência Composta} = \frac{\text{Eficiência Padrão} + (1 - \text{Eficiência Invertida})}{2} \quad (1)$$

$$\text{Eficiência Composta Normalizada} = \frac{\text{Eficiência Composta}}{\text{Max(Eficiência Composta)}} \quad (2)$$

O método de Savage, utilizado em tomadas de decisão em ambientes com incerteza, permite uma ponderação entre os métodos otimista e pessimista. A equação 3 representa este método, onde  $\alpha$  corresponde ao coeficiente de ponderação,  $i$  representa cada opção de decisão (variável controlável),  $j$

representa os cenários possíveis (variável não controlável),  $\text{Min}K_{i,j}$  representa os valores favoráveis ao decisor e  $\text{Max}K_{i,j}$  representa os valores desfavoráveis.

$$\alpha \text{Min}K_{i,j} + (1 - \alpha) \text{Max}K_{i,j} \quad (3)$$

Pode-se reescrever a equação 3 para adaptá-la aos valores de eficiência obtidos através do DEA. Tem-se então a equação 4, onde  $Ef_o$  é o

valor de eficiência otimista e  $Ef_p$  é o valor de eficiência calculado pela fronteira invertida (pessimista).

$$Ef_{savage} = \alpha Ef_o + (1 - \alpha)(1 - Ef_p) \quad (4)$$

O valor de  $\alpha$  deve estar entre o intervalo (0-1). Quanto maior o valor de  $\alpha$ , mais benevolente será a avaliação.

Um exemplo da utilização do método de Savage em modelos DEA é a Eficiência Composta calculada pelo software SIAD [Meza et al. (2004) Meza, Biondi-Neto, MELLO, Gomes, e Coelho], onde o valor do coeficiente de ponderação  $\alpha$  é de 0,5, levando os métodos otimista e pessimista a possuírem o mesmo peso no cálculo da eficiência composta.

Para realizar a ponderação dos resultados de eficiência, será efetuada uma análise de sensibilidade da eficiência das DMUs para diversos valores do coeficiente  $\alpha$ . As DMUs que mantiverem seu índice de eficiência alto independente do valor de  $\alpha$  serão consideradas verdadeiramente eficientes, enquanto que as DMUs que demonstrarem uma diminuição do seu índice de eficiência conforme o valor de  $\alpha$  é reduzido (dando-se ênfase à avaliação pessimista) serão consideradas ineficientes.

### 3 ESTUDO DE CASO: ANÁLISE DE EFICIÊNCIA DAS POLICLÍNICAS DO ESTADO DO CEARÁ

#### 3.1 INTRODUÇÃO

O Estado do Ceará é precursor em iniciativas de conformação regional da saúde, com forte presença de planejamento e estruturas regionais sólidas e com considerável autonomia [Viana e Lima (2011)]. Dentro da visão de gestão pública por resultados, o Sistema Único de Saúde – SUS, através da Secretaria Estadual de Saúde do Ceará – SESA-CE, vem desenvolvendo ações no sentido de melhorar a qualidade dos serviços prestados à população e, ao mesmo tempo, reduzir os gastos financeiros envolvidos nos seus processos de atendimento [Linhares et al. (2014) Linhares, Lira, e Albuquerque]. Em consonância com o Plano Estadual de Saúde 2012-2015 [Ceara (2011)], esta secretaria busca assegurar atendimento de ações e serviços de saúde universal, integral

e humanizado a todos os cearenses, além de fortalecer a capacidade de planejamento e gestão do Sistema Estadual de Saúde de forma cooperativa e solidária observando o princípio de eficiência e equidade com participação da sociedade. Com isso, a preocupação na adoção de medidas de garantia de acesso da população à assistência à saúde nos níveis de média e alta complexidade torna-se importante item da agenda dos formuladores e gestores de saúde do estado do Ceará, até mesmo pela demanda gerada por essa maior efetividade do sistema.

Fundamentado em indicadores epidemiológicos e ainda considerando as lacunas assistenciais e o envelhecimento populacional, o Governo do Estado do Ceará lançou em 2009 o Programa de Expansão e Melhoria da Assistência Especializada à Saúde do Estado do Ceará – PROEXMAES. Como definido em [Ceara (2009)], este programa objetiva contribuir para a ampliação do acesso e da qualidade dos serviços especializados de saúde, promovendo a integração entre os distintos níveis de atenção, com a expansão da cobertura da assistência especializada em todas as microrregiões do interior do estado. Inseridas neste contexto estão as Policlínicas, que oferecem atendimentos nas principais especialidades médicas de interesse epidemiológico e com serviços de suporte ao diagnóstico e reabilitação dos pacientes atendidos. Entretanto, os custos de fornecimento e manutenção desses serviços são altos, cabendo aos gestores a utilização do orçamento destinado aos mesmos de forma otimizada e eficiente.

#### 3.2 MODELAGEM

O objetivo deste trabalho é medir a eficiência de 19 Policlínicas no Estado do Ceará no período de 2011 a 2014. As unidades tomadoras de decisão (DMUs) serão as próprias Policlínicas. Os inputs serão o tempo (em meses) em que cada Policlínica

funcionou durante os anos de 2011 a 2014 e o número de especialidades que oferecem atendimento em cada uma das DMUs. O único output do modelo será a quantidade de atendimentos realizados em todas as

especialidades oferecidas em cada DMU durante seu período de funcionamento. A tabela 1 apresenta os dados que serão utilizados no cálculo da eficiência dessas unidades.

Table 1: Dados utilizados para cálculo da eficiência das Policlínicas

| Dmu               | (i)Tempo | (i)Especialidades | (o)Atendimentos |
|-------------------|----------|-------------------|-----------------|
| Acaraú            | 12       | 9                 | 4.682           |
| Aracati           | 23       | 11                | 19.936          |
| Barbalha          | 6        | 10                | 2.786           |
| Baturité          | 39       | 12                | 52.105          |
| Brejo Santo       | 13       | 11                | 13.093          |
| Camocim           | 36       | 14                | 27.815          |
| Campos Sales      | 12       | 11                | 5.968           |
| Caucaia           | 7        | 13                | 7.474           |
| Crateús           | 2        | 11                | 1.735           |
| Icó               | 6        | 9                 | 6.939           |
| Iguatu            | 7        | 13                | 4.276           |
| Itapipoca         | 23       | 11                | 12.597          |
| Limoeiro do Norte | 2        | 14                | 1.506           |
| Pacajus           | 28       | 15                | 18.418          |
| Quixadá           | 6        | 10                | 4.499           |
| Russas            | 6        | 12                | 3.678           |
| Sobral            | 23       | 13                | 17.993          |
| Tauá              | 42       | 15                | 55.784          |
| Tianguá           | 6        | 10                | 874             |

O número reduzido de fatores (inputs e outputs) aumenta o poder discriminatório do modelo DEA [Soares-de Mello et al.(2002)Soares-de Mello, Gomes, Soares-de Mello, e Estellita-Lins]. Isso se dá pelo fato de que é recomendado que o número de DMUs analisadas seja pelo menos três vezes maior que a soma de inputs e outputs, conforme sugerem os trabalhos de [Bowlin(1998)] e [Nunamaker(1985)].

Embora várias formas da fronteiras possam ser determinadas, existem dois modelos que são considerados clássicos: o CCR (acrônimo para Charnes, Cooper e Rhodes) e o BCC (acrônimo para Banker, Charnes e Cooper). O modelo CCR constrói uma superfície linear por partes, envolvendo os dados. Trabalha com retornos constantes de escala, ou seja, qualquer variação nas entradas (inputs) produz variação proporcional nas saídas (outputs) [Charnes et al.(1978)Charnes, Cooper, e Rhodes]. Este modelo também é conhecido como CRS (do inglês Constant

Returns to Scale). O modelo BCC considera retornos variáveis de escala, substituindo o axioma da proporcionalidade entre inputs e outputs pelo axioma da convexidade [Banker et al.(1984)Banker, Charnes, e Cooper]. Por isso, esse modelo também é conhecido como VRS (do inglês Variable Returns to Scale). Ao obrigar que a fronteira seja convexa, o modelo BCC permite que DMUs que operam com baixos valores de inputs tenham retornos crescentes de escala e as que operam com altos valores tenham retornos decrescentes de escala. Ambos os modelos trabalham com duas formas básicas de uma DMU não eficiente tornar-se eficiente. A primeira é reduzindo os recursos, mantendo constantes os produtos (orientação a inputs). A segunda é fazendo o inverso (orientação a outputs). Neste trabalho considera-se o modelo BCC de escala variável objetivando maximizar a eficiência orientada a outputs, que são os serviços de saúde prestados pelas Policlínicas à população. Assim, considera-se o seguinte modelo BCC orientado a outputs.

$$\text{Min } h_0 = \sum_{k=1}^r v_i x_{i0} - u_*$$

Sujeito a:

$$\begin{aligned} \sum_{i=1}^r u_j y_{jk} &= 1 \\ \sum_{i=1}^r v_i x_{ik} - \sum_{j=1}^s u_j y_{jk} - v_* &\leq 0, \quad K = 1, 2, \dots, n \\ u_j e v_i &\geq 0 \quad \forall j, i \\ v_* &\in R \end{aligned}$$

Onde:

$h_0$ : eficiência;  
 $u_j, v_i$ : pesos de *outputs* e *inputs* respectivamente;  
 $x_{ik}, y_{jk}$ : *inputs*  $i$  e *outputs*  $j$  da  $DMU_k$ ;  
 $x_{i0}$ : *inputs*  $i$  da  $DMU_0$ .

### 3.3 RESULTADOS

O modelo DEA BCC orientado a outputs foi aplicado aos dados da tabela 1. Os resultados obtidos através do software OSDEA (solver de código livre desenvolvido

por Juergen Ebert e Mark James (<http://opensourcedea.org>) podem ser observados abaixo na tabela 2.

Table 2: Eficiência das Policlínicas do Estado do Ceará

| Dmu               | Padrão | Invertida | Composta | Composta normalizada |
|-------------------|--------|-----------|----------|----------------------|
| Acaraú            | 0,6746 | 1         | 0,3373   | 0,4919               |
| Aracati           | 0,6589 | 0,7852    | 0,4368   | 0,6370               |
| Barbalha          | 0,3950 | 0,8575    | 0,2687   | 0,3919               |
| Baturité          | 1      | 0,9464    | 0,5268   | 0,7682               |
| Brejo Santo       | 0,7851 | 0,7491    | 0,5180   | 0,7554               |
| Camocim           | 0,5792 | 1         | 0,2896   | 0,4223               |
| Campos Sales      | 0,3896 | 0,9648    | 0,2124   | 0,3097               |
| Caucaia           | 0,8750 | 0,9057    | 0,4846   | 0,7067               |
| Crateús           | 1      | 0,7927    | 0,6037   | 0,8803               |
| Icó               | 1      | 0,6284    | 0,6858   | 1                    |
| Iguatu            | 0,5006 | 0,9701    | 0,2653   | 0,3868               |
| Itapipoca         | 0,4163 | 1         | 0,2082   | 0,3035               |
| Limoeiro do Norte | 0,8680 | 1         | 0,4340   | 0,6329               |
| Pacajus           | 0,4960 | 1         | 0,2480   | 0,3617               |
| Quixadá           | 0,6378 | 0,7605    | 0,4387   | 0,6397               |
| Russas            | 0,5122 | 0,9129    | 0,2997   | 0,4370               |
| Sobral            | 0,5934 | 0,8681    | 0,3626   | 0,5288               |
| Tauá              | 1      | 1         | 0,5000   | 0,7291               |
| Tianguá           | 0,1239 | 1         | 0,0620   | 0,0903               |

Analisando-se a tabela 2, percebe-se pelos resultados da fronteira padrão que as Policlínicas de Baturité, Crateús, Icó e Tauá são eficientes. Em contrapartida, tem-se Tianguá, Campos Sales e Barbalha indicando menor eficiência. Pelos resultados da fronteira invertida, entretanto, percebe-se que Tauá é ineficiente, evidenciando que a fronteira padrão indicou uma falsa eficiência para esta DMU. Tal resultado é interessante e merece atenção especial por parte dos gestores responsáveis pois esta Policlínica é

considerada como referência no Estado e o território onde está inserida tem sido historicamente beneficiado com projetos e programas piloto tais como a Planificação da Atenção Primária em Saúde e o recente QualificaAPSUS.

Os resultados obtidos através da fronteira composta permitem uma melhor avaliação das DMUs. A tabela 3 mostra rankings comparando a eficiência padrão e a eficiência composta normalizada.

Table 3: Comparativo dos rankings das Policlínicas do Estado do Ceará

| DMU               | Rank Padrão | Ef. Padrão | Rank Comp. Norm. | Ef. Comp. Norm. |
|-------------------|-------------|------------|------------------|-----------------|
| Icó               | 1           | 1          | 1                | 1               |
| Crateús           | 1           | 1          | 2                | 0,8803          |
| Baturité          | 1           | 1          | 3                | 0,7682          |
| Tauá              | 1           | 1          | 5                | 0,7291          |
| Caucaia           | 5           | 0,8750     | 6                | 0,7067          |
| Limoeiro do Norte | 6           | 0,8680     | 9                | 0,6329          |
| Brejo Santo       | 7           | 0,7851     | 4                | 0,7557          |
| Acaraú            | 8           | 0,6746     | 11               | 0,4919          |
| Aracati           | 9           | 0,6589     | 8                | 0,6370          |
| Quixadá           | 10          | 0,6378     | 7                | 0,6397          |
| Sobral            | 11          | 0,5934     | 10               | 0,5288          |
| Camocim           | 12          | 0,5792     | 13               | 0,4223          |
| Russas            | 13          | 0,5122     | 12               | 0,4370          |
| Iguatu            | 14          | 0,5006     | 15               | 0,3868          |
| Pacajus           | 15          | 0,4960     | 16               | 0,3617          |
| Itapipoca         | 16          | 0,4163     | 18               | 0,3035          |
| Barbalha          | 17          | 0,3950     | 14               | 0,3919          |
| Campos Sales      | 18          | 0,3896     | 17               | 0,3097          |
| Tianguá           | 19          | 0,1239     | 19               | 0,0903          |

Após análise desta tabela, percebe-se que Icó e Tianguá (respectivamente a primeira e última colocadas no ranking padrão) permaneceram em suas posições. Destaca-se ainda a Policlínica de Tauá, que perdeu 4 posições no ranking, ficando em 5º lugar no ranking da Eficiência Composta Normalizada. Brejo Santo e Barbalha obtiveram ganho de 3 posições, passando de 7º para 4º lugar e de 17º para 14º lugar respectivamente. Limoeiro do Norte, Acaraú e Quixadá obtiveram uma perda considerável de 3 posições no ranking. As demais policlínicas oscilaram apenas 1 posição.

### 3.4 APLICAÇÃO DO MÉTODO PONDERADO

A aplicação da ponderação ao caso das Policlínicas através do uso da metodologia de Savage permite que seja realizada uma análise de sensibilidade da eficiência dessas unidades. Realiza-se tal ponderação variando-se o coeficiente de otimismo  $\alpha$ . É importante observar que a Eficiência Composta calculada anteriormente corresponde ao modelo ponderado com coeficiente  $\alpha=0,5$ , conforme é possível observar na equação 5.

$$E_{fsavage} = \alpha E_{fo} + (1 - \alpha)(1 - E_{fp}) = 0,5E_{fo} + (1 - 0,5)(1 - E_{fp}) = 0,5E_{fo} + 0,5(1 - E_{fp}) = \frac{E_{fo} + (1 - E_{fp})}{2} \quad (5)$$

Na tabela 4 são mostradas as eficiências obtidas através da aplicação da equação 4, variando-se o valor do coeficiente  $\alpha$  de 1 a 0,1

em intervalos de 0,1. Os valores apresentados na tabela estão normalizados.



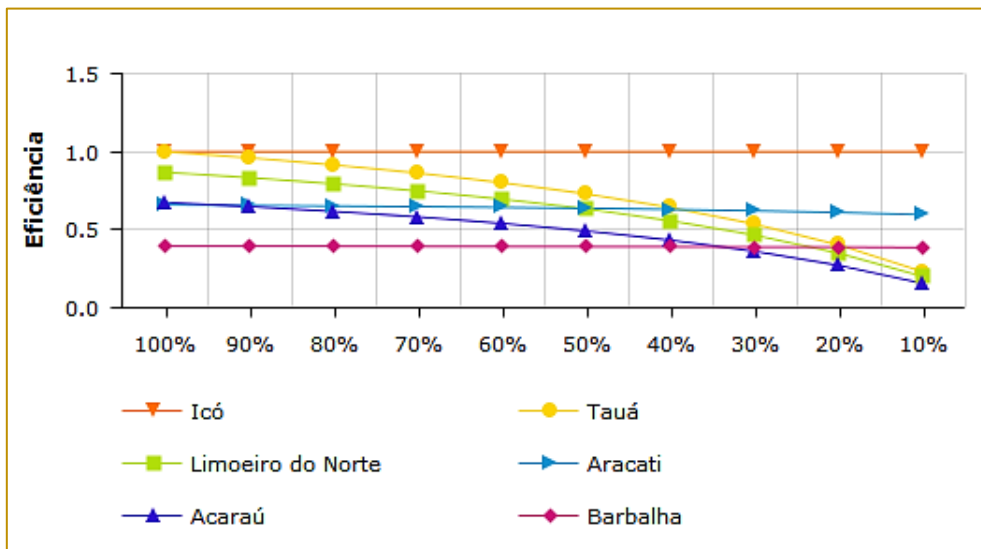
Table 4: Eficiências das Policlínicas do Estado do Ceará para diversos valores de  $\alpha$

| Alfa (coeficiente do modelo ponderado) |       |       |       |       |       |       |       |       |       |       |
|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| DMU                                    | 100%  | 90%   | 80%   | 70%   | 60%   | 50%   | 40%   | 30%   | 20%   | 10%   |
| Icó                                    | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     | 1     |
| Crateús                                | 1     | 0,983 | 0,962 | 0,939 | 0,912 | 0,88  | 0,842 | 0,795 | 0,736 | 0,66  |
| Baturité                               | 1     | 0,966 | 0,927 | 0,883 | 0,83  | 0,768 | 0,694 | 0,603 | 0,489 | 0,341 |
| Tauá                                   | 1     | 0,96  | 0,915 | 0,863 | 0,802 | 0,729 | 0,642 | 0,536 | 0,502 | 0,23  |
| Caucaia                                | 0,875 | 0,85  | 0,822 | 0,79  | 0,752 | 0,707 | 0,653 | 0,587 | 0,504 | 0,397 |
| Limoeiro do Norte                      | 0,868 | 0,834 | 0,794 | 0,749 | 0,696 | 0,633 | 0,557 | 0,465 | 0,349 | 0,2   |
| Brejo Santo                            | 0,785 | 0,781 | 0,776 | 0,77  | 0,763 | 0,755 | 0,746 | 0,734 | 0,72  | 0,701 |
| Acaraú                                 | 0,675 | 0,648 | 0,617 | 0,582 | 0,541 | 0,492 | 0,433 | 0,361 | 0,271 | 0,155 |
| Aracati                                | 0,659 | 0,656 | 0,652 | 0,648 | 0,643 | 0,637 | 0,63  | 0,621 | 0,611 | 0,597 |
| Quixadá                                | 0,632 | 0,638 | 0,638 | 0,639 | 0,639 | 0,64  | 0,64  | 0,641 | 0,642 | 0,643 |
| Sobral                                 | 0,593 | 0,584 | 0,573 | 0,561 | 0,546 | 0,529 | 0,508 | 0,483 | 0,451 | 0,41  |
| Camocim                                | 0,579 | 0,556 | 0,53  | 0,5   | 0,464 | 0,422 | 0,372 | 0,31  | 0,233 | 0,133 |
| Russas                                 | 0,512 | 0,501 | 0,489 | 0,474 | 0,457 | 0,437 | 0,413 | 0,383 | 0,346 | 0,298 |
| Iguatu                                 | 0,501 | 0,484 | 0,465 | 0,443 | 0,417 | 0,387 | 0,35  | 0,306 | 0,25  | 0,177 |
| Pacajus                                | 0,496 | 0,476 | 0,454 | 0,428 | 0,398 | 0,362 | 0,319 | 0,266 | 0,2   | 0,114 |
| Itapipoca                              | 0,416 | 0,4   | 0,381 | 0,359 | 0,334 | 0,304 | 0,267 | 0,223 | 0,167 | 0,096 |
| Barbalha                               | 0,395 | 0,395 | 0,394 | 0,393 | 0,393 | 0,392 | 0,391 | 0,39  | 0,388 | 0,386 |
| Campos Sales                           | 0,39  | 0,378 | 0,365 | 0,349 | 0,331 | 0,31  | 0,284 | 0,253 | 0,213 | 0,163 |
| Tianguá                                | 0,124 | 0,119 | 0,113 | 0,107 | 0,099 | 0,09  | 0,08  | 0,066 | 0,05  | 0,029 |

Na figura 1 é mostrada a representação gráfica da evolução das eficiências de algumas Policlínicas. É possível perceber que as Policlínicas de Icó, Aracati e Barbalha apresentam pouca ou nenhuma variação em seu valor de eficiência independente do valor

atribuído ao coeficiente de ponderação. Outras como Tauá, Acaraú e Limoeiro do Norte tem suas eficiências reduzidas à medida em que o coeficiente diminui, dando maior ênfase às más práticas dessas unidades.

Figure 1: Evolução das eficiências das Policlínicas



#### 4 CONCLUSÕES

A metodologia DEA aplicada na área de saúde para medir desempenho das Policlínicas de atenção secundária do SUS no estado do Ceará se mostrou altamente robusta, confiável, precisa e de fácil entendimento por parte dos gestores públicos nos processos de apoio às decisões gerenciais e operacionais.

A combinação entre as fronteiras padrão e invertida, formando a fronteira composta aumenta o poder discriminatório do DEA, evidenciando as DMUs que são falsas eficientes.

A análise de sensibilidade resultando da combinação entre o DEA e o modelo de Savage se constitui numa rica e poderosa ferramenta de gestão pois permite uma avaliação de desempenho mais precisa das DMUs. Nesta análise, as DMUs que

mantiveram seu nível de eficiência alta independente do valor do coeficiente de ponderação são consideradas verdadeiras eficientes.

As próximas etapas envolvem o desenvolvimento do software Eficiência.Poli, o qual proporcionará uma fácil interação com os gestores das Policlínicas do Estado através de ambientes amigáveis para composição de análises de dados reais e de simulações via mudança de parâmetros controláveis de *inputs* e *outputs*. O software será desenvolvido em linguagem JAVA e usará outros modelos DEA além do BCC-O utilizado neste trabalho. Espera-se com este trabalho poder contribuir para qualidade e celeridade do processo de atendimento das Policlínicas do estado do Ceará.

#### REFERÊNCIAS

- [1] [Banker et~al.(1984)Banker, Charnes, e Cooper] Banker, R. D., Charnes, A., e Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 300 (9):0 1078–1092.
- [2] [Bowlin(1998)] Bowlin, W. F. (1998). Measuring performance: An introduction to data envelopment analysis (dea). *The Journal of Cost Analysis*, 150 (2):0 3–27.
- [3] [Ceara(2009)] Ceara (2009). Portaria n. 6493, de 03 de dezembro de 2009. diário oficial do estado do ceara. <http://imagens.seplag.ce.gov.br/PDF/20091203/do20091203p02.pdf> . Acessado: 2016-02-29.
- [4] [Ceara(2011)] Ceara (2011). Secretaria da saude do estado do ceara - plano estadual de saude 2012-2015. <http://tinyurl.com/jhwayux> . Acessado: 2016-02-29.
- [5] [Charnes et~al.(1978)Charnes, Cooper, e Rhodes] Charnes, A., Cooper, W. W., e Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal Operational Research*, Amsterdam, North-Holland, 2:0 429–444.
- [6] [Debreu(1951)] Debreu, G. (1951). The measurement of productive efficiency. *Journal of Econometric Society*, 190 (3):0 "273–292".
- [7] [Entani et~al.(2002)Entani, Maeda, e Tanaka] Entani, T., Maeda, Y., e Tanaka, H. (2002). Dual models of interval dea and its extension to interval data. *European Journal of Operational Research*, 1360 (1):0 32–45.

- [8] [Farrel(1957)] Farrel, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistic Society*.
- [9] [Koopmans(1951)] Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. in:; wiley, jone. activity analysis of production and allocation. New York Cowles Comission for Research in Economics.
- [10] [Linhares et~al.(2014)Linhares, Lira, e Albuquerque] Linhares, P. H. A., Lira, G. V., e Albuquerque, I. M. N. (2014). Avaliação do programa nacional de melhoria do acesso e da qualidade da atenção básica no estado do ceará. *Saúde debate*.
- [11] [Meza et~al.(2004)Meza, Biondi~Neto, MELLO, Gomes, e Coelho] Meza, L. A., Biondi Neto, L., MELLO, J., Gomes, E., e Coelho, P. (2004). FsdA-free software for decision analysis (slad-software livre de apoio à decisão): a software package for data envelopment analysis models. In CONGRESO LATINO-IBEROAMERICANO DE INVESTIGACIÓN OPERATIVA-CLAIO, volume 12.
- [12] [Nunamaker(1985)] Nunamaker, T. R. (1985). Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. *Managerial and decision Economics*, 60 (1):0 50–58.
- [13] [Pimenta e Soares~de Mello(2005)] Pimenta, H. e Soares de Mello, J. (2005). Modelo dea-savage para análise de eficiência do parque de refino brasileiro. *Relatórios de Pesquisa em Engenharia de Produção*, 50 (5).
- [14] [Savage(1972)] Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.
- [15] [Soares~de Mello et~al.(2002)Soares~de Mello, Gomes, Soares~de Mello, e Estellita~Lins]

Soares de Mello, J., Gomes, E. G., Soares de Mello, M. H. C., e Estellita Lins, M. (2002). Método multicritério para seleção de variáveis em modelos de. Pesquisa Naval, 15:0 55–66.

[16] [Viana e Lima(2011)] Viana, A. L. d. e Lima, L. D. d. (2011). Regionalização e relações

federativas na política de saúde do Brasil. In Regionalização e relações federativas na política de saúde do Brasil. Contra Capa.

# Capítulo 24

## *ANÁLISE DO COMPARTILHAMENTO DE VIAS ENTRE O MODAL RODOVIÁRIO E FERROVIÁRIO POR MEIO DE SIMULAÇÃO DE EVENTOS DISCRETOS*

*Laura Luiza Pereira David*

*Gisele Tessari Santos*

**Resumo:** Sabe-se que o planejamento a longo prazo nas empresas permite sua perenidade. O planejamento de investimentos em infraestrutura é um dos diversos pontos de atuação dessa área dentro de uma empresa. Ele permite analisar quais investimentos são necessários na empresa, além de averiguar se ao desembolsar um determinado valor em investimento a empresa terá um retorno favorável. Este artigo visou analisar a necessidade de investimento em um provável gargalo no cruzamento entre um modal ferroviário e rodoviário por meio de um modelo de simulação de eventos discretos. Ao desenvolver o modelo de simulação, todas as premissas rodoviárias e ferroviárias do trecho, e todo o volume de veículos e de trens esperados tanto na rodovia quanto na ferrovia respectivamente, até 2021, foram considerados. Estes volumes aumentam consideravelmente por ano, e por terem que esperar a passagem de trem, os caminhões e carros que estão em sentido entrada e saída de um porto não poderiam ter grandes interferências em seus processos. Portanto, como fator de decisão sobre a necessidade de investimento, analisou-se a variável fila máxima de veículos a qual não deveria ultrapassar o valor de 40 unidades em fila em ambos os sentidos. Três cenários foram simulados. Os resultados em todos os cenários demonstraram que a fila de veículos que esperam a passagem de trens no cruzamento não ultrapassa 40 unidades, e, portanto, constatou-se que não é necessário um investimento em infraestrutura, como a construção de um viaduto no cruzamento, no período analisado.

**Palavras chave:** Simulação de eventos discretos, planejamento de transporte, compartilhamento de via.

## 1. INTRODUÇÃO

A necessidade de planejamento no sistema de transporte no Brasil tem se tornado cada vez mais evidente. De acordo com o Departamento Nacional de Trânsito, entre os anos de 2005 a 2015, a frota brasileira aumentou 116% (DENATRAN, 2017). Esse crescimento vertiginoso exige uma maior eficiência no uso das vias para que o fluxo de veículos flua de maneira contínua.

O trânsito de veículos sem interrupções está diretamente relacionado à capacidade das vias em que eles transitam. Segundo Araujo (2009), "Gargalo é qualquer recurso cuja capacidade é inferior à demanda". Portanto, planejar o sistema como um todo é importante para que a ocupação de uma via seja menor que sua capacidade, evitando assim, os congestionamentos e os impactos negativos no trânsito causados pelos gargalos.

Uma realidade no Brasil são as soluções paliativas dadas aos problemas de trânsito que não são sanados à longo prazo. Isso ocorre justamente pela falta de planejamento e de investimento em infraestrutura. Em um estudo feito no ano 2000 pela Fundação Getúlio Vargas e pela União Nacional da Construção, o Brasil ficou classificado abaixo de países como Namíbia, Estônia e Hungria em relação à infraestrutura e à vascularização da malha rodoviária (FGV, 2006).

O *U.S. Department of Transportation* (USDT, 2007) afirma que o planejamento de transporte deve ser um processo cooperativo concebido para promover o envolvimento de todos os usuários do sistema, como a comunidade empresarial, grupos da comunidade de cidadãos, organizações ambientais, os operadores de transporte, os viajantes e o público em geral. Coletar informações relacionadas à população, renda, empresas, indústrias, comércio e aeroportos, auxiliam na análise sobre qual é a ocupação que determinado número de pessoas demanda de um espaço compartilhado.

Segundo a Associação Nacional de Transportes Públicos (ANTP, 1997), o planejamento de transportes é a "atividade que define a infraestrutura necessária para assegurar a circulação de pessoas e mercadorias". Já Litman (2011), define que o processo de planejamento de transporte, normalmente, estabelece um nível de serviço a ser ofertado e avalia os impactos relacionados aos custos de infraestrutura para

o governo, tempos de viagem, congestionamento mínimo, risco de acidentes por quilometro, impactos ambientais, e o custo de utilização do automóvel. Um sistema de transporte eficiente permite crescimento e desenvolvimento da nação, além de disponibilizar recursos e facilitar a mobilidade (VASCONCELOS, 2008).

A infraestrutura de um país é fundamental para seu desenvolvimento econômico. Porém, as atividades econômicas relacionadas ao transporte geram externalidades associadas à impactos positivos e negativos. As externalidades positivas são os ganhos de eficiência para a economia, vantagens sociais da conexão de regiões e vantagens no segmento da atividade econômica. As negativas dizem respeito a congestionamento de tráfego, poluição e acidentes (OLIVEIRA, 2013).

Considerado o modal mais expressivo no transporte de cargas, as rodovias atingem praticamente todos os pontos do território nacional (RIBEIRO, 2002). Mas o espaço urbano é o fator escasso na provisão dos serviços rodoviários de transporte já que se pode constatar uma oferta escassa em recurso e uma demanda em excesso (LACERDA, 2006). A medida que a capacidade viária é ocupada, o tempo de trânsito dos veículos em viagem aumentam (LACERDA, 2006). As soluções aplicadas às externalidades negativas como congestionamento de tráfego, ou, aumento no tempo de viagem, estão relacionadas às alternativas que buscam aumentar a capacidade do recurso via, ou, diminuir a ocupação da mesma. No Brasil, as seguintes soluções têm sido propostas: ampliação de vias, rodizio de placas, construção de anéis rodoviários, viadutos, trincheiras e túneis.

A simulação de eventos discretos é uma importante ferramenta que facilita o planejamento de um sistema como um todo e, portanto, pode ser utilizada para a análise e planejamento do compartilhamento de vias a fim de evitar as externalidades negativas. Para Prado (1999), dentre as técnicas disponíveis, a simulação é a mais utilizada na modelagem de sistemas que apresentam gargalos. Esta técnica busca identificar gargalos, ou seja, os possíveis pontos em que a capacidade não é suficiente para atender a demanda. Esses pontos são onde há formação de grandes filas. Esse método tem sido utilizado na análise de vários sistemas

reais por auxiliar na tomada de decisão permitindo um melhor dimensionamento entre recursos, demanda e tempos de processo, e, conseqüentemente, proporcionando um aumento no nível de serviço. A metodologia tem sido utilizada em análises de fluxos de produção em indústrias, em serviços oferecidos por hospitais e bancos, no planejamento logístico e militar, dentre outros (HILL, 2001; LANNONI, 2006; GUNAL, 2010).

Um estudo de simulação pode economizar tempo e recursos financeiros (FREITAS, 2008). Utilizar a simulação de eventos discretos facilita o entendimento sobre a iteração das variáveis dentro do sistema, além de ser uma alternativa para teste de novas hipóteses operacionais sem comprometer o funcionamento do sistema. Entretanto, uma dificuldade ao se utilizar este método é a falta de dados dos sistemas reais. Cabe destacar que sem os dados reais é mais difícil afirmar que o modelo de simulação desenvolvido é devidamente válido e pode ser utilizado para análises sobre o sistema real. Assim, para se realizar uma simulação, é preciso de uma base de dados históricos do sistema real ou de um sistema considerado *benchmark*. Obter dados adequados para utilizar no modelo de simulação é geralmente demorado, difícil e dispendioso. Esses problemas são, muitas vezes, a razão que leva a validação do modelo ao fracasso (SARGENT, 2011).

Diante do exposto, o presente trabalho tem como objetivo desenvolver um modelo de simulação para o cruzamento, tecnicamente chamado de passagem em nível, situado entre uma ferrovia e uma estrada de acesso ao Terminal Integrador Portuário Luiz Antônio Mesquita, na região metropolitana de Santos (SP). A partir da simulação, analisar-se-á se a capacidade do recurso "cruzamento" limitando-se o fluxo de veículos que passa por ele de acordo com a quantidade de veículos e trens esperados para os próximos anos. O objetivo com essa simulação é definir para quando um investimento em infraestrutura deve ser provisionado de acordo com o planejamento da infraestrutura necessária para que o cruzamento não seja um gargalo. Para isso, diferentes cenários foram simulados e analisados.

## 2. REFERENCIAL TEÓRICO

### 2.1 CONTEXTUALIZAÇÃO DO PROBLEMA REAL

O cruzamento, estudado neste artigo, entre a ferrovia e a estrada de acesso ao porto. Está situado depois da alça da BR-101, da portaria do porto e do início da estrada de acesso ao porto, que passa pelo cruzamento com a ferrovia e continua em direção ao porto.

O porto está em obras de expansão que serão finalizadas em abril de 2017. A expansão levará a um aumento do volume processado pelo porto, conseqüentemente, um número maior de trens e caminhões entrará no sistema utilizando a ferrovia, o cruzamento e a estrada de acesso ao porto. Isso levantou um ponto de atenção em relação a necessidade de investimento no cruzamento visando um aumento de capacidade.

O fluxo de veículos na estrada de acesso ao porto devem esperar antes ou depois do cruzamento caso tenha algum trem na ferrovia cruzando a estrada de acesso ao porto. O número máximo de 40 veículos em ambos os sentidos é o limite para que os veículos não ocupem a BR-101 e/ou interfiram negativamente algum processo portuário. Considera-se que o cruzamento seja gargalo no sistema se a fila em qualquer sentido ultrapassar 40 veículos.

Visto isso, visa-se analisar o nível de ocupação do cruzamento e o número de veículos em fila a fim de permitir um planejamento para determinar quando um possível investimento em infraestrutura será necessário. Uma solução para diminuir esta ocupação será a construção de um viaduto para que os veículos não esperem a passagem do trem.

### 2.2 SIMULAÇÃO DE EVENTOS DISCRETOS

Simular é a tentativa de representar algo real. Segundo Shannon (1976), a simulação é o processo para desenvolver um modelo do sistema real e conduzi-lo a experimentos com o propósito de analisar o comportamento do sistema e avaliar as melhores estratégias para opera-lo. Já Kelton (2009), se refere à simulação como um amplo grupo de métodos que visam permitir representação do sistema real utilizando softwares computacionais.

As ocorrências ou os processos neste sistema

real são chamados de eventos. Eles interferem no tipo de simulação a ser utilizada. Para Silva (2009), os eventos para utilizarmos a simulação discreta podem ser acontecimentos como a chegada de caminhões programada ou não, e que nos modelos discretos as variáveis mantêm um comportamento ao longo de um intervalo de tempo com alterações em momentos bem definidos e conhecidos do tempo.

Ingalls (2008) defende que o poder da simulação de eventos discretos para as demais simulações, ou a principal diferença é que a simulação de eventos discretos tem a habilidade de imitar as dinâmicas de um sistema real. Já para Shannon (1976), é uma ferramenta disponível para planejamento e controle de eventos complexos.

A simulação pode ser aplicada em várias áreas como sistemas de manufatura, sistemas públicos ou sistemas que oferecem serviços (BANKS, 1998). Encontra-se na literatura, o método de simulação de eventos discretos aplicado à otimização de serviços de saúde (GUNAL, 2010), em análises da logística de sistemas de suprimentos (LANNONI, 2006), em simulação de fluxo de passageiros em um aeroporto (GUIZZI, 2009), em problemas de modelos militares (HILL, 2001), em simulação do transporte logístico de petróleo bruto (CHENG, 2004), entre outros.

### 3. METODOLOGIA

As análises dos resultados da simulação de eventos discretos pelo software Arena, neste trabalho, buscam identificar se a construção de um viaduto no cruzamento é necessária. O primeiro passo para construir um modelo que simule o sistema real foi desenvolver um fluxograma que contenha todos os processos que utilizam o recurso cruzamento. O cruzamento é utilizado pelo fluxo dos produtos enxofre, rocha e fosfato que saem do porto via caminhão, por carros de passeio tanto de funcionários como de visitantes, e pelos trens que passam na ferrovia.

Os fluxogramas foram divididos por tipo de elementos que requisitam o recurso e o

sentido que eles transitam. Sendo assim, tem-se o grupo de trens importação, grupo de trens exportação, grupo de carros/caminhões importação, e grupo de carros/caminhões exportação.

O fluxo de trens foi dividido em dois grupos de acordo com a direção que eles passam no cruzamento. O sentido é importação se o trem está no sentido porto-país ou exportação se o sentido for país-porto. A rota é a mesma para ambos os fluxos de trem porém elas tem o sentido invertido.

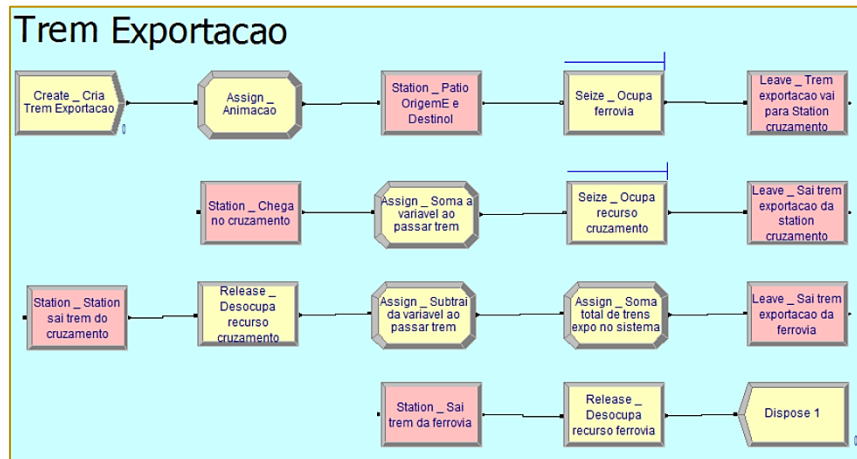
Dentro desse sistema, considerando o sentido exportação, o trem percorre uma distância de 3 km entre sua origem e o cruzamento, e de 13 km entre o cruzamento e seu destino. Na importação, por ser o sentido oposto, o trem percorre 13km entre sua origem e cruzamento, e 3km entre cruzamento e seu destino.

Durante esse percurso em ambos os sentidos, o recurso ferrovia, que tem a capacidade de um trem por vez, é utilizado. A velocidade permitida entre a origem e o destino é de 25 km/h, e de 10 km/h durante o cruzamento. A distância do cruzamento tem por base o comprimento do trem (1,5 km) e a margem de segurança (0,1 km) para a parte inicial e final do trem.

O recurso ferrovia tem a capacidade de 1 trem, e ele é ocupado durante o tempo total necessário para percorrer 17,7 km considerando a velocidade necessária em cada trecho. O recurso cruzamento é ocupado durante o tempo que o trem precisa para percorrer 1,7 km considerando a velocidade de 10 km/h e o recurso ferrovia é ocupado durante o tempo que o trem precisa para percorrer 16km a uma velocidade de 25km/h. Os recursos serão desocupados quando o tempo de ocupação estiver completo.

A seguir pode-se observar na figura 1 o fluxograma de trens no sentido exportação. O fluxograma de trens no sentido importação é igual ao da figura 1 se diferenciando apenas pela ordem em que os processos estão dispostos, uma vez que eles são os mesmos em sentidos opostos.

Figura 1 – Fluxo de trens no sentido exportação



Os *inputs* do modelo tiveram como base os dados previstos por meio do estudo realizado pelas operadoras ferroviárias que utilizam a ferrovia. Neste estudo, é exposta a quantidade de pares de trens por dia. Os resultados deste estudo são 16,3 pares em 2017, 18,9 pares em 2019 e 19,6 pares em 2021.

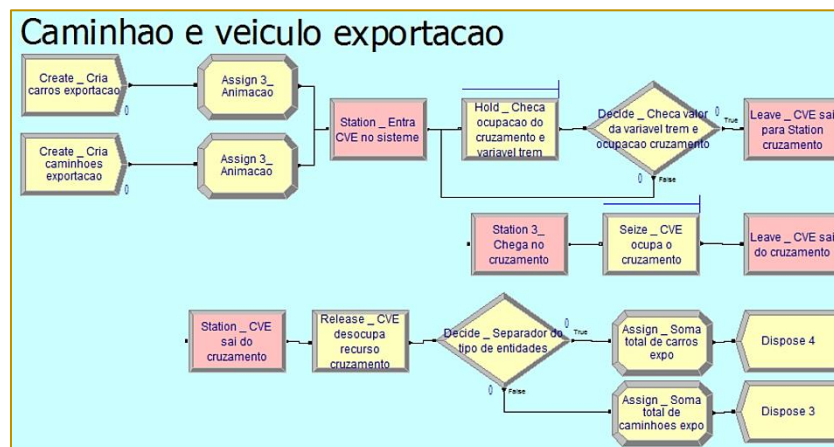
O fluxo de carros e caminhões também tem a mesma rota e são separados pelo sentido exportação que é a chegada dos veículos e importação que é a saída.

O recurso compartilhado nesse sistema é o

cruzamento que é ocupado por 12 segundos tanto por caminhões quanto pelos carros. Esse tempo foi obtido por meio de observações do sistema real. Por se tratar de uma estrada duplicada, não existem o impedimento que os veículos no sentido exportação travem os veículos do sentido importação e vice-versa.

Observa-se na figura 2 o fluxograma de carros e caminhões no sentido exportação. O fluxograma de carros e caminhões no sentido importação é igual ao da figura 2 alterando-se apenas o sentido dos processos.

Figura 2 – Fluxo de carros e caminhões no sentido exportação



Os *inputs* para o fluxo de caminhões tiveram como base o volume de produtos processados no porto e a capacidade média de 25 toneladas por caminhão. O volume processado no porto é definido pelo Plano de Negócios da empresa, que é um documento desenvolvido e dimensionado sobre a

demanda esperada de produtos para os próximos anos, e a capacidade dos ativos.

O cálculo do intervalo de chegada de caminhões por dia foi realizado a partir do volume de produtos esperado no mês de maior demanda, da capacidade média do



caminhão, e do número de caminhões necessários para transportar esse volume. Essa taxa do mês de maior pico foi representada por uma curva exponencial e fixada para todos os meses do ano, respeitando o comportamento esperado dentro das 24 horas do dia de um porto.

O *input* para o fluxo de carros teve por base as observações no porto. De acordo com essas observações, espera-se em média cerca de 50 a 60 carros por dia. E esse número é o esperado até o ano de 2021.

Considerando que o porto está em processo de expansão com finalização das obras em 2017, o modelo que simulou o sistema real para o ano de 2017 foi validado utilizando-se

dados históricos de 2015 dos processos que não se alteraram. Portanto, para a validação, utilizaram-se os dados relacionados ao volume dos produtos via rodoviário realizado pelo porto no ano de 2015, o número de carros de passeio que circulavam no mesmo ano e o dado histórico de 3 pares de trens dia que circulavam na margem esquerda em 2015.

A amostra piloto foi rodada com estes dados com um tempo total de simulação igual a 365 dias e 1 replicação. A partir da amostra piloto, calculou-se o número adequado de replicações considerando um intervalo de confiança de 95% de acordo com a expressão a seguir:

$$n^* = \left[ n \left( \frac{h}{h^*} \right)^2 \right] \quad (1)$$

em que,

$h^*$  = precisão desejada;

$h$  = precisão da amostra piloto;

$n^*$  = número de replicações necessárias;

$n$  = número de replicações da amostra piloto.

Após definir o número de replicações adequado, o modelo foi simulado novamente a fim de checar se o número de caminhões, carros e pares de trens dia gerados pelo sistema condizem com os dados históricos disponíveis. Após a validação do modelo, este foi utilizado para simular os cenários atuais e futuros.

As premissas como distância entre os pátios ferroviários, a velocidade permitida em cada trecho, e o tempo de ocupação do recurso foram as mesmas por tipo de entidade em todos os cenários.

Os parâmetros de entrada foram o intervalo entre chegadas de trens, obtido no estudo da ferrovia até 2021; o intervalo de caminhões, refere-se a demanda para os próximos anos até 2021 e a capacidade dos caminhões; e o intervalo de carros, que se manterão com uma média total entre 50 e 60 carros esperados por dia até 2021. Todos os parâmetros de intervalo de chegada utilizados em cada cenário são referentes ao ano da simulação.

As principais medidas de desempenho do sistema coletadas e analisadas foram a fila máxima de carros e caminhões no

cruzamento que não pode exceder 40 veículos, e a fila de trens no cruzamento que deve ser igual à zero. Assim, garante-se que o trânsito de trens é preferencial ao de carros e caminhões, e que a fila de caminhões é menor que o limite de veículos em fila. Adicionalmente, desejou-se saber a taxa de ocupação do cruzamento.

O primeiro cenário foi definido como cenário base, no qual simulou-se o modelo com os *inputs* do ano de 2017 e 2019. Os demais cenários seriam definidos à partir da análise das saídas do seu anterior. Visa-se chegar ao cenário em que a fila de veículos em qualquer sentido ultrapassasse 40 unidades utilizando os dados previstos até 2021. E se na simulação do ano de 2021 o cruzamento não se apresentar como um gargalo, com uma das fila maior que 40 veículos, será necessário estressar o sistema e descobrir o seu limite.

Os resultados dos diferentes cenários foram utilizados para sugerir um melhor dimensionamento dos processos, e concluir se mesmo com um sistema melhor dimensionado o investimento em um viaduto é necessário.

## 4. RESULTADOS E DISCUSSÃO

### 4.1 VALIDAÇÃO DO MODELO

A validação do modelo consiste em verificar se o sistema computacional é confiável para representar o sistema real. Para isso, compararam-se os dados de saída em

relação ao número de trens, carros, caminhões nos sentidos exportação e importação gerados pelo modelo de simulação com os dados históricos observados em 2015, e calcula-se o erro relativo com a seguinte equação:

$$\% \text{ erro} = \frac{|x_0 - X|}{x_0} \times 100 \quad (2)$$

em que,

% erro = Erro ou desvio relativo percentual;

$X_0$  = Valor desejado;

$X$  = Valor obtido por meio da simulação.

O sistema foi simulado com os seguintes parâmetros: duração de 365 dias (considerando 24 horas por dia) e 8 replicações. O número de replicações foi calculado de acordo com a equação (1), considerando um intervalo de confiança de 95% para todos os indicadores de tempo dos veículos em fila no sistema. Esse número de replicações garantiu uma precisão dos

indicadores analisados menor que 10% da média amostral.

A quantidade média de entidades criadas por meio da simulação foi confrontada com os dados históricos a fim de calcular o erro relativo por meio da equação (2). Observe os resultados de 2015 nas tabelas 1, 2 e 3 a seguir:

Tabela 1 – Erro relativo entre o número de caminhões nos sentidos importação e exportação em 2015

| Ano  | Exportação            |                       |                 | Importação            |                       |                 |
|------|-----------------------|-----------------------|-----------------|-----------------------|-----------------------|-----------------|
|      | # Caminhões Simulados | # Caminhões Real 2015 | Erro Relativo % | # Caminhões Simulados | # Caminhões Real 2015 | Erro Relativo % |
| 2015 | 76.244                | 72.360                | 5,37%           | 76.254                | 72.360                | 5,38%           |

Fonte: Elaboração própria

Tabela 2 – Erro relativo entre o número de carros nos sentidos importação e exportação em 2015

| Ano  | Exportação         |                    |                 | Importação         |                    |                 |
|------|--------------------|--------------------|-----------------|--------------------|--------------------|-----------------|
|      | # Carros Simulados | # Carros Real 2015 | Erro Relativo % | # Carros Simulados | # Carros Real 2015 | Erro Relativo % |
| 2015 | 21.224             | 21.600             | 1,74%           | 21.163             | 21.600             | 2,02%           |

Fonte: Elaboração própria

Tabela 3 – Erro relativo entre o número de trens nos sentidos importação e exportação em 2015

| Ano  | Exportação        |                   |                 | Importação        |                   |                 |
|------|-------------------|-------------------|-----------------|-------------------|-------------------|-----------------|
|      | # Trens Simulados | # Trens Real 2015 | Erro Relativo % | # Trens Simulados | # Trens Real 2015 | Erro Relativo % |
| 2015 | 1.090             | 1.080             | 0,95%           | 1.097             | 1.080             | 1,06%           |

Fonte: Elaboração própria

Observou-se que o erro relativo percentual anual foi inferior a 5,5% em relação à criação de caminhões, inferior a 2% em relação à criação de carros e inferior a 1,7% em relação à criação de trens. Todos inferiores a 10% que é o parâmetro de referência da empresa.

Utilizou-se também, da animação do modelo de simulação no software arena para a validação. Observou-se que a animação dos processos do modelo de simulação condizem com os processos observados e esperados no cruzamento.

Portanto, pode-se considerar o modelo válido para representar a realidade do sistema real, visto que a animação do modelo condizia com os processos do sistema, e que os erros relativos obtidos são valores aceitáveis de acordo com o padrão trabalhado pela empresa.

#### 4.2 RESULTADOS DO CENÁRIO BASE

O cenário base consistiu, primeiramente, na simulação do modelo utilizando os dados de 2017 como *inputs* e, posteriormente, os dados de 2019. Ambas simulações foram feitas conforme descrito na metodologia.

Observou-se que as quantidades de carros, caminhões e trens que passam pelo cruzamento condizem com as quantidades dessas mesmas entidades previstas no orçamento de 2017 da empresa com um desvio relativo anual inferior a 7,42%. O menor desvio relativo anual observado foi em relação ao número de carros sendo inferior a 2,2%.

De acordo com esse cenário de 2017 e suas criações, temos como saída a ocupação média anual do cruzamento de 30%, e o número máximo de veículos em fila igual a 23 na exportação e 21 na importação.

Portanto, constata-se por meio dos resultados da simulação do ano 2017 que o cruzamento ainda não é considerado um gargalo por não ter excedido sua capacidade de fila máxima de 40 veículos. Assim, conclui-se que neste ano o investimento em obras de aumento de capacidade não é ainda necessário.

Simulou-se, então, o modelo considerando as taxas de geração de caminhões e trens previstas para o ano de 2019. Os resultados das criações de caminhões e carros que passam pelo cruzamento apresentam um desvio relativo inferior a 5,3%. Entretanto, o desvio relativo referente ao número de trens

foi em torno de 20%. Vale lembrar que a contagem de criações utilizada no cálculo do desvio compara o número de trens simulados e previstos está situada após o cruzamento, portanto, são contabilizados todos os trens que passam pelo cruzamento e não todos criados pelo sistema.

Sabe-se que de acordo com o orçamento de 2019, o número esperado de trens são 18,9 pares por dia, e com a simulação obteve-se uma média de apenas 15 pares por dia que passam pelo cruzamento e uma média de 18,9 pares criados por dia.

Portanto, conclui-se que a capacidade da ferrovia entre os pátios origem e destino não permite a passagem de mais do que 15 pares de trens por dia no cruzamento. Uma vez que, ao diminuir o intervalo entre as criações o número de trens criados aumentam, mas o número de trens que passam pelo cruzamento se estabiliza em 15 pares por dia, e a fila de trens na origem importação e origem exportação, esperando para passar pela ferrovia, aumenta significativamente. E confirma-se que a capacidade da ferrovia limita o fluxo de trens no cruzamento ao calcular a divisão entre as horas operacionais da ferrovia e o tempo que um trem gasta para percorrer o trecho.

Mesmo com o número de trens estagnados em 15 pares por dia no modelo de 2019, e não 18,9 pares como esperado no estudo da ferrovia, impossibilitando a análise do escopo principal do projeto, que visa analisar o cruzamento com o aumento de trens e caminhões durante os próximos anos. Há no modelo um aumento no número de caminhões e, portanto, observa-se que a ocupação média anual do cruzamento é de 32,4% e que a fila máxima é igual a 28 na exportação e 22 na importação.

Conclui-se que, a simulação do cenário base de 2019, assim como no de 2017, o cruzamento não é considerado gargalo no sistema uma vez que a fila máxima não ultrapassa a capacidade de 40 veículos.

#### 4.3 CENÁRIO 1 – AUMENTO DA CAPACIDADE DA FERROVIA

Considerando os resultados da simulação do cenário base não foi possível chegar ao objetivo do trabalho. Nem concluir que há um gargalo no cruzamento devido à preferência do trânsito de trens na ferrovia, ou definir para

quando um investimento em infraestrutura deve ser provisionado para que a fila de veículos seja inferior a 40. Portanto, neste cenário, simulou-se um aumento na capacidade de ferrovia já que ela estava limitando a capacidade média de 15 trens por dia no cruzamento.

Ao aumentar a capacidade da ferrovia para que ela atenda a demanda de trens prevista no cruzamento para os próximos anos. Simulou-se o modelo com os dados de 2019 e 2021 a fim de verificar os reais reflexos que o aumento de número de trens e caminhões causa na taxa de ocupação do cruzamento.

Uma das formas utilizadas na prática para o aumento da capacidade da ferrovia é a construção de pátios de cruzamento. Os pátios de cruzamento se instalados exatamente entre a origem e o destino da ferrovia duplicam a capacidade da mesma. Assim, enquanto anteriormente um trem deveria esperar em fila um outro passar por toda a ferrovia para depois seguir viagem, depois da construção do pátio de cruzamento, os trens, um na origem e outro no destino, podem se deslocar ao mesmo tempo e com velocidade semelhante até o pátio de cruzamento. Ao chegar no meio do trecho, onde o pátio está, há uma linha duplicada que permitirá que os trens em sentidos diferentes possam passar ao mesmo tempo. O cálculo de aumento na capacidade da ferrovia se baseia na teoria de Colson (KRUEGER, 1999). A representação desse aumento de capacidade com a criação do pátio pode ser feita no modelo de simulação com a diminuição dos tempos de trânsito pela metade, então o trem não utilizará na média 40 minutos para percorrer todo o trecho, mas sim 20 minutos.

Após o aumento da capacidade da ferrovia foi simulado o ano de 2019 e observou-se que o modelo gerou as quantidades adequadas de veículos e trens, apresentando um desvio relativo anual inferior à 5,1% na criação das entidades. Com a taxa média anual de ocupação igual a 38,1%, e as filas máximas com 25 veículos na exportação e 24 na importação.

Logo, o cruzamento não é um gargalo em 2019. Então, simulou-se o modelo considerando os dados de 2021 conforme descrito na metodologia, e a melhoria de capacidade na ferrovia. Os resultados das criações no cruzamento apresentaram um desvio relativo inferior à 5,4% e, portanto, as

criações do modelo estão adequadas. Quanto às análises das saídas, tem-se a taxa de ocupação de 41% e as filas máximas de veículos na exportação de 28 e na importação de 31. Conclui-se que, considerando os dados previstos no orçamento da empresa até o ano de 2021, o cruzamento ainda não é um gargalo já que fila de veículos não supera 40 unidades, e nenhum investimento em infraestrutura será necessário.

#### 4.4 CENÁRIO 2 – ESTRESSANDO O SISTEMA

Este cenário tem como objetivo saber quando o cruzamento será um gargalo uma vez que os demais cenários não apresentaram ele como tal. Para isso, utilizou-se os dados do ano de 2021 para os carros e caminhões e observou-se um desvio inferior a 5,61% nessas criações. E aumentou-se gradativamente o número de pares de trem até que o cruzamento fosse considerado um gargalo. Neste cenário, assim como no anterior, a capacidade da ferrovia foi aumentada para que ela não limite a quantidade de trens gerada durante a simulação.

Neste cenário, a quantidade de trens para que o cruzamento passe a ser um gargalo deve ser uma média de 58 pares de trens por dia. Este valor é maior que o dobro do valor previsto pelas operadoras ferroviárias, no seu estudo de pares de trens diários na ferrovia, para o ano de 2037.

Ao aumentar o número de pares de trens para 58, a taxa de ocupação do cruzamento atingiu um valor médio anual de 94,4%, e o número máximo de veículos em fila chegou a 40 no sentido exportação e a 41 no sentido importação, ultrapassando, portanto, o espaço disponível para o estacionamento prejudicando as operações portuárias.

### 5. CONCLUSÃO

A partir da simulação, analisaram-se se a capacidade do recurso “cruzamento” limita o fluxo de veículos que passa por ele de acordo com a quantidade de veículos e trens previstos no orçamento da empresa de 2017 a 2021. O objetivo foi definir o ano que um investimento em infraestrutura deveria ser provisionado. Para isso, diferentes cenários foram simulados e analisados.

Em cada cenário os resultados obtidos por

meio das simulações e análises permitem concluir que:

a) A análise do cruzamento por meio de um modelo de simulação de eventos discretos foi viável e relevante para prever os possíveis gargalos do sistema nos próximos anos;

b) No cenário base as principais conclusões são que nos anos de 2017 e 2019 com o número médio de trens limitado a 15 pares, o cruzamento não se apresenta como um gargalo no sistema. Obteve-se uma taxa de ocupação do cruzamento de 30,3% em 2017 e 32,4% em 2019. E uma fila máxima de 23 em 2017 e 28 em 2019;

c) No cenário 1, alterou-se a capacidade ferroviária. Considerou-se, em 2019, uma média de 19 pares de trens que resultou em uma taxa de ocupação do cruzamento de 38,1% e uma fila máxima de 25 veículos. E em 2021, com uma média de 20 pares de trens, obteve-se a taxa de ocupação de 41,6% com a fila máxima de 31

veículos;

d) Quanto ao cenário 2, realizou-se o estresse do sistema uma vez que até o ano de 2021 o cruzamento não é considerado um gargalo, ou seja, a fila máxima é inferior à 40 veículos. Utilizou-se as mesmas premissas e parâmetros do ano de 2021 e aumentou-se o número de trens. Observou-se que o número de pares de trens que passam pelo cruzamento deve ser igual ou maior que uma média de 58 pares por dia para que a fila fosse superior a 40 unidades, implicando, assim, a necessidade de investimento.

Como conclusão final, tem-se que a técnica de simulação de eventos discretos se mostrou adequada e eficaz para o estudo do comportamento do sistema analisado. A eficácia desta técnica foi comprovada por meio da validação do modelo de simulação que teve por base dados históricos da empresa.

## REFERÊNCIAS

- [1] ANTP – ASSOCIAÇÃO NACIONAL DE TRANSPORTES PÚBLICOS. Transporte Humano: Cidades com Qualidade de Vida. São Paulo: ANTP, 1997.
- [2] ARAUJO, M.A. Administração de produção e operações – Uma abordagem prática. Rio de Janeiro: Brasport, 2009.
- [3] BANKS, J. Handbook of simulation – Principles, methodology, advances, applications, and practice. Georgia: Institute of Technology, 1998.
- [4] CHENG, L.; DURAN, A.M. Logistics for world-wide crude oil transportation using discrete event simulation and optimal control. Computers & Chemical Engineering. Vol.28, n.6-7, p.897-911, 2004.
- [5] DENATRAN. Departamento Nacional de trânsito, disponível em: <<http://www.denatran.gov.br/index.php/estatistica/237-frota-veiculos>>. Acesso em: 16 de Março de 2017.
- [6] FGV – FUNDAÇÃO GETULIO VARGAS. A Construção do Desenvolvimento Sustentado: A importância da construção na vida econômica e social do país. São Paulo: FGV Projetos e União Nacional da Construção, 2006.
- [7] FREITAS F., P. J. Introdução à modelagem e simulação de sistemas: com aplicações em Arena.. 2. ed. Florianópolis: Visual Book, 2008.
- [8] GUIZZI, G.; MURINO, T.; ROMANO, E. A Discrete Event Simulation to model Passenger Flow in the Airport Terminal. Mathematical Methods and Applied Computing. Vol.2, n.1, p.427-434, 2009.
- [9] GUNAL, M.M.; PIDD, M. Discrete event simulation for performance modelling in health care: a review of the literature. Journal of Simulation. Vol.4, n.1, p.42-51, 2010.
- [10] HILL, R.R.; MILLER, J.O.; MCINTYRE, G.A. Simulation analysis: applications of discrete event simulation modeling to military problems. 33rd conference of Winter simulation, 2001.
- [11] INGALLS, G.R. Introduction to Simulation. 40th Winter Simulation Conference, 2008.
- [12] KELTON, W.D.; SADOWSKY, R.P.; SADOWSKY D.A. Simulation with Arena.. 5 Ed. Boston: McGraw-Hill, 2009.
- [13] KRUEGER, H.; Parametric modeling in rail capacity planning. 31st Winter Simulation Conference. New York, 1999.
- [14] LACERDA, S.M. Precificação de congestionamento e o transporte coletivo urbano. Ministério do Desenvolvimento, indústria e comércio exterior. Rio de Janeiro: BNDES Setorial, 2006.
- [15] LANNONI, A.P.; MORABITO, R. A discrete simulation analysis of a logistics supply system. Elsevier. Vol.42, n.3, p.191-210, 2006.
- [16] LITMAN, T. Introduction to multi-modal

transportation planning Vol.15. Victoria: Victoria transport policy institute, 2011.

[17] OLIVEIRA, A.V.M.; TUROLLA, F.A. Financiamento da infraestrutura de transportes. Journal of transport literature Vol.7, n.1 , p.103-126, 2013.

[18] PRADO, D.S.D. Usando Arena em Simulação. Série Pesquisa Operacional. Volume 3. Editora DG-Desenvolvimento Gerencial. Belo Horizonte: Falconi, 1999.

[19] RIBEIRO, P.C.C.; FERREIRA, K.A. Logística e transportes: Uma discussão sobre os modais de transporte e o panorama brasileiro. XXII Encontro Nacional de Engenharia de Produção em Curitiba, 2002.

[20] SHANNON, R.E. Systems simulation. The Art and Science.. IEEE Transactions on Systems, Man, and Cybernetics. Vol.6, n.10, p.723-724,

1976.

[21] SARGENT, R.G. Verification and validation of simulation models. Winter Simulation conference, 2011.

[22] SILVA, A.K.; BOTTER, R.C. Method for assessing and selecting discrete event simulation software applied to the analysis of logistic systems. Journal of Simulation. Vol.3, n.2, p.95-106, 2009.

[23] USDT, United States department of transportation. The transportation planning process key issues – a briefing book for transportation decision makers, officials, and staff. U.S.: Federal Highway Administration, 2007.

[24] VASCONCELOS, G.M.; BASSO, L. Sistemas de transportes brasileiros: origens, estado atual e desafios. Jovens pesquisadores Mackenzie. Vol.5, n.8, p.141-158, 2008.

# Capítulo 25

## APLICAÇÃO DA TÉCNICA DE SIMULAÇÃO COMPUTACIONAL EM UMA EMPRESA DO SETOR DE PANIFICAÇÃO

*Jarbas Ancelmo da Silva Júnior*

*Natália Cibele de Sousa Santos*

*Nayara Felício de Oliveira*

*Lara Cristine Pereira dos Santos*

**Resumo:** jCondigno ao cenário altamente competitivo em que as indústrias estão inseridas, é de extrema importância que as empresas adotem ferramentas que lhes auxiliem nas tomadas de decisões. Neste contexto, a simulação de sistemas é uma ótima ferramenta para auxiliar as empresas sem que seja necessário realizar mudanças em seu espaço físico. Com base nesta conjuntura, o objetivo do presente trabalho é simular o comportamento de um sistema real, utilizando as técnicas de simulação de sistemas, a fim de propor melhorias ao desempenho do processo e das operações necessárias para realização da fabricação de pães de queijo. Para isso, empregou-se o procedimento de pesquisa experimental devido ao uso de simulação. Foram modelados e simulados dois cenários, um cenário atual e um proposto, verificou-se, com isso, que o cenário proposto apresentou uma melhoria no total médio de saídas do sistema. Este trabalho contribuiu para um maior entendimento e divulgação acerca do uso de simulação de sistemas, como uma ferramenta que auxilia as empresas na tomada de decisões.

**Palavras chave:** Simulação de sistemas, Promodel®, panificadora.

\* Artigo originalmente apresentado no Congresso Brasileiro de Engenharia de Produção Conbrepro (2017)

## 1. INTRODUÇÃO

A Engenharia de Produção evoluiu-se no decorrer do século XX. Isso ocorreu, principalmente, como um retorno a importância de uma busca constante, por melhoria nos procedimentos e técnicas que envolvem o espaço produtivo, além, da busca por novos diferenciais competitivos, como por exemplo, a preocupação em atender as exigências e necessidades do cliente (ABREPO, 2002).

Com um foco em ferramentas e métodos que possibilitem a otimização e uma melhor utilização dos recursos produtivos, a Engenharia de Produção é subdividida em várias áreas, intitulada como grande área. Ou seja, o curso é organizado em vários ramos do conhecimento, e essa interdisciplinaridade contribui com um melhor desenvolvimento produtivo (ABREPO, 2002).

Dentro da chamada grande área da Engenharia de Produção está a Simulação de Sistemas que é muito importante, principalmente quando se deseja reproduzir um processo real. A Simulação de Sistemas segundo Banks (2000), é uma reprodução do desempenho de um sistema concreto, em um período estabelecido, e inclui a elaboração similar do sistema real, objetivando tirar resultados e conclusões do processo executado.

De acordo com Freitas Filho (2008), a simulação tem sido cada vez mais utilizada e reconhecida como um método que auxilia profissionais de diferentes meios como, por exemplo, engenharia, biologia, dentre outros, a encontrarem soluções dos desafios encontrados cotidianamente. Essa técnica é adotada por vários autores como Buffa e Sarin (1987), Law e Kelton (2000), Berends e Romme (1999) Chwif e Medina (2007).

A empresa a ser estudada é uma Panificadora que atua no mercado em aproximadamente quatro anos. Segundo o Instituto Tecnológico da Panificação e Confeitaria (2017), o ramo das panificadoras vem crescendo nos últimos anos e apresenta uma importante potencialidade tanto socialmente como economicamente, fornecendo trabalho e renda para a comunidade.

Desta forma, o objetivo geral do trabalho é aplicar as técnicas envolvendo a simulação de sistemas, analisando um ambiente real, nesse caso, uma Panificadora, a fim de buscar e propor melhorias quanto ao

desempenho do processo e das operações necessárias para fabricação de pães de queijo.

A estruturação do trabalho será organizada da seguinte forma: a segunda seção, referente à revisão bibliográfica sobre Simulação de Sistemas e sobre a área de aplicação do modelo; a terceira seção, os métodos e procedimentos realizados na pesquisa; quarta seção, os resultados obtidos das análises e a discussão sobre os mesmos; quinta seção serão apresentados os erros encontrados durante a elaboração do modelo; sexta seção será descrito as dificuldades encontradas ao longo do trabalho e por fim será encerrada com a sétima seção, apresentando as conclusões.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 SIMULAÇÃO DE SISTEMAS

Organizações dos mais variados portes vem fazendo uso da técnica de simulação, por ser um método que as auxiliam nas tomadas de decisão. Encontra-se na literatura da área diversas formas de definir simulação. Harrel et al (2002) a descreve como a importação da realidade, através de um modelo detalhado, lógico e matemático para um ambiente controlável, com o objetivo de visualizar como este sistema se comportará diante às mudanças realizadas em sua estrutura, ambiente ou condição de contorno.

Harrell, Ghosh e Bowden (2010) define a simulação de sistemas como uma ferramenta de suporte a decisão, podendo ser aplicada em diversas áreas da Engenharia de Produção.

Chwif e Medina (2007) definem um modelagem de simulação como uma compreensível representação de um sistema físico real, o qual, pode-se determinar as interações e relações das variadas partes que constituem o sistema.

Qualquer sistema pode ser representado e simulado através de um modelo, entretanto, quanto mais complexo for o sistema, mais tempo será necessário para realização da modelagem e simulação. Isto posto, tem-se a necessidade de definir os termos "sistema" e "modelo".

Law e Kelton (2000), definem um sistema como um conjunto de entidades, pessoas ou máquinas, que operam e interagem juntas por



meio de suas atribuições visando atingir um objetivo específico. Conforme o mesmo autor, um estado do sistema pode ser compreendido como um agrupamento de variáveis necessárias para retratar um sistema em um intervalo de tempo.

Bratley, Bennet e Schrage (1987) definem modelo como uma descrição de sistemas com o intuito de antever o que acontecerá quando alguma decisão é tomada. Portanto o modelo deve ser válido e estar harmônico com o sistema. Para Law e Kelton (2000), um modelo pode ser compreendido como uma representação de um sistema com um objetivo de estudo. O modelo deve estar suficientemente detalhado, de uma forma que permita que o analista tome decisões que possam ser aplicadas no sistema real.

Freitas Filho (2008) enumera algumas vantagens da aplicação de um estudo de simulação:

- i) facilidade de aplicação quando comparada aos métodos analíticos;
- ii) possibilidade de identificação de gargalos;
- iii) evita retrabalhos irrelevantes, por ser um método que permite analisar antecipadamente como o sistema se comportará diante das possíveis mudanças a serem realizadas;

Além das diversas vantagens de sua aplicação, Law e Kelton (2000) expõem algumas de suas desvantagens:

- i) Os modelos de simulação consomem tempo para serem desenvolvidos;
- ii) Um estudo de simulação fornece uma grande quantidade de números, fazendo com que as pessoas confiem no modelo mais do que deveriam. Pois, se não houver a representação exata do sistema, as informações obtidas terão uma tênue utilidade;
- iii) Modelar um sistema requer treinamento especializado, porém, a confiabilidade do estudo e qualidade do modelo dependem das habilidades e capacidade do modelador.

Para Saliby (1989) a abordagem de estudo da simulação apresenta uma grande expansão e tem sido cada vez mais utilizada em diversas áreas de conhecimento, o autor acredita que tal propagação corresponde ao fato da maior

disponibilidade de recursos computacionais e da crescente complexidade dos problemas.

Indubitavelmente, a simulação computacional vem retratando constante evolução, principalmente por causa dos avanços nas áreas de hardware e software. Como afirmado por Chwif e Medina (2015), a simulação de sistemas é reproduzida por um modelo computacional, através de linguagens de simulação, programação ou uso de um simulador. Existem diversos pacotes de simulação, os quais, possuem o objetivo de diminuir a lacuna entre o entendimento que o usuário tem do modelo e sua forma executável. Arena<sup>®</sup>, AweSim<sup>®</sup>, SimFactory e ProModel<sup>®</sup> são alguns dos pacotes de simulação existentes. Dentre estes, o Promodel<sup>®</sup> será o software de simulação utilizado para realização do modelo apresentado neste trabalho.

Antes da construção do modelo computacional e utilização do simulador, Freitas Filho (2008) sugere que seja elaborado um modelo conceitual, o qual, trata-se de uma representação do funcionamento do sistema, bem como, representação de seus componentes, variáveis e interações lógicas. A modelagem conceitual reduz o tempo gasto na fase de construção dos modelos computacionais e aumenta a qualidade dos modelos de simulação.

### 3. METODOLOGIA

A abordagem de pesquisa utilizada foi a quantitativa. Segundo Richardson (1999) a abordagem quantitativa é caracterizada pelo emprego da quantificação, tanto na coleta de dados quanto no tratamento das informações, aplicando-se técnicas estatísticas.

Quanto ao procedimento de pesquisa, utilizou-se a pesquisa experimental, tal procedimento, segundo Creswell (1994) busca verificar as relações de causa e efeito de um determinado fato, e ao mesmo tempo possibilita que o pesquisador manipule as variáveis independentes. Segundo Bryman (1989) este procedimento é umas metodologias de pesquisa mais aconselháveis para abordagens quantitativas. Segundo Ferreira, Hochman e Barbosa (2005), o modelo experimental deve representar o objetivo de estudo da forma mais exata possível, por isso, a pesquisa experimental emprega algumas ferramentas, tais como: modelagens matemáticas,

experimentos controlados em laboratório e simulações computacionais. Desta forma, utilizou-se este procedimento devido ao uso de simulação.

Empregou-se técnicas matemáticas e computacionais para simulação do sistema de produção de pães de queijo. Para Law e Kelton (2000), a empresa, objeto no nosso estudo é conceituada como um sistema terminal, pois apresenta condições iniciais fixas e um evento que determina um fim natural para o processo de simulação.

A simulação computacional foi realizada com o auxílio do Software Promodel® versão Student 2010, o qual foi escolhido em função da familiaridade entre o programa e os pesquisadores.

Para facilitar a elaboração do modelo computacional, elaborou-se um modelo conceitual, através da técnica IDEF-SIM, desenvolvida por Leal, Almeida e Montevechi (2008). Segundo, Sargent (2007), a modelagem conceitual deve ser feita visando facilitar a visualização e interpretação do sistema em estudo.

Diante da complexidade de um estudo de simulação, Law e Kelton (2000) sugerem uma proposta de estudo constituída por 10 etapas, as quais auxiliaram nosso estudo:

1. Formulação do problema e planejamento do estudo;
2. Coleta de dados e definição do modelo;
3. Validação do modelo;
4. Construção e verificação do programa computacional;

5. Realização de execuções piloto;
6. Validação do modelo programado;
7. Projeto dos experimento;
8. Realização das execuções de simulação;
9. Análise de resultados;
10. Documentação, apresentação e implementação dos resultados.

Os dados foram coletados na empresa de panificação durante dois dias consecutivos. Para realização das observações e cronometragem dos tempos a produção de pães de queijo foi dividida em 4 atividades, a saber: (preparação da massa, modelagem da massa, tempo de forno e deslocamento do forno até o expositor). O tempo da última atividade foi considerado como constante de 5 minutos para melhor visualização da simulação.

Foram realizadas 15 observações de cada atividade, De acordo com Barnes (1977), através da equação 1, considerando um nível de confiança de 95% e um erro relativo de 5%, quando  $N > N'$  pode-se aferir que o número de observações realizadas garantem a confiabilidade dos dados. Dessa forma, ao aplicar a equação 1 para as atividades cronometradas, foi possível identificar que  $N'$  apresentou valor menor que 15. Garantindo a confiabilidade dos dados coletados. Os valores utilizados na equação 1, bem como o valor dos  $N'(s)$  estão representados no apêndice A.

$$N' = \left( \frac{40 \sqrt{[N \Sigma X^2 - (\Sigma X)^2]}}{\Sigma X} \right) \quad (1)$$

Onde:

$N'$  = número de observações necessárias

$N$  = número de observações realizadas;

$X$  = valor da amostra.

Para este estudo definiu-se um nível de confiança de 95%. Segundo Kelton, Sadowski, Sadowski (2002), para obtenção de um intervalo de confiança no valor de 95% é necessário que o resultado médio da simulação, em 95% das replicações, esteja

dentro do intervalo da média obtida  $\pm$  o semi-intervalo (*half-width*).

Utilizou-se o software Sketchup® para produzir o *layout* da empresa e outras figuras utilizadas como apoio à biblioteca gráfica. O modelo foi validado e os resultados fornecidos pelo

*Output Viewer*, ferramenta integrada ao software ProModel® foram avaliados.

## 4. RESULTADOS E DISCUSSÕES

### 4.1 INFORMAÇÕES GERAIS NO MODELO SIMULADO

O estudo foi realizado em uma empresa do ramo de panificação. Sendo comercializado na mesma, grande variedade de quitandas, como, roscas, pães, broas de milho, pães de queijo, dentre outros. Porém, a partir de uma análise prévia dos mesmos e entrevistas com a proprietária foi possível constatar o produto mais vendido: o pão de queijo. Desta forma, o trabalho terá como foco, a produção desta entidade. O processo produtivo é conduzido por uma “cozinheira”, responsável por toda as etapas de desenvolvimento. A produção do “pão de queijo” é realizada da seguinte forma, primeiro a padaria coleta a “matérias prima” que será utilizada no processo, deslocando-a até a “maseira”, onde haverá a junção da matéria prima originando a “massa”, posteriormente, esta massa é transportada pela cozinheira até o “posto de trabalho 2” para serem modelados no formato desejado e prontos para serem assados, formando assim a “massa”, consistindo neste o penúltimo processo, levando-os até o “forno” para serem assados durante 30 minutos. Como derradeira etapa, há o transporte do produto acabado, pão de queijo, até o “expositor”, prontos para serem comercializados. Portanto, o processo consiste de um total de cinco locais e uma padaria sendo está representada como recurso “cozinheira” no modelo elaborado. Para uma visualização simplificada do modelo há a modelagem conceitual, vide Figura 1. E, também, a tabela com os dados coletados no Apêndice A.

Como a maior demanda pelo produto é realizada no período matinal, e pela limitação do espaço físico e dos equipamentos utilizados são feitas quantidades específicas de pães de queijo ao longo da manhã, portanto, analisou-se um período de 5 horas. Considerando, ainda, os pequenos descansos de 10 minutos entre as produções. A unidade de tempo e de distância utilizadas foram minutos e metros, respectivamente.

### 4.2 ETAPAS PARA CONDUÇÃO DE UM ESTUDO DE SIMULAÇÃO

Foram utilizadas as etapas propostas por Law e Kelton (2000) a fim de elaborar este estudo de simulação, as quais estão apresentadas abaixo:

**Etapa 1** – Nesta etapa, formulação da problemática e planejamento do estudo, foram designados os responsáveis por cada atividade, além das datas a serem respeitadas para entrega das respectivas tarefas, foi definido o objetivo a ser atingido através do presente estudo: simular o processo de fabricação de pão de queijo de uma empresa de panificação, posteriormente, o planejamento do estudo foi apresentado à responsável pela disciplina.

**Etapa 2** – Na parte de coleta de dados e definição do modelo, foram realizadas 15 cronometragens de cada etapa do processo, além dos tempos de deslocamentos da cozinheira, com o propósito de colher dados que garantam uma certa confiança nas amostragens. A fim de encontrar a melhor distribuição probabilística a ser utilizada, foram utilizadas bibliografias da área, sendo encontrada a distribuição triangular a que melhor representaria o modelo. Que de acordo com Chwif e Medina (2015) é utilizado quando há o conhecimento ou se tem uma estimativa de um valor mínimo, de moda e de máximo apresentados no sistema. Dados estes presentes de forma clara nos dados coletados, apresentados no Apêndice A. Nesta etapa, ainda, foi elaborado o modelagem conceitual, por meio da técnica IDEF-SIM, com o auxílio do software DIA®, disponibilizado no Apêndice B.

**Etapa 3** – Para a validação do modelo, houve uma pesquisa aprofundada acerca da elaboração da modelagem conceitual, sendo executado com o auxílio de todos os integrantes do estudo, e, posteriormente, apresentado à docente onde foi proposta as adequações pertinentes.

**Etapa 4** – A partir do modelo conceitual elaborado e validado, foi possível construir o programa computacional e verificá-lo. Para produzir o *layout* da empresa e outras figuras utilizadas como apoio à biblioteca gráfica, utilizou-se o software Sketchup®. Já para a modelagem computacional utilizou-se o software ProModel®, versão Student 2010.

**Etapa 5** – Foram realizadas execuções piloto, com o propósito de testar a funcionalidade do

modelo e, também, realizar possíveis inadequações e correções necessárias. Verificando se as distribuições utilizadas representaram o modelo de forma realística.

**Etapa 6** – Com o intuito de validar o modelo, utilizou-se os conhecimentos adquiridos em aulas práticas e materiais teóricos, além, do auxílio da docente responsável, a qual possui amplo conhecimento no software utilizado, posteriormente, o modelo foi apresentado a todos os integrantes do grupo para verificação de possíveis incoerências com a realidade da empresa estudada. E, ainda, foram testadas e executadas a quantidades de replicações adequadas.

**Etapa 7** – No passo de projeto dos experimentos, foram definidos as perícias a serem executadas, sendo optado em realizar de acordo com sistema real da empresa e, também, uma outra execução, porém, usando a capacidade máxima do “Posto de trabalho 2” e utilizando um outro forno para depois ser comparado e verificado o possível aumento na produtividade. Utilizou-se uma duração de 5 horas de trabalho e um número de 8 ocorrências, ou seja, dentre os dois parâmetros utilizados, o que ocorrer primeiro será considerado.

**Etapa 8** – Realização das execuções de simulação, sendo feitas várias replicações da simulação a fim de que os dados encontrados sejam usados na validação do modelo computacional. E encontrar os dados a serem utilizados na etapa seguinte.

**Etapa 9** – Nesta etapa, realizou-se a análise dos resultados, ao final das 30 replicações realizadas, utilizando um índice de confiança de 95% para indicar o intervalo de confiança de cada medida de desempenho, sendo

necessário um valor menor ou igual a 10% da média amostral.

**Etapa 10** – Documentação, apresentação e implementação dos resultados, nesta etapa é realizada a documentação de forma adequada do estudo a fim de contribuir para um melhor entendimento do trabalho executado, facilitar modificações futuras, além de promover uma maior credibilidade do trabalho realizado. Como a licença do software indica fins educativos e não lucrativos a parte de implementação foi desconsiderada. Os dados obtidos serão apresentados nas seções seguintes.

#### 4.3 PRINCIPAIS RESULTADOS ENCONTRADOS NA SIMULAÇÃO, CONSIDERANDO CENÁRIO ATUAL

Para a análise dos valores médios alcançados nos relatórios, foram efetuadas 30 replicações e um tempo de simulação de 4,78 horas. Os semi-intervalos de confinça (*half-width*) obtidos foram inferiores a 10% das médias amostrais, o que aponta que a quantidade de replicações realizadas são consideradas satisfatórias, garantindo a precisão dos valores médios encontrados. Os dados obtidos no relatório do *OutPut Viewer* estão apresentados no Apêndice D.

A Tabela 1 expõe os dados relativos a média total de saída da entidade “pão de queijo”, são apresentados, também, os resultados referentes aos valores mínimos, máximos, desvio padrão intervalos de confiança de 95% inferior e superior, além, do *half-width*.

Pode-se inferir que ao longo das 4,78 horas de simulação, foram processados em média 1280 pães de queijo.

Tabela 1. Total de saídas de “pão de queijo” do sistema, no cenário atual

| Entidade      | Média | Máximo | Mínimo | Desvio padrão | I.C. de 95% superior | I.C. de 95% confiança inferior | <i>Half-width</i> |
|---------------|-------|--------|--------|---------------|----------------------|--------------------------------|-------------------|
| Pão de queijo | 1280  | 1280   | 1280   | 0             | 1280                 | 1280                           | 0                 |

Vide Tabela 2, pode-se observar os tempos médios no sistema (que é a duração desde a entrada da entidade no sistema até ser transformada em outra ou até sair do sistema) para as duas entidades utilizadas na

simulação, onde a entidade “massa”, possui um tempo médio de 52 minutos podendo variar 0,13 para mais ou para menos. E, também, a entidade “pão de queijo” com um tempo médio de 5,23 minutos.

Tabela 2. Tempo médio no sistema das entidades (em minutos).

| Entidade      | Média | Máximo | Mínimo | Desvio padrão | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|---------------|-------|--------|--------|---------------|----------------------|----------------------|------------|
| Massa         | 52    | 52,56  | 51,37  | 0,20          | 52,13                | 51,87                | 0,13       |
| Pão de queijo | 5,23  | 5,23   | 5,22   | 0             | 5,23                 | 5,23                 | 0          |

Analisando a Tabela 3, é possível ver a taxa de utilização dos locais. Indicando “Forno” como o local com maior taxa de uso, em que durante as 4,78 horas simuladas esteve 84,78% sendo utilizado, podendo variar 0,05 para mais ou para menos. O local com menor utilização foi o estoque inicial com 0,06% e tendo a possibilidade de variar 0,01 para mais

ou para menos, este valor deve-se à configuração de simulação, sendo programada sua chegada para a cada 30 minutos. O expositor teve um uso ínfimo, também, com uma média 1,12%, pois conforme os pães de queijo chegam ao local são instantaneamente vendidos, devido à alta demanda do período simulado.

Tabela 3. Taxa de utilização dos locais no atual cenário da empresa.

| Local               | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|---------------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Estoque inicial     | 0,06      | 0,08       | 0,03       | 0,02              | 0,06                 | 0,05                 | 0,01       |
| Posto de trabalho 1 | 24,43     | 24,65      | 24,09      | 0,16              | 24,49                | 24,37                | 0,06       |
| Masseira            | 44,50     | 45,97      | 42,82      | 0,93              | 44,85                | 44,15                | 0,35       |
| Forno               | 84,78     | 85         | 84,53      | 0,13              | 84,83                | 84,73                | 0,05       |
| Expositor           | 1,12      | 1,12       | 1,11       | 0                 | 1,12                 | 1,12                 | 0          |

Quanto ao recurso “cozinheira” foi analisado a sua taxa de utilização, pois como há apenas um operário havia a preocupação em relação a este dado. Porém, analisando o relatório do *Output Viewer* foi possível verificar uma taxa média de uso de 6,30% com uma variabilidade de 0,04 para mais ou para

menos. Ou seja, durante as 4,78 horas analisadas, a padeira esteve sendo utilizada em média 6,30% do tempo. Este valor deve-se ao fato do recurso simulado ser responsável apenas pelos deslocamentos das entidades no sistema.

Tabela 4. Utilização do recurso “cozinheira” no cenário atual.

| Recurso    | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Cozinheira | 6,30      | 6,46       | 6,18       | 0,10              | 6,34                 | 6,26                 | 0,04       |

Outro dado interessante a ser analisado do recurso “cozinheira” é a taxa de ociosidade da mesma, haja vista a Tabela 5, onde pode ser analisado que em média 92,35% do

tempo programado ela ficou ociosa, podendo variar, para menos ou mais, 0,04. Legitimando a proposta de melhoria proposta a ser apresentada na seção a seguir.

Tabela 5. Taxa de ociosidade do recurso “cozinheira” no cenário atual.

| Recurso    | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Cozinheira | 92,32     | 92,43      | 92,15      | 0,10              | 92,35                | 92,28                | 0,04       |

#### 4.4 PROPOSTA DE MELHORIA

A fim de aumentar a produtividade da empresa estudada, é proposto como melhoria

produzir o dobro de massa dos atuais 2 quilogramas para 4 quilogramas, utilizando a capacidade máxima da masseira e a ativação de um segundo forno já existente no

estabelecimento. Utilizando os mesmos parâmetros do cenário atual de 5 horas de simulação ou 8 ocorrências, prevalecendo o primeiro a ocorrer, encontramos os seguintes resultados, que estão expostos nas tabelas a seguir.

#### 4.4.1 PRINCIPAIS RESULTADOS ENCONTRADOS NA PROPOSTA DE MELHORIA

Para o modelo de melhoria foram utilizadas as mesmas 30 replicações, analisando as mesmas medidas de desempenho, média,

máximo, mínimo, desvio padrão e intervalos de confiança de 95%. Apresentando em todos os casos *half width* inferior a 10% da média amostral.

Analisando a partir do total de saída das entidades foram encontrados os dados apresentados na Tabela 6, onde foi possível obter 2080 pães de queijo, em 5 horas de programação, obtendo um aumento considerável se comparado ao modelo atual operado pela empresa, justificado pelo aumento de massa a ser produzida, além do aumento em uma unidade de forno para assadura da massa modelada.

Tabela 6. Total de saídas de “pães de queijo” do sistema.

| Entidade      | Média | Máximo | Mínimo | Desvio padrão | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|---------------|-------|--------|--------|---------------|----------------------|----------------------|------------|
| Pão de queijo | 2080  | 2080   | 2080   | 0             | 2080                 | 2080                 | 0          |

A partir da Tabela 7, é possível perceber que o tempo aumentou frente aos tempos encontrados no cenário atual, porém se justifica haja visto que a quantidade produzida agora é mais elevada. No modelo proposto o tempo médio da entidade “massa” seria de 64,22 minutos, podendo variar, para mais ou menos, 0,07 minutos. Lembrando que a massa originará os pães de queijo. Já a entidade “pão de queijo” agora tem um tempo médio de 10,21 minutos.

Tabela 7. Tempo médio no sistema da entidades, no cenário proposto (em minutos).

| Entidade      | Média | Máximo | Mínimo | Desvio padrão | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|---------------|-------|--------|--------|---------------|----------------------|----------------------|------------|
| Massa         | 64,22 | 64,52  | 63,96  | 0,18          | 64,28                | 64,15                | 0,07       |
| Pão de queijo | 10,21 | 10,21  | 10,21  | 0             | 10,21                | 10,21                | 0          |

Pode-se inferir, analisando a Tabela 8, que a taxa de utilização média do forno diminui, fundamentado pela quantidade agora de duas unidades em funcionamento, sendo que agora a utilização média deste ficou em 77,97%, variando 0,12 para mais ou para

menos, e 71,81%, variando 0,06, para “Forno.1” e “Forno.2”, respectivamente. Outro dado interessante é o “Posto de trabalho 2”, segundo mais utilizado com 65,60% normalmente, variando este valor em 0,26.

Tabela 8. Taxa de utilização dos locais no cenário proposto.

| Local               | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|---------------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Estoque inicial     | 0,32      | 0,38       | 0,28       | 0,04              | 0,34                 | 0,30                 | 0,02       |
| Posto de trabalho 1 | 33,78     | 34,44      | 33,07      | 0,36              | 33,92                | 33,65                | 0,14       |
| Posto de trabalho 2 | 65,60     | 64,34      | 66,89      | 0,70              | 65,34                | 65,86                | 0,26       |
| Forno.1             | 77,97     | 78,47      | 77,46      | 0,30              | 78,08                | 77,85                | 0,12       |
| Forno.2             | 71,81     | 72,07      | 71,43      | 0,15              | 71,87                | 71,75                | 0,06       |
| Expositor           | 2,41      | 2,42       | 2,39       | 0,01              | 2,41                 | 2,41                 | 0          |

É possível verificar que o recurso “cozinheira” teve uma taxa de utilização média de 8,99%, podendo variar, este valor em 0,04 para mais

ou menos. Houve um pequeno aumento na taxa de uso do mesmo devido a utilização de um segundo forno, onde agora teria que

abastecer mais um forno com massa e retirá-la como pão de queijo e expô-lo à venda.

Tabela 9. Utilização do recurso “cozinheira” no cenário proposto

| Recurso    | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Cozinheira | 8,99      | 9,25       | 8,90       | 0,11              | 9,03                 | 8,95                 | 0,04       |

Vide Tabela 10, percebe-se que houve uma diminuição na taxa média de ociosidade do recurso, justificado pelo dado anteriormente citado, pois como houve um aumento na taxa média de utilização da mesma, diminuiu o

tempo ocioso, como, também, a taxa ociosa média da mesma. Sendo esta taxa de 90,22%, com uma variância de 0,04 para mais ou menos.

Tabela 10. Taxa de ociosidade do recurso “cozinheira” no cenário proposto.

| Recurso    | Média (%) | Máximo (%) | Mínimo (%) | Desvio padrão (%) | I.C. de 95% superior | I.C. de 95% inferior | Half-width |
|------------|-----------|------------|------------|-------------------|----------------------|----------------------|------------|
| Cozinheira | 90,22     | 90,28      | 90,00      | 0,10              | 90,26                | 90,18                | 0,04       |

#### 4.4.2 COMPARAÇÃO DE CENÁRIOS

Ao comparar os resultados obtidos através dos dois cenários simulados, exposto na Tabela 11, foi possível perceber que, primeiramente, o total de saídas aumentaram consideravelmente, garantindo a proprietária uma maior produtividade de pães de queijo,

um acréscimo de 800 unidades do produto. Entretanto, aliado a este resultado houve um pequeno acréscimo na taxa de utilização do recurso. Houve, ainda, um aumento no tempo médio do sistema e da taxa de utilização da “Masseira” e “Posto de trabalho 2”, porém, com uma taxa de saída maior, ou seja, um *trade off* desejável na linha de produção.

Tabela 11. Comparação entre os modelos simulados.

| Variável/ Cenário                                    | Cenário atual       | Cenário proposto    |
|--|---------------------|---------------------|
| Total de saídas                                      | 1280 pães de queijo | 2080 pães de queijo |
| Tempo médio no sistema para entidade “massa”         | 52 minutos          | 64,22 minutos       |
| Tempo médio no sistema para entidade “pão de queijo” | 5,23 minutos        | 10,21 minutos       |
| Taxa média de utilização do “Posto de trabalho 2”    | 44,50 %             | 65,60 %             |
| Taxa média de utilização do “Forno.1”                | 84,78 %             | 77,97 %             |
| Taxa média de utilização do “Forno.2”                | -                   | 71,81 %             |
| Taxa média de utilização do recurso “cozinheira”     | 6,30 %              | 8,99 %              |
| Taxa média de ociosidade do recurso “cozinheira”     | 92,32 %             | 90,22 %             |

#### 5. CONCLUSÕES

Os objetivos inicialmente propostos foram atingidos. Realizou-se a modelagem e simulação da produção de pães de queijo de um ambiente real, denominado de cenário atual, posteriormente, após a análise dos resultados deste cenário, propostas de melhorias foram feitas, e outro cenário contendo tais propostas foi modelado e simulado, denominado de cenário proposto, e por fim, realizou-se a comparação entre estes dois cenários. Ambos foram desenvolvidos no software ProModel® e se diferenciam em relação a capacidade da masseira e do acréscimo de um forno.

Desta forma, os resultados obtidos apontaram que, em relação a quantidade de pães de queijo produzidos, o modelo proposto apresentou maior produtividade. A simulação do modelo proposto apresentou um aumento de produção de aproximadamente 61,63%, comparado com o cenário atual que apresentou, em média, 1280 unidades de pães de queijo produzidos ao final de 5 horas do turno de trabalho. Isto posto, verificou-se o seguinte *trade-off*: para aumentar a quantidade de saídas no modelo, é preciso aumentar a utilização média do recurso e dos locais.

Com a realização deste trabalho, pode-se aferir o quão vantajosa a simulação é, por permitir que o comportamento do sistema

fosse analisado sem realizar alterações físicas no sistema real.

Este trabalho pôde contribuir para a comunidade acadêmica uma vez que acrescentará à literatura mais um trabalho acerca de uma área que cada vez mais está ganhando espaço e reconhecimento da sua potencial contribuição para aumentar a eficiência e produtividade de uma empresa, que é a área de simulação computacional. O mesmo também contribuiu para a

comunidade empresarial, ao passo que estimula as empresas a utilizarem a técnica de simulação computacional como uma ferramenta de auxílio da tomada de decisão.

Por fim, para um trabalho futuro sugere-se simular a produção de outros produtos da empresa, os quais também possuem uma boa saída, como roscas, broas e biscoito de queijo.

## REFERÊNCIAS

- [1] ABEPRO. Um Panorama Atual da Engenharia de Produção, Porto Alegre, 2002. Disponível em: <http://www.abepro.org.br/arquivos/websites/1/PanoramaAtualEP4.pdf>. Acesso: 02/01/2017.
- [2] Banks, J. Introduction to simulation. Proceedings of the 2000 Winter Simulation Conference, 2000.
- [3] BARNES, R. M. Estudo de movimentos e de tempos: Projeto e Medida do trabalho, 8ª ed. São Paulo. Edgard blucher, 1977.
- [4] BERENDS, P.; ROMME, G. Simulation as a research tool in management studies European Management Journal, v.17, n.6, pp.576-583, 1999.
- [5] BRATLEY P.; BRENNET L.; SCHRAGE L. E. A Guide to Simulation, Second Edition. New York, Springer-Verlang, 1987.
- [6] BRYMAN, A. Research methods and organization studies. 1ª ed. London: Uniwin Hyman, 1989.
- [7] BUFFA, E. S.; SARIN, R. K. Modern Production/Operations Management. 8ª ed. John Wiley & Sons, 1987.
- [8] CHWIF, L.; MEDINA, A. C. Modelagem e Simulação de Eventos Discretos: Teoria e Aplicações. 4. ed. Rio de Janeiro: Elsevier, 2015.
- [9] CRESWELL, J. W. Research design: qualitative & quantitative approaches. 4ª ed. London: Sage, 1994.
- [10] FERREIRA, L. M.; HOCHMAN, B.; BARBOSA, M. V, J, Modelos experimentais em pesquisa. Acta Cir. Bras., São Paulo, v. 20, supl.2 ,2005.
- [11] FREITAS FILHO, P. J. Introdução à modelagem e simulação de sistemas: com aplicação em Arena. 2ª ed Florianópolis: Visual Books, 2008.
- [12] HARRELL, C.; GHOSH, B.K.; BOWDEN, R. O. Simulation using ProModel. 3. ed. New York: McGraw Hill, 2010.
- [13] HARREL, C. R.; MOTT, J. R. A.; BATEMAN, R. E.; BOWDEN, R. G. & GOGG, T. J. Simulação: otimizando os sistemas. São Paulo: Belge Simulação e IMAM. 2ª ed. 2002.
- [14] INSTITUTO TECNOLÓGICO DE PANIFICAÇÃO E CONFEITARIA - ITPC. Performance do setor de panificação e confeitaria em 2016. (2017). Disponível em <http://institutoitpc.jimdo.com/indicadores-do-setor/>. Acesso: 02/09/2017.
- [15] KELTON, W. D.; SADOWSKI, R. P.; SADOWSKI, D. A. Simulation with Arena. 2ª ed. London: McGraw-Hill, 2002.
- [16] LAW, A. M.; KELTON, W. D. Simulation modeling and analysis. 3. ed. New York: McGraw-Hill, 2000.
- [17] LEAL, F., ALMEIDA, D. A.; MONTEVECHI, J. A. B. Uma proposta de técnica de modelagem conceitual para a simulação através de elementos do IDEF. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 50., 2008. João Pessoa-PB. Anais... João Pessoa: SBPO, 2008.
- [18] SALIBY, E. Repensando a simulação: a amostragem descritiva. São Paulo: Atlas; Rio de Janeiro: Editora de UFRJ, 1989.
- [19] SARGENT, R. G. Verification and validation of simulation models. In: Winter Simulation Conference, USA, 2007.
- [20] RICHARDSON, R. J. Pesquisa social: métodos e técnicas. São Paulo: Atlas, 1999.



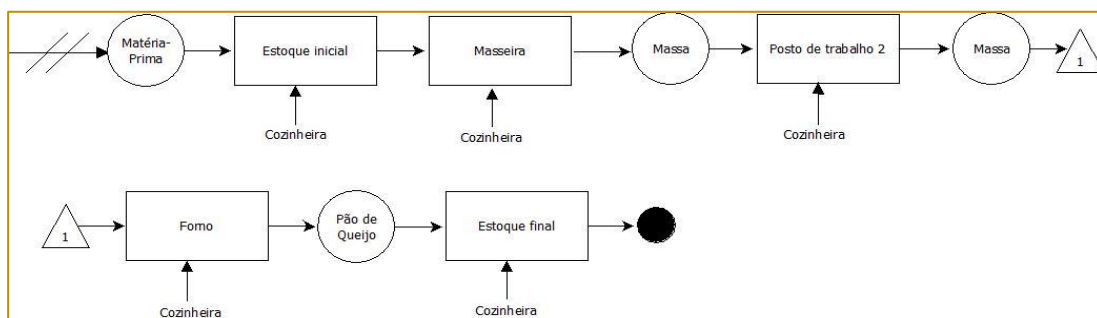
## ANEXOS

## ANEXO A – Dados coletados da pesquisa.

| Atividade/ Tempos cronometrados (em minutos) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  |
|--|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Preparo da massa                             | 8,5 | 8,4 | 8,2 | 8,3 | 8,4 | 8,6 | 8,8 | 9,1 | 8,8 | 8,7 | 8,5 | 9   | 8,4 | 8,8 | 8,7 |
| Modelagem da massa                           | 2,4 | 2,4 | 2,5 | 2,3 | 2,3 | 2,6 | 2,3 | 2,4 | 2,3 | 2,4 | 2,4 | 2,6 | 2,5 | 2,3 | 2,5 |
| Tempo de forno                               | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  | 30  |

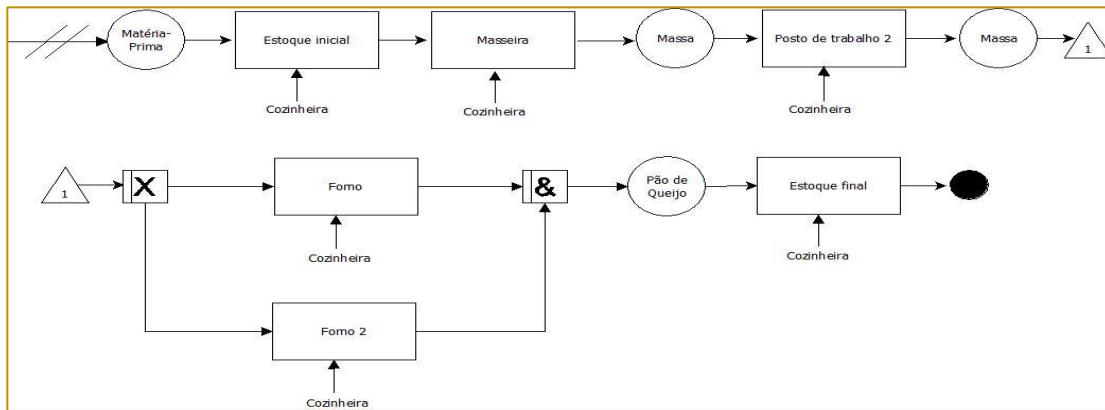
## ANEXO B – Modelos conceituais usando a técnica IDEF-SIM

Figura B.1 – Modelo conceitual do cenário atual usando a técnica IDEF-SIM.



Fonte: Dados da pesquisa

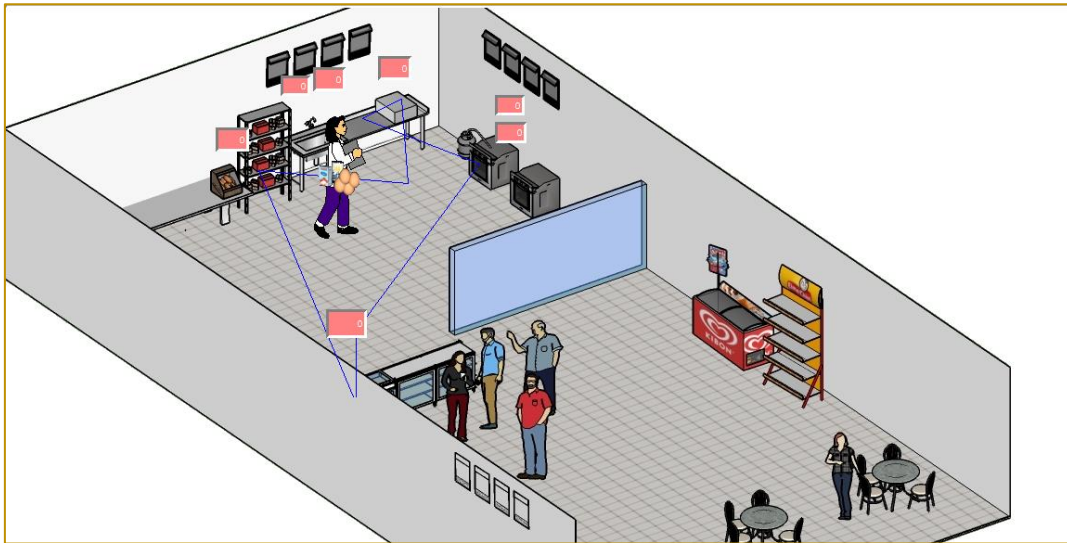
Figura B.2 – Modelo conceitual do cenário proposto usando a técnica IDEF-SIM.



Fonte: Dados da pesquisa

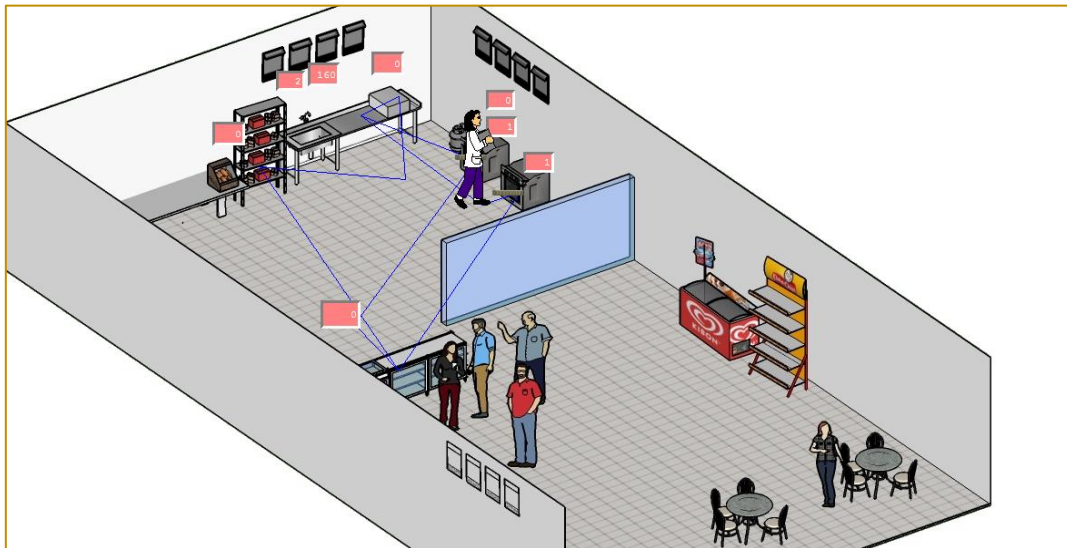
ANEXO C – Representação do sistema simulado no ProModel®.

Figura C.1 – Representação do sistema simulado do cenário atual no ProModel®.



Fonte: Elaborado pelos autores.

Figura C. – Representação do sistema simulado do cenário proposto no ProModel®.



Fonte: Elaborado pelos autores.

# Capítulo 26

## ABORDAGENS GRASP APLICADAS AO PROBLEMA QUOTA CARS

*Matheus da Silva Menezes*

*Marco César Goldberg*

*Elizabeth Ferreira Goldberg*

*Vanessa Elionara Souza Ferreira*

*Gabriela Colaço*

**Resumo:** Este artigo apresenta um estudo de aplicação de abordagens GRASP hibridizadas com *Variable Neighborhood Search* e *Path Relinking*, bem como novos operadores para a montagem da Lista de Candidatos, aplicados e adaptados ao problema do Caixeiro Alugador com Quotas. Neste problema são disponibilizados um conjunto de vértices, cada um com um bônus associado e um conjunto de veículos. O objetivo do problema é determinar um ciclo que visite alguns vértices coletando, pelo menos, um bônus pré-definido e minimizando os custos de viagem através da rota, que pode ser feita com veículos de diferentes tipos. São apresentadas quatro versões GRASP produzindo resultados em oitenta instâncias. É realizada uma comparação com resultados do modelo exato do problema, onde as abordagens GRASP encontraram o melhor resultado em 62 instâncias, mostrando o potencial de aplicação ao problema.

**Palavras chave.** Otimização Combinatória. GRASP. Quota CaRS

\*Artigo originalmente apresentado no Simpósio Brasileiro de Pesquisa Operacional (SBPO 2017)

## 1. INTRODUÇÃO

A aplicação de métodos heurísticos e técnicas de modelagem matemática como meio de solucionar problemas logísticos, de bens e/ou pessoas, tem se tornado um tema cada vez mais recorrente na literatura. Tendo em vista que tais soluções impactam positivamente na tomada de decisão do cotidiano de diversas empresas e pessoas, [DA SILVA e OCHI, 2016] explicam que muitos pesquisadores tem se empenhado no desenvolvimento de algoritmos e modelos matemáticos para solucionar problemas de gerenciamento de frotas, onde cada uma dessas apresenta características, comportamento e objetivos distintos. De acordo com [BODIN, 1990] existem diversos métodos na literatura que realizam de forma eficiente a definição e o aprimoramento de rotas de transporte, porém a maior transformação em relação a esses métodos se desenvolveu no ambiente computacional, quando muitos algoritmos baseados em modelagem matemática atingiram resultados com baixo percentual de erro em relação as soluções exatas.

O Problema do Caixeiro Viajante (PCV) é um problema de otimização combinatória clássico e tema de diversos estudos e consiste em definir o menor custo associado a distância percorrida entre  $n$  vértices, iniciando e finalizando a roteirização no mesmo ponto, com cada vértice sendo visitado uma única vez. Conforme apresentado em [GOLDBARG; LUNA, 2005], o PCV tem aplicações em problemas de coleta e distribuição de produtos, sistemas de produção, roteamentos, seqüenciamento de tarefas, dentre outros. No trabalho de [GOLDBARG, ASCONAVIETA E GOLDBARG, 2012] foi proposta uma generalização do PCV, considerando como elemento central o cliente que pretende percorrer uma rota utilizando-se de veículos alugados, devendo-se agora considerar não apenas os custos associados às distancias percorridas, como também o custo associado ao aluguel de acordo com o modelo de carro utilizado, bem como suas taxas associadas, denominado Problema do Caixeiro Alugador ou *Traveling CarRenter Problem* (CaRS).

No CaRS, existem diversas possibilidades de escolhas de veículo para o cliente, que deve buscar em cada ponto obter a opção de carro com melhor combinação possível no que se refere aos custos associados À distancia

percorrida, valor de aluguel e taxa de retorno, onde este último é um custo associado a entrega do carro em um ponto diferente do de origem de aluguel de tal veículo. Admitindo-se que o CaRS apresenta uma maior quantidade de variantes a serem consideradas, e o PCV é reconhecido como um subcaso do mesmo e é NP-difícil [GUTIN; PUNNEN, 2002], o CaRS também é um problema NP-difícil.

De acordo com [GOLDBARG et al. 2011] o CaRS possui grande aplicação no segmento de transporte turístico, segmento este que vem apresentando um crescente volume de estudos. Além disso, diversos estudos, tais como [VANSTEENWEGEN, SOUFFRIAU e OUDHEUSDEN, 2011], estão direcionados para a otimização de definição de rotas turísticas, buscando obter o máximo de satisfação de um indivíduo que apresenta diferentes restrições em visitar variados pontos de turísticos de uma cidade. Este artigo trata de uma variante do problema CaRS, denominda Caixeiro Alugador com Quota ou *Quota CaRS* (QCaRS) que foi introduzido em [MENEZES et al, 2014] e reformulado em [GOLDBARG et al, 2016]. Este problema é uma generalização do Problema do Caixeiro Viajante com Quota ou *Quota TSP* (QTSP). O Problema QCaRS é uma variante do CaRS que, assim como no QTSP, uma quota é associada a cada cidade. Esta quota define um nível de satisfação em visitar a cidade associada. Esse parâmetro é relevante, pois em muitos casos o turista não pode visitar todas as atrações existentes durante uma viagem, devendo escolher as que possuam maior interesse.

No problema QCaRS uma satisfação presumida de visita é atribuída a cada cidade e uma satisfação cumulativa pré-definida mínima deve ser atendida. Além disso, um vértice é escolhido para ser a cidade onde o passeio começa e termina, denominada cidade base. O objetivo é selecionar um subconjunto de cidades (vértices) a serem visitados, visando a minimização do custo total da viagem, incluindo o aluguel dos veículos e que a satisfação mínima de visitar as cidades, representada pelo parâmetro  $\omega$  seja atendida. Nesses casos, é interessante tentar maximizar a satisfação visitando os pontos mais atraentes.

Como o problema QCaRS é relativamente novo na literatura, poucas metaheurísticas foram aplicadas ao mesmo. Em [MENEZES et AL, 2014] foi apresentado um algoritmo

memético e em [GOLDBARG et al, 2016] um algoritmo evolucionário híbrido para a resolução do problema. O presente artigo propõe a implementação de estratégias GRASP adaptadas ao problema, e hibridizadas com *Variable Neighborhood Search* (VNS) e *Path Relinking*, ainda não relatadas na literatura.

O resto do artigo está organizado de acordo com a seguinte estrutura: Na seção 2 é definido o problema QCaRS. Na seção 3, as estratégias GRASP e suas hibridizações são relatadas em detalhes. Já na seção 4, temos os resultados dos experimentos computacionais enquanto na seção 5 temos as considerações finais sobre o estudo realizado.

## 2. O PROBLEMA QUOTA-CARS

Seja  $G = (V, A)$  um grafo completo, onde  $V$  é um conjunto com  $n$  nós (cidades) e  $A$  é um conjunto de arcos (estradas entre as cidades). Um bônus  $SV_i$ ,  $i = 1, \dots, n$ , é atribuído a cada cidade  $i \in V$ . Neste problema, um conjunto  $C$  de veículos está disponível para aluguel. Custos operacionais específicos estão associados a cada carro, incluindo o consumo de combustível, pagamento de pedágio e custos de aluguel. Uma vez que as taxas de pedágio normalmente dependem do tipo de veículo e do comprimento do caminho percorrido, é possível assumir, sem perda de generalidade, que o custo operacional de cada carro para atravessar aresta  $(i, j) \in A$  é uma função daquele carro. Além disso, o QCaRS considera que o bônus mínimo a ser recolhido é dado pelo parâmetro  $\omega$ , e nesse estudo tem valor igual  $0.8ST$ , onde,  $ST = \sum SV_i$ ,  $i = 1, \dots, n$ .

O objetivo do QCaRS é completar um caminho Hamiltoniano em  $G$  pagando o mínimo possível. Como o QCaRS é uma variante do Quota TSP quando apenas um carro é utilizado, e o Quota TSP é NP-Difícil, então o QCaRS também é.

Neste estudo, investigamos a mesma variante do problema QCaRS relatada em outros estudos: o grafo que modela o problema possui todas as ligações (**completo**); qualquer carro pode ser alugado em qualquer cidade (**total**); o carro alugado pode ser devolvido em qualquer cidade (**irrestrito**); cada carro só pode ser locado uma vez (**sem repetição**); os custos de retorno independem da topologia das estradas ou restrições de rede (**livre**); os

custos de utilizar um carro  $c$ , para ir de uma cidade  $i$  para uma cidade  $j$  são simétricos aos de  $j$  para  $i$  utilizando  $c$  (**simétrico**)

Uma definição formal detalhada e um modelo matemático para o problema foram apresentados e implementados no *solver* GLPK em [MENEZES et al, 2014] e atualizado com novas linearizações em [GOLDBARG et al, 2016].

## 3. ESTRATÉGIAS GRASP E VNS APLICADAS AO QUOTA CARS

GRASP é a abreviação da nomenclatura Greedy Randomized Adaptive Search Procedure e proposto por [FEO; RESENDE, 1989] e [FEO; RESENDE, 1995]. Seu diferencial está na geração da solução inicial, cuja abordagem está referenciada nas três primeiras iniciais de sua sigla: gulosa (*Greedy*), aleatória (*Randomized*) e adaptativa (*Adaptive*). Essa estratégia tem sido amplamente utilizada em problemas relacionados ao roteamento de veículos, a exemplo de [LABADIE; MELECHOVSKÝ; CALVO, 2011], [SOUFFRIAU et al., 2008], [CHAOVALITWONGSE; KIM; PARDALOS, 2003], [CHAVES et al., 2007], [MARINAKIS; MIGDALAS; PARDALOS, 2005].

Essa metaheurística busca manter um equilíbrio entre ser guloso e aleatório, e envolve duas fases distintas: a Fase Construtiva, onde busca-se construir boas soluções, utilizando-se para isso uma lista de candidatos, listando todos os vizinhos, a partir da qual é gerada a Lista Restrita de Candidatos - LRC, onde sua composição é formada pelos  $\alpha$  % melhores vizinhos disponíveis, onde  $0 \leq \alpha \leq 1$ . Valores de  $\alpha$  próximos a zero conduzem a soluções muito próximas aquela obtida pela escolha gulosa, enquanto valores de  $\alpha$  próximos a 1 conduzem a soluções praticamente aleatórias. A heurística é adaptativa porque os benefícios associados com a escolha de cada elemento são atualizados em cada iteração da fase de construção, de forma a refletir as mudanças oriundas da seleção do elemento anterior. O componente probabilístico do procedimento reside no fato de que cada elemento é selecionado de forma aleatória, a partir de um subconjunto restrito formado pelos melhores elementos que compõem a LRC.

Na segunda fase aplica-se a Busca Local. Procura-se com ela melhorar as soluções encontradas, através de um método de

exploração da vizinhança. Um algoritmo de busca local define uma vizinhança para cada solução, fazendo uma transformação da solução atual através de uma técnica ou heurística, gerando um conjunto de soluções com características próximas.

O principal desafio entre os métodos heurísticos que usam a busca local é definir uma estratégia eficiente para cobrir o espaço de busca, explorando principalmente regiões promissoras [VANSTEENWEGEN et al., 2011].

Muitas vezes temos o GRASP hibridizado com outra metaheurística. A heurística *Variable Neighborhood Search* (VNS) foi proposta por [MLADENOVIC, 1995] e consiste na ideia de mudar sistematicamente a vizinhança de busca. O método VNS funciona de forma pouco convencional se comparado a outras metaheurísticas [CHAVES et al., 2007]. Ele explora vizinhanças mais distantes da solução corrente e concentra a busca em torno de uma nova solução. A solução é atualizada apenas caso um movimento de melhora seja realizado. O método pode incluir também rotinas de busca local que utilizam diferentes estruturas de vizinhança.

Outra estratégia de uso comum em conjunto com o GRASP é o *Path-Relinking*, proposto de forma pioneira por [GLOVER, 1996] e é uma estratégia de intensificação de soluções que possui o objetivo de gerar novas soluções a partir da utilização de trechos de soluções de elite que são inseridos nas soluções do sistema e assim encontrar soluções de melhor qualidade, que possuam as características relevantes das soluções de elite. Para gerar os caminhos, os movimentos são selecionados para apresentar os atributos que estão presentes na solução de elite, passando-os para a solução atual. O *Path Relinking* pode ser visto como uma estratégia que pretende incorporar atributos de soluções de elevada qualidade, ao favorecer estes atributos nas novas soluções através de movimentos selecionados. Em [RESENDE; RIBEIRO, 2003], é feita uma revisão das várias alternativas consideradas para o *Path-Relinking*.

De posse dos comentários acima, foram propostos no presente trabalho quatro algoritmos com abordagem GRASP, sendo o primeiro em sua formulação tradicional hibridizado com VNS, o segundo hibridizado com VNS e *Path-Relinking*. Os outros dois algoritmos utilizarão um novo operador de intensificação que será descrito em detalhes

posteriormente. Os algoritmos de abordagem GRASP utilizados no presente trabalho são descritos a seguir.

### 3.1. GRASP + VNS

Nesta seção, é apresentado o algoritmo GRASP hibridizado com VNS, proposto para a resolução do problema pCaRS. A fase inicial do algoritmo corresponde a fase construtiva do GRASP. A segunda fase engloba a aplicação de técnicas de busca local com abordagem VNS nas soluções encontradas.

O pseudo-código do algoritmo é apresentado na Figura 1, onde suas partes constitutivas são explicadas nos itens subsequentes. Os parâmetros de entrada para o algoritmo são: *nomeInstancia*, que identifica o problema que será executado, *imax*, que representa o número de iterações que o algoritmo vai executar, e  $\alpha$  que é o tamanho da Lista Restrita de Candidatos. O conjunto de soluções encontrados é representado por  $X = \{x_0, x_1, \dots, x_n\}$  e a melhor solução encontrada por  $x^*$ . A função que avalia o custo de cada solução é representada por *f* e *numSol* representa o número de soluções que serão geradas na fase construtiva do algoritmo.

A fase construtiva está implementada no método *geraSolConstrutiva*( $\alpha$ ); que recebe o tamanho da lista restrita de candidatos como parâmetro de entrada. A cidade inicial escolhida no algoritmo é a cidade base (nó-1). Um carro  $c_1$  é escolhido aleatoriamente para a primeira cidade. Em seguida, a cidade *i* onde o carro  $c_1$  será entregue é também escolhida de forma aleatória. Um caminho é criado entre o nó-1 e a cidade *i* através da metodologia construtiva do GRASP, considerando os custos de viagem com o carro  $c_1$  para montar a LRC. A cidade *i* é a cidade inicial da próxima parte da rota. Em seguida, um novo carro  $c_2$  e uma nova cidade *j* são escolhidos aleatoriamente. Obviamente, a cidade *j* não pode ser uma cidade visitada anteriormente. Um caminho é construído entre as cidades *i* e *j* com as cidades não visitadas, utilizando a mesma abordagem. Caso não existam mais carros disponíveis, utiliza-se o último carro sorteado. O procedimento continua sorteando outra cidade aleatoriamente e construindo o caminho até que o nível mínimo de satisfação seja alcançado, ou seja, 80% da satisfação total disponível. Então o ciclo é fechado,

considerando a ligação com a cidade de partida.

Figura 1: Algoritmo 1: GRASP + VNS proposto para o QCaRS

**Entrada:** nomeInstancia,  $i_{\max}$ ,  $\alpha$

**Saída:**  $x^*$

```

 $f(x^*) \leftarrow \infty$ 
para  $i=1 \dots i_{\max}$  faça
|    $x \leftarrow \text{geraSolConstrutiva}(\alpha)$ ;
|    $x \leftarrow \text{buscaLocal}(x)$ ;
|   se  $f(x) < f(x^*)$  então
|   |    $f(x^*) \leftarrow f(x)$ 
|   |    $x^* \leftarrow x$ 
|   fim
fim
retorne  $x^*$ 

```

O VNS corresponde à fase de busca local, que no caso do problema QCaRS, pode atuar em três áreas distintas: a melhor rota, a ordem de uso dos carros e o nível de satisfação. A fase de busca local do algoritmo proposto visa lidar com estas três áreas, afim de

melhorar os candidatos a solução. A busca local é implantada no procedimento *buscaLocal()* apresentado na Figura 2 e é similar ao proposto em [MENEZES et al, 2014], sendo composto por seis buscas locais aplicadas em seqüência.

Figura2: Procedimento *buscaMultiOperadores*

1. *buscaLocal*( $x$ )
2.  $x \leftarrow \text{removeCidade}(x)$
3.  $x \leftarrow \text{inverteSol}(x)$
4.  $x \leftarrow \text{insereCidade}(x)$
5.  $x \leftarrow \text{substituiCidade}(x)$
6.  $x \leftarrow \text{substituiCarro}(x)$
7.  $x \leftarrow 2\text{Opt}(x)$
8. **fim**

A cada aplicação de procedimentos constantes na busca local, a solução  $x$  é atualizada com o melhor valor encontrado no procedimento. Dessa forma, se a solução de entrada é melhorada durante a aplicação de um procedimento da fase de busca local, então a nova solução substitui a antiga. Ao final do procedimento de busca local é retornada a melhor solução da vizinhança da solução de entrada  $x$ . Os procedimentos de busca possuem o seguinte funcionamento básico:

***removeCidade.*** Este método consiste em remover da solução as cidades com os menores índices de satisfação. As cidades vão sendo removidas enquanto a restrição referente a satisfação mínima é atendida.

***InverteSol.*** Este método inverte a ordem de visita das cidades na solução. Os mesmos carros estão associados com as mesmas cidades, mas os pontos de aluguel e entrega são trocados.

**substituiCarro.** Este procedimento concentra-se em veículos que não estão ainda na solução a examinando a inserção de um novo carro, se possível, substituindo outro na solução.

**substituiCidade.** Este procedimento examina a substituição de cidades na solução por cidades que não estão presentes na solução. Todas as cidades que estão fora da solução de entrada são consideradas para a inserção.

**insereCidade.** Este procedimento testa a inserção de uma nova cidade entre cada par de cidades da solução de entrada. Ele também foca em cidades que ainda não estão na solução de entrada. O carro atribuído a nova cidade é o mesmo da cidade imediatamente anterior. .

**2-opt.** é um algoritmo de busca local simples proposto por Croes (1958) para o PCV. Um movimento 2-opt consiste em eliminar duas arestas e religar os dois caminhos, resultando em uma maneira diferente de obter uma nova rota. Entre todos os pares de arestas cuja troca 2-opt atua, a que resultar no menor custo de rota é escolhido. Este procedimento é então iterado até que o par de arestas mais econômico seja encontrado. A rota resultante é chamada de 2-ótima. Neste caso, a seqüência de carros associados permanece a mesma

Durante a aplicação dos procedimentos de busca local e Path Relinking, podemos ter a geração de soluções inviáveis. Tais soluções são reparadas através do mesmo procedimento adotado em [MENEZES et al, 2014].

### 3.2.. GRASP + VNS + PATH RELINKING

O procedimento *Path Relinking* é bastante utilizado em conjunto com a metaheurística GRASP, como estratégia de intensificação de soluções [RESENDE; RIBEIRO, 2003]. Duas estratégias básicas de utilização *do Path-Relinking* são: a aplicação a todos os pares de soluções elite, ou periodicamente, durante as iterações do GRASP, ou após todas as iterações GRASP em uma etapa de pós-otimização.

O *Path Relinking* proposto considera que uma quantidade pré-definida das melhores soluções geradas pelo algoritmo GRASP formarão o conjunto de elite. Este conjunto é atualizado e ordenado a cada iteração até atingir o número máximo de soluções pré-determinado. Se uma nova solução for melhor do que a última solução do conjunto (e conseqüentemente a pior solução deste conjunto), então a nova boa solução é adicionada ao conjunto e a última é removida, reordenando o conjunto.

Figura 3: Algoritmo 2: GRASP + VNS + *Path Relinking* proposto para o QCaRS

**Entrada:**  $nomInstancia, imax, \alpha$

**Saída:**  $x^*$

$f(x^*) \leftarrow \infty$

**para**  $i=1 \dots i_{max}$  **faça**

$x \leftarrow geraSolConstrutiva(\alpha);$

$x \leftarrow buscaLocal(x);$

**se**  $f(x) < f(x^*)$  **então**

$f(x^*) \leftarrow f(x)$

$x^* \leftarrow x$

**fim**

$Elite \leftarrow montaElite(x);$

**fim**

$x \leftarrow PathRelinking(Elite);$

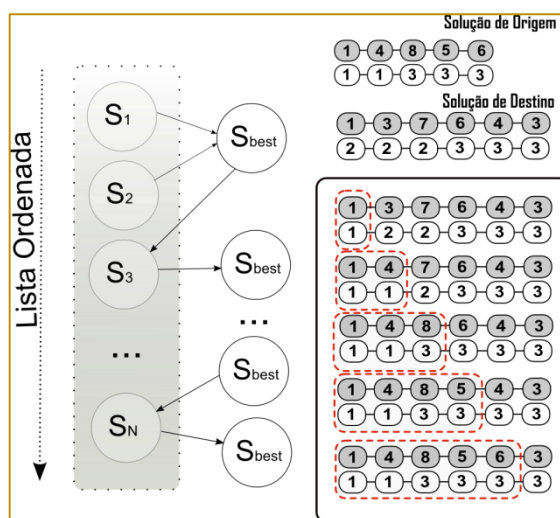
**retorne**  $x^*$



De posse dessas soluções, o procedimento funciona utilizando a melhor solução do conjunto como solução base. A solução de destino é escolhida como sendo a segunda melhor solução do conjunto. Em seguida busca-se inserir cada cidade, com seu respectivo carro na solução de destino, sobrepondo o par cidade-carro anteriormente existente. O algoritmo considera a inserção acumulativa das cidades constantes na solução base na solução destino, na mesma ordem em que estão presentes na solução

base. Ao final desse processo, a melhor solução obtida permanece e será a nova solução base. A terceira melhor solução do conjunto é então escolhida como solução de destino, repetindo todo o processo. O procedimento continua de forma semelhante até que não existam mais soluções destino no conjunto de elite. A Figura 4 a seguir ilustra o funcionamento da escolha das soluções base e destino e do operador de *Path Relinking* proposto.

Figura 4: Exemplo de aplicação do Procedimento *Path Relinking*

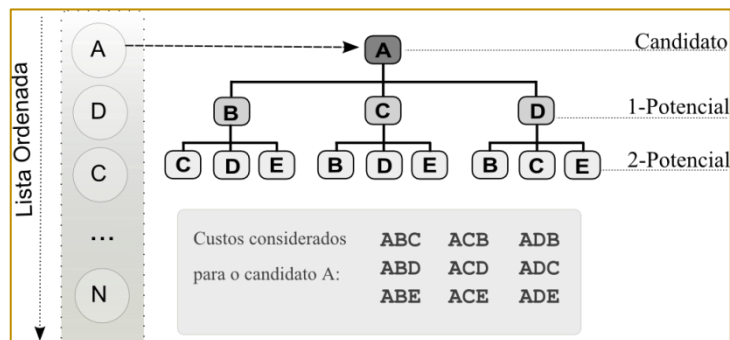


### 3.3. GRASP 2-POTENCIAL + VNS + PATH RELINKING

Este terceiro experimento com o GRASP considerou a utilização de uma nova abordagem na construção da lista de candidatos. Para cada candidato, verificou-se a adequação para inserção considerando dois passos adiante, considerando as três

idades com menor custo de inserção em cada passo. A intenção dessa abordagem é de tentar evitar que seja efetuada uma inserção que é boa considerando a LRC atual, mas que não será uma boa opção para o restante da solução a longo prazo.

Figura 5: Exemplo de aplicação do Procedimento *2Potencial*



Para cada cidade na Lista de Candidatos, são consideradas as próximas três melhores cidades (com menor custo), e a partir de cada uma dessas cidades, as três próximas de menor custo. A montagem da lista é feita considerando o somatório cumulativo de cada caminho com todos os carros disponíveis. Cada cidade da lista de candidatos terá  $(nCar * 9)$  opções geradas para a lista. A LRC é montada a partir da lista de candidatos, considerando os  $\alpha$  % melhores. A Figura 5 ilustra o funcionamento dessa abordagem a partir de uma entrada de cidade A da lista de candidatos da abordagem tradicional.

É importante ressaltar que as soluções podem ter tamanhos (quantidade de cidades) diferentes, o que pode tornar o procedimento apenas parcial, conforme ilustrado na Figura 5. Após cada inserção, pode ser necessário aplicar o procedimento de restauração de soluções, já descrito anteriormente. Foram consideradas para implementação duas variações dessa abordagem, com diferenças na forma como o sorteio na LRC é efetuado, sendo:

**Versão 1:** O sorteio na LRC foi efetuado considerando igual probabilidade a todas as candidatas.

**Versão 2:** O sorteio na LRC foi do tipo "roleta", onde a probabilidade foi proporcional ao valor de inserção de cada candidato.

#### 4. EXPERIMENTOS COMPUTACIONAIS

As metaheurísticas propostas e seus respectivos experimentos computacionais são apresentados com detalhes no presente capítulo. Com a finalidade de validar os algoritmos propostos, os mesmos foram implementados em um PC Intel Core i5, 16G de RAM, rodando Ubuntu 12.04 64bits, utilizando a linguagem C++ e executados trinta vezes para cada instância. Os testes preliminares para a definição de parâmetros do algoritmo, consideraram seis instâncias da biblioteca QCaRSLib, disponível em [http://www.dimap.ufrn.br/lae/en/projects/CaRS\\_php](http://www.dimap.ufrn.br/lae/en/projects/CaRS_php). com um número de cidades na faixa de 9 a 25 e 2 a 4 veículos, onde foram executadas 20 vezes cada uma. Dessa forma, os parâmetros adotados no GRASP foram uma lista LCR de tamanho  $\alpha=0.25$  e como critério de parada, foi estabelecido um limite de 5000 iterações.

Os algoritmos com abordagem GRASP propostos obtiveram uma boa adaptação aos problemas analisados. Os algoritmos variaram na fase de pós otimização, onde tivemos a proposta com e sem *Path-Relinking*. Além disso, foi proposta a abordagem 2-Potencial com duas formas de escolha do candidato na LCR: através de sorteio e através de roleta. Os resultados são exibidos nas tabelas 1 e 2. O menor tempo de processamento em todas as instâncias provém da abordagem GRASP sem estratégias de pós-otimização, representado pela sigla GRASP1. A abordagem GRASPPR também obteve um tempo de processamento ligeiramente superior ao da abordagem GRASP1, indicando que o operador de *Path Relinking* é leve e demanda pouco esforço computacional. A abordagem 2Potencial obteve um tempo de processamento pouco maior que as outras duas abordagens.

O menor valor encontrado para o problema, incluindo os resultados do modelo exato obtidos em [GOLDBARG et al, 2016] são tidos como valor de referência para o problema. Como o problema considerado é de minimização, quanto maior o resultado desta operação, pior é a solução encontrada pelo algoritmo. Em relação a essa variável, verifica-se que a versão GRASPPR foi mais eficiente, encontrando, de forma geral, os menores valores da abordagem considerada. O GRASP1 também obteve um bom desempenho, com soluções próximas ao mínimo encontrado. A abordagem 2-Potencial obteve desempenho pior do que o algoritmo com a abordagem tradicional. Os menores valores médios foram encontrados pela abordagem GRASPPR, que utiliza o operador de *Path-Relinking* como pós-otimização. As abordagens de GRASPPR e GRASP1 obtiveram soluções médias muito próximas, e com comportamento semelhante. Observe que estas duas abordagens se destacaram em relação a utilização da metaheurística GRASP com abordagem 2Potencial. É importante ressaltar que quanto maior o número de cidades na instância, maior é a diferença entre a solução mínima encontrada e a solução média dos algoritmos, indicando uma perda de eficiência na busca dos valores mínimos em instâncias com maior número de cidades.

## 5. CONSIDERAÇÕES FINAIS

O problema do caixeiro alugador com coleta de prêmios aqui analisado possui uma maior dificuldade na fase de busca local, pois possui três regiões onde as buscas podem atuar: na rota da solução, nos pontos de troca dos veículos, e no recolhimento de bônus. Cada busca aplicada pode ser efetiva para um dos aspectos acima, porém sem garantias de melhoria dos demais, sendo um dos aspectos mais sensíveis do problema, pois eles são conflitantes entre si.

As abordagens GRASP demandaram pouco tempo de processamento entre quando comparada às abordagens utilizadas em [MENEZES et al, 2014] e [GOLDBARG et al, 2016]. O Algoritmo Grasp2PRol não obteve resultados qualitativos satisfatórios, indicando que a abordagem 2-Potencial com escolha através de roleta não foi bem adaptada a esta variante do problema. Para a abordagem GRASP-VNS, analisamos quatro variações onde apenas a primeira versão não contou com uma fase de pós-otimização. Nas demais, analisamos o operador de *Path-*

*Relinking*. A segunda versão do algoritmo mostrou que o operador foi eficiente na melhoria da solução ótima encontrada pelo algoritmo, a um custo de tempo de processamento relativamente baixo. Foi proposta uma nova abordagem de montagem da Lista de Candidatos denominada 2Potencial, que não teve resultados promissores. Nessa proposta, o sorteio da LCR também foi feito através de roleta, cujos resultados foram inferiores a todas as outras abordagens. O ponto positivo dessa metaheurística foi relativa ao baixo tempo de processamento, quando comparado à abordagens analisadas em outros estudos para este problema. De uma forma geral, os algoritmos analisados tiveram uma boa adaptação para as instâncias menores, tanto Euclidianas, quanto não Euclidianas. Ao aumentar o tamanho da instância, os algoritmos vão perdendo eficiência na busca pelo valor ótimo, e tendem a demandar maior tempo de processamento. A perda de eficiência é mais notória nas instâncias não Euclidianas.

## REFERÊNCIAS

- [1] BODIN, L.D. Twenty years of routing and scheduling. *Operations Research*, v. 38, n. 4, pp. 571-579, 1990.
- [2] CHAOVALITWONGSE, W.; KIM, D.; PARDALOS, P. Grasp with a new local search scheme for vehicle routing problems with time windows. *Journal of Combinatorial Optimization*, Kluwer Academic Publishers, v. 7, n. 2, p. 179\_207, 2003. ISSN 1382-6905.
- [3] CHAVES, A. A. et al. Metaheurísticas híbridas para resolução do problema do caixeiro viajante com coleta de prêmios. *Produção, sciELO*, v. 17, p. 263 - 272, 08 2007. ISSN 0103-6513
- [4] DA SILVA, A.R.V.; OCHI, L.S. An efficient hybrid algorithm for the Traveling Car Renter Problem. *Expert Systems With Applications*, v. 64, pp. 132-140, 2016.
- [5] FEO, T. A.; RESENDE, M. G. A probabilistic heuristic for a computationally difficult set covering problem. *Operations Research Letters*, v. 8, n. 2, p. 67 - 71, 1989. ISSN 0167-6377.
- [6] FEO, T. A.; RESENDE, M. G. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, v. 6, p. 109-133, 1995.
- [7] GLOVER, F. Tabu search and adaptive memory programming: Advances, applications and challenges. In: *Interfaces in Computer Science and Operations Research*. [S.l.]: Kluwer, 1996. p. 1-75.
- [8] GOLDBARG, M. C.; ASCONAVIETA, P. H.; GOLDBARG, E. F. G. Memetic algorithm for the traveling car renter problem: an experimental investigation. *Memetic Computing*, v. 4, n. 2, p. 89-108, 2012.
- [9] GOLDBARG, M. C.; ASCONAVIETA, P. H.; GOLDBARG, E. F. G. Algoritmos evolucionários na solução do problema do caixeiro alugador. In: LOPES, H. S.; TAKAHASHI, R. H. C. (Org.). *Computação Evolucionária em Problemas de Engenharia*. Curitiba: Omnipax, 2011, p. 301-330.
- [10] GOLDBARG, M. C.; GOLDBARG, E. F. G., MENEZES, M.S., PACCA LUNA, Henrique P. Quota traveling car renter problem: Model and evolutionary algorithm. In: *Information Sciences. Volumes 367-368*, 1 Novembro 2016, Pags. 232-245, ISSN 0020-0255.
- [11] GUTIN, G.; PUNNEN, A. *The traveling salesman problem and its variations*. Kluwer Academic Publishers, 2002.
- [12] LABADIE, N.; MELECHOVSKÝ, J.; CALVO, R. W. Hybridized evolutionary local search algorithm for the team orienteering problem with time windows. *Journal of Heuristics*, Springer US, v. 17, n. 6, p. 729-753, 2011. ISSN 1381-1231.
- [13] MARINAKIS, Y.; MIGDALAS, A.; PARDALOS, P. Expanding neighborhood grasp for

---

the traveling salesman problem. Computational Optimization and Applications, Kluwer Academic Publishers, v. 32, n. 3, p. 231-257, 2005. ISSN 0926-6003.

[14] MENEZES, Matheus da Silva, GOLDBARG, Marco César, GOLDBARG, Elizabeth F. G. A Memetic Algorithm for the Prize-collecting Traveling Car Renter Problem. IEE Congress on Evolutionary Computation (IEEE CEC 2014) Beijing, China.

[15] MLADENOVIC, N. A variable neighborhood algorithm: a new metaheuristics for combinatorial optimization. In: Abstracts of Papers Presented at Optimization Days. Montreal. 1995

[16] SOUFFRIAU, W. et al. A greedy randomised adaptive search procedure for the team orienteering problem. In: EU/MEeting 2008 on metaheuristics for logistics and vehicle routing, Troyes, France, 23-24 October 2008.

[17] VANSTEENWEGEN, P.; SOUFFRIAU, W.; OUDHEUSDEN, D. V. The orienteering problem: A survey. European Journal of Operational Research, v. 209, n. 1, p. 1-10, 2011.

ANEXO

Tabela 1: Resultados para as instâncias Euclidianas

| Problemas<br>Instância | nCid | nCar | Exato  |       |       | GRASP1 |       |       | GraspPR |        |       | Grasp2PBoleto |       |        | Grasp2PProleta |       |       |        |       |
|------------------------|------|------|--------|-------|-------|--------|-------|-------|---------|--------|-------|---------------|-------|--------|----------------|-------|-------|--------|-------|
|                        |      |      | GLPK   | Min   | Media | #vezes | T (s) | Min   | Media   | #vezes | T (s) | Min           | Media | #vezes | T (s)          | Min   | Media | #vezes | T (s) |
| Mauritania10e          | 10   | 2    | 422    | 422   | 422   | 30     | 1     | 422   | 422     | 30     | 1     | 422           | 422   | 30     | 1              | 422   | 422   | 30     | 1     |
| Colombia11e            | 11   | 2    | 326    | 326   | 326   | 30     | 1     | 326   | 326     | 30     | 1     | 326           | 326   | 30     | 1              | 326   | 326   | 30     | 1     |
| Angola12e              | 12   | 2    | 490    | 490   | 490   | 30     | 1     | 490   | 490     | 30     | 1     | 490           | 490   | 30     | 1              | 490   | 490   | 30     | 2     |
| Peru13e                | 13   | 2    | 556    | 556   | 556   | 25     | 1     | 556   | 556     | 25     | 1     | 556           | 559   | 7      | 2              | 556   | 560   | 16     | 2     |
| Libia14e               | 14   | 2    | 521    | 521   | 529   | 10     | 2     | 521   | 521     | 16     | 2     | 533           | 533   | 26     | 2              | 533   | 544   | 5      | 2     |
| BrazilRJ14e            | 14   | 2    | 230    | 230   | 233   | 1      | 2     | 230   | 232     | 15     | 2     | 230           | 230   | 28     | 2              | 230   | 230   | 26     | 2     |
| Congo15e               | 15   | 2    | 513    | 513   | 513   | 28     | 2     | 513   | 514     | 24     | 2     | 513           | 513   | 30     | 2              | 513   | 520   | 4      | 2     |
| Argentina16e           | 16   | 2    | 719    | 719   | 729   | 6      | 2     | 719   | 727     | 9      | 2     | 719           | 727   | 3      | 2              | 719   | 729   | 2      | 3     |
| EUA17e                 | 17   | 2    | 602    | 610   | 639   | 4      | 3     | 610   | 642     | 1      | 3     | 602           | 623   | 3      | 3              | 602   | 620   | 2      | 3     |
| Bolivia10e             | 10   | 3    | 384    | 384   | 384   | 30     | 1     | 384   | 384     | 30     | 1     | 384           | 384   | 30     | 1              | 384   | 386   | 22     | 2     |
| AfricaSul11e           | 11   | 3    | 402    | 402   | 402   | 30     | 1     | 402   | 402     | 30     | 1     | 402           | 402   | 30     | 2              | 402   | 402   | 30     | 2     |
| Niger12e               | 12   | 3    | 564    | 564   | 564   | 30     | 1     | 564   | 564     | 30     | 2     | 564           | 564   | 30     | 2              | 564   | 567   | 24     | 2     |
| Mongolia13e            | 13   | 3    | 543    | 544   | 545   | 1      | 2     | 545   | 545     | 15     | 2     | 543           | 545   | 1      | 2              | 545   | 545   | 28     | 2     |
| Indonesia14e           | 14   | 3    | 504    | 504   | 517   | 15     | 2     | 504   | 515     | 10     | 2     | 504           | 519   | 14     | 3              | 504   | 532   | 8      | 2     |
| Argelia15e             | 15   | 3    | 487    | 487   | 492   | 13     | 2     | 487   | 493     | 12     | 2     | 487           | 489   | 21     | 3              | 487   | 501   | 1      | 3     |
| India16e               | 16   | 3    | 705    | 705   | 724   | 5      | 2     | 705   | 725     | 5      | 2     | 705           | 705   | 28     | 3              | 705   | 721   | 6      | 3     |
| China17e               | 17   | 3    | 735*   | 741   | 762   | 3      | 2     | 736   | 762     | 1      | 3     | 736           | 744   | 6      | 3              | 736   | 766   | 1      | 3     |
| Etiopia10e             | 10   | 4    | 283    | 283   | 283   | 30     | 1     | 283   | 283     | 30     | 1     | 283           | 383   | 30     | 1              | 283   | 283   | 30     | 3     |
| Mali11e                | 11   | 4    | 428    | 428   | 428   | 30     | 1     | 428   | 428     | 30     | 2     | 428           | 428   | 30     | 2              | 428   | 428   | 30     | 2     |
| Chade12e               | 12   | 4    | 655    | 655   | 676   | 1      | 2     | 655   | 673     | 4      | 2     | 655           | 673   | 3      | 2              | 675   | 682   | 1      | 2     |
| Ira13e                 | 13   | 4    | 532    | 532   | 541   | 12     | 2     | 532   | 540     | 17     | 2     | 532           | 534   | 20     | 3              | 532   | 583   | 1      | 3     |
| Mexico14e              | 14   | 4    | 492    | 492   | 492   | 30     | 2     | 492   | 492     | 29     | 2     | 492           | 492   | 30     | 3              | 492   | 492   | 30     | 3     |
| Sudao15e               | 15   | 4    | 422    | 422   | 422   | 30     | 2     | 422   | 422     | 30     | 2     | 422           | 422   | 30     | 3              | 422   | 422   | 30     | 3     |
| Australia16e           | 16   | 4    | 682    | 711   | 777   | 1      | 3     | 711   | 775     | 1      | 3     | 730           | 762   | 1      | 4              | 727   | 788   | 1      | 4     |
| Canada17e              | 17   | 4    | 783    | 810   | 878   | 1      | 3     | 817   | 879     | 1      | 3     | 783           | 836   | 1      | 4              | 837   | 924   | 1      | 5     |
| Arabia14e              | 14   | 5    | 482    | 482   | 482   | 28     | 2     | 482   | 482     | 30     | 2     | 482           | 482   | 30     | 3              | 482   | 490   | 15     | 4     |
| Cazaquistao15e         | 15   | 5    | 574    | 589   | 611   | 2      | 2     | 585   | 607     | 1      | 3     | 585           | 600   | 2      | 4              | 589   | 640   | 1      | 3     |
| Brazil16e              | 16   | 5    | 619    | 637   | 647   | 3      | 2     | 637   | 648     | 3      | 3     | 637           | 652   | 1      | 4              | 638   | 658   | 9      | 4     |
| Russia17e              | 17   | 5    | 760*   | 790   | 821   | 1      | 3     | 781   | 818     | 1      | 3     | 775           | 814   | 1      | 4              | 799   | 821   | 1      | 5     |
| BrazilAM26e            | 25   | 3    | 371*   | 372   | 392   | 1      | 6     | 349   | 366     | 1      | 7     | 342           | 373   | 1      | 9              | 349   | 375   | 1      | 9     |
| BrazilMG30e            | 26   | 3    | 412*   | 431   | 445   | 1      | 8     | 391   | 423     | 1      | 9     | 382           | 423   | 1      | 14             | 394   | 439   | 1      | 13    |
| BrazilCO40e            | 40   | 5    | --     | 636   | 677   | 1      | 22    | 599   | 632     | 1      | 26    | 597           | 648   | 1      | 36             | 609   | 640   | 2      | 32    |
| at48eA                 | 48   | 3    | 25234* | 26535 | 27903 | 1      | 30    | 23268 | 26569   | 1      | 30    | 26663         | 28347 | 1      | 42             | 30457 | 32545 | 1      | 37    |
| at48eB                 | 48   | 4    | 26555* | 27719 | 29658 | 1      | 35    | 26522 | 28678   | 1      | 36    | 27125         | 29705 | 1      | 49             | 29101 | 32801 | 1      | 44    |
| berlins2aA             | 52   | 3    | 5802*  | 6514  | 6803  | 1      | 37    | 6154  | 6694    | 1      | 37    | 6063          | 6807  | 1      | 51             | 7381  | 7888  | 1      | 46    |
| berlins2aB             | 52   | 4    | 6871*  | 6130  | 6537  | 1      | 39    | 5687  | 6240    | 1      | 39    | 6390          | 6652  | 1      | 55             | 7031  | 7294  | 1      | 49    |
| ci70eA                 | 70   | 3    | 1494*  | 1804  | 1902  | 1      | 87    | 1801  | 1867    | 1      | 89    | 1817          | 1906  | 1      | 105            | 2028  | 2093  | 1      | 103   |
| ci70eB                 | 70   | 4    | 1708*  | 1689  | 1809  | 1      | 116   | 1689  | 1752    | 1      | 118   | 1707          | 1792  | 1      | 133            | 1894  | 1991  | 1      | 132   |
| eil76eA                | 76   | 3    | 1563*  | 1898  | 1938  | 1      | 107   | 1853  | 1910    | 1      | 109   | 1833          | 1919  | 1      | 148            | 1920  | 1976  | 1      | 129   |
| eil76eB                | 76   | 4    | 1216*  | 1657  | 1734  | 1      | 124   | 1588  | 1692    | 1      | 126   | 1607          | 1687  | 1      | 171            | 1712  | 1821  | 1      | 145   |

Tabela 2: Resultados para as instâncias não Euclidianas

| Problemas<br>Instância | nCid | nCar | Exato |     |       | GRASP1 |       |      | GraspPR |        |       | Grasp2PBoleto |       |        | Grasp2PProleta |     |       |        |       |
|------------------------|------|------|-------|-----|-------|--------|-------|------|---------|--------|-------|---------------|-------|--------|----------------|-----|-------|--------|-------|
|                        |      |      | GLPK  | Min | Media | #vezes | T (s) | Min  | Media   | #vezes | T (s) | Min           | Media | #vezes | T (s)          | Min | Media | #vezes | T (s) |
| Mauritania10n          | 10   | 2    | 306   | 306 | 306   | 30     | 1     | 306  | 306     | 30     | 1     | 306           | 306   | 30     | 1              | 306 | 306   | 30     | 1     |
| Colombia11n            | 11   | 2    | 461   | 461 | 461   | 30     | 1     | 461  | 461     | 30     | 1     | 461           | 461   | 30     | 1              | 461 | 461   | 30     | 1     |
| Angola12n              | 12   | 2    | 409   | 409 | 409   | 30     | 1     | 409  | 409     | 30     | 1     | 409           | 409   | 30     | 2              | 409 | 409   | 30     | 2     |
| Peru13n                | 13   | 2    | 502   | 502 | 502   | 30     | 1     | 502  | 502     | 30     | 1     | 502           | 502   | 25     | 2              | 502 | 503   | 18     | 2     |
| Libia14n               | 14   | 2    | 504   | 504 | 504   | 10     | 2     | 504  | 504     | 7      | 2     | 504           | 504   | 30     | 2              | 504 | 504   | 13     | 2     |
| BrazilRJ14n            | 14   | 2    | 101   | 101 | 101   | 30     | 2     | 101  | 101     | 30     | 2     | 101           | 101   | 24     | 2              | 101 | 101   | 5      | 2     |
| Congo15n               | 15   | 2    | 573   | 573 | 573   | 21     | 2     | 573  | 573     | 25     | 2     | 573           | 573   | 29     | 2              | 573 | 573   | 30     | 2     |
| Argentina16n           | 16   | 2    | 647*  | 642 | 645   | 4      | 2     | 643  | 645     | 3      | 2     | 648           | 658   | 1      | 3              | 646 | 666   | 1      | 3     |
| EUA17n                 | 17   | 2    | 579   | 589 | 602   | 1      | 2     | 589  | 600     | 2      | 2     | 583           | 590   | 1      | 3              | 580 | 588   | 3      | 3     |
| Bolivia10n             | 10   | 3    | 448   | 448 | 448   | 30     | 1     | 448  | 448     | 30     | 1     | 448           | 448   | 30     | 1              | 448 | 448   | 30     | 2     |
| AfricaSul11n           | 11   | 3    | 537   | 537 | 537   | 30     | 1     | 537  | 537     | 30     | 1     | 537           | 537   | 30     | 2              | 537 | 537   | 30     | 2     |
| Niger12n               | 12   | 3    | 607   | 608 | 614   | 1      | 1     | 609  | 614     | 1      | 1     | 607           | 608   | 11     | 2              | 607 | 613   | 6      | 2     |
| Mongolia13n            | 13   | 3    | 551   | 551 | 551   | 28     | 1     | 551  | 551     | 27     | 1     | 551           | 551   | 30     | 2              | 551 | 551   | 26     | 3     |
| Indonesia14n           | 14   | 3    | 522   | 522 | 522   | 30     | 1     | 522  | 522     | 28     | 2     | 522           | 522   | 30     | 2              | 522 | 522   | 30     | 2     |
| Argelia15n             | 15   | 3    | 619*  | 616 | 620   | 2      | 2     | 617  | 621     | 3      | 2     | 617           | 620   | 2      | 3              | 620 | 628   | 1      | 3     |
| India16n               | 16   | 3    | 734*  | 723 | 729   | 2      | 2     | 723  | 728     | 5      | 3     | 723           | 726   | 5      | 3              | 723 | 731   | 1      | 3     |
| China17n               | 17   | 3    | 651*  | 646 | 659   | 1      | 3     | 646  | 659     | 2      | 3     | 651           | 661   | 2      | 4              | 656 | 665   | 1      | 4     |
| Etiopia10n             | 10   | 4    | 403   | 403 | 403   | 30     | 1     | 403  | 403     | 30     | 1     | 403           | 403   | 30     | 2              | 403 | 403   | 30     | 2     |
| Mali11n                | 11   | 4    | 494   | 494 | 494   | 22     | 2     | 494  | 494     | 17     | 2     | 494           | 494   | 26     | 2              | 494 | 494   | 22     | 2     |
| Chade12n               | 12   | 4    | 654*  | 649 | 650   | 12     | 2     | 649  | 650     | 20     | 2     | 649           | 650   | 18     | 3              | 649 | 655   | 4      | 3     |
| Ira13n                 | 13   | 4    | 697*  | 693 | 693   | 24     | 2     | 693  | 693     | 20     | 2     | 693           | 693   | 12     | 3              | 693 | 694   | 3      | 3     |
| Mexico14n              | 14   | 4    | 620*  | 610 | 610   | 27     | 2     | 610  | 610     | 28     | 2     | 610           | 611   | 21     | 3              | 610 | 611   | 24     | 3     |
| Sudao15n               | 15   | 4    | 793*  | 769 | 773   | 9      | 2     | 769  | 774     | 4      | 3     | 769           | 774   | 3      | 4              | 771 | 778   | 2      | 4     |
| Australia16n           | 16   | 4    | 551*  | 525 | 537   | 1      | 3     | 525  | 534     | 3      | 3     | 525           | 537   | 2      | 4              | 534 | 545   | 1      | 4     |
| Canada17n              | 17   | 4    | 827*  | 827 | 898   | 1      | 3     | 834  | 886     | 1      | 3     | 825           | 859   | 2      | 4              | 834 | 931   | 1      | 4     |
| Arabia14n              | 14   | 5    | 701*  | 688 | 705   | 1      | 3     | 688  | 698     | 1      | 3     | 688           | 694   | 14     | 4              | 688 | 712   | 5      | 4     |
| Cazaquistao15n         | 15   | 5    | 843*  | 833 | 859   | 1      | 3     | 830  | 859     | 1      | 3     | 831           | 855   | 1      | 4              | 834 | 870   | 1      | 4     |
| Brazil16n              | 16   | 5    | 769*  | 742 | 742   | 13     | 3     | 742  | 742     | 16     | 3     | 742           | 742   | 25     | 4              | 742 | 742   | 24     | 4     |
| Russia17n              | 17   | 5    | 820*  | 781 | 790   | 3      | 3     | 779  | 792     | 1      | 4     | 778           | 789   | 1      | 5              | 780 | 803   | 1      | 5     |
| BrazilAM26n            | 25   | 3    | 107*  | 116 | 121   | 1      | 6     | 110  | 65      | 2      | 7     | 109           | 112   | 3      | 9              | 110 | 117   | 1      | 9     |
| BrazilMG30n            | 26   | 3    | 179*  | 171 | 186   | 1      | 9     | 173  | 180     | 4      | 10    | 171           | 179   | 2      | 14             | 171 | 179   | 1      | 13    |
| BrazilCO40n            | 40   | 5    | --    | 444 | 469   | 1      | 25    | 416  | 436     | 2      | 28    | 407           | 439   | 1      | 35             | 452 | 482   | 1      | 35    |
| at48nA                 | 48   | 3    | 616*  | 671 | 751   | 1      | 32    | 684  | 723     | 2      | 33    | 648           | 736   | 1      | 38             | 738 | 799   | 1      | 37    |
| at48nB                 | 48   | 4    | 484*  | 655 | 712   | 1      | 45    | 618  | 673     | 1      | 42    | 659           | 699   | 1      | 49             | 729 | 775   | 1      | 50    |
| berlins2nA             | 52   | 3    | 812*  | 953 | 1026  | 1      | 38    | 938  | 995     | 1      | 39    | 923           | 983   | 1      | 46             | 912 | 1077  | 2      | 45    |
| berlins2nB             | 52   | 4    | 520*  | 638 | 716   | 1      | 49    | 646  | 691     | 1      | 49    | 651           | 695   | 1      | 58             | 679 | 772   | 1      | 59    |
| ci70nA                 | 70   | 3    | 764*  | 993 | 1203  | 1      | 81    | 1055 | 1162    |        |       |               |       |        |                |     |       |        |       |

*Autares*

**ALINE MAKI IGARI**

Cursando 10º período de Engenharia de Produção no Centro Universitário Senac com previsão de conclusão em 2018; Atualmente Analista na área de Mercados Capitais no Banco Original; Participação na ENEGEP no ano de 2017 com a apresentação de um pôster do projeto: Modelos matemáticos para otimização do processo de montagem da cabine de pintura. Realização de monitoria em Física 1 no ano de 2015.

**ALVARO FARIA**

Concluiu o PhD em Estatística em Warwick no ano de 1996. Atualmente é professor de Estatística no Departamento de Matemática e Estatística da OU (Universidade Aberta, Milton Keynes, Buckinghamshire, Inglaterra). Tem interesse em modelos de séries temporais Bayesianas multivariadas, modelagem Bayesiana de processos espaço-temporais, modelos Bayesianos Dinâmicos para processos auto-regressivos não-lineares, Estatística Multivariada, Estatística Bayesiana.

**ANA PAULA NUNES DUARTE**

Graduada em Engenharia de Transportes em Logística pela Universidade Federal de Santa Catarina. Desenvolveu pesquisas nas áreas de roteirização e programação de veículos na área da saúde e simulação de transportes públicos.

**ANDERSON LEVATI AMOROSO**

Possui graduação em Engenharia Elétrica pela Universidade Federal de Santa Catarina (1995) e mestrado em Engenharia e Tecnologia Espaciais pelo Instituto Nacional de Pesquisas Espaciais (1999). Atualmente é professor assistente da Universidade Tecnológica Federal do Paraná (UTFPR) com atuação na área de Automação.

**ANDERSON RUFINO**

Graduado em Ciência da Computação pela Universidade Federal da Paraíba, onde cursa o Mestrado na mesma área, com ênfase em Otimização Combinatória. Atualmente, pesquisa sobre métodos computacionais exatos e heurísticos para composição de solos de guitarra.

**ANGELA CRISTINA GOMES**

Possui mestrado em Contabilidade Gerencial pela FUCAPE- Fundação Instituto Capixaba de Pesquisa em Contabilidade, Economia e Finanças (2014). MBA em Gestão Empresarial pela FGV -Fundação Getúlio Vargas (2006), Graduação em Ciências Contábeis (2001). Doze anos de experiência como Contadora. Atualmente é Professora e Coordenadora dos Cursos de Graduação de Ciências Contábeis, Logística e Gestão de Recursos Humanos da FAVI- Associação Vitoriana de Ensino Superior, Perita Contadora da 10ª Vara do Trabalho de Vitória/ES e integrante da Comissão de Coordenadores de Cursos de Ciências Contábeis do Conselho Regional de Contabilidade - CRC/ES (Portaria Nº49/2016).

**ÂNGELO MÁRCIO OLIVEIRA SANTANNA**

Possui graduação em Estatística pela Universidade Federal da Bahia (2003), mestrado em Engenharia de Produção pela Universidade Federal do Rio Grande do Sul (2006) e doutorado em Engenharia de Produção pela Universidade Federal do Rio Grande do Sul (2009). Atualmente, é professor adjunto da Universidade Federal da Bahia. Atua como professor colaborador da Pontifícia Universidade Católica do Paraná (Programa de Pós-Graduação em Engenharia de Produção e Sistemas). É revisor de periódicos internacionais e nacionais. Tem experiência na área de Estatística e Engenharia de Produção, com ênfase em Planejamento,

Otimização e Controle de Sistemas de Produção; atuando principalmente nos temas: engenharia da qualidade, engenharia do produto, modelagem e otimização de processos.

#### **ANTÔNIO CLÉCIO FONTELLES THOMAZ**

Possui graduação em Matemática pela Universidade Federal do Ceará (1969), mestrado em Informática pela Pontifícia Universidade Católica do Rio de Janeiro (1973) e doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1986). Pós-Doutorado na Universidade Joseph Fourier - Gfenoble França em 1995. Atualmente é Prof. Titular Aposentado da Universidade Federal do Ceará e Prof. adjunto da Universidade Estadual do Ceará. Tem experiência na área de Matemática, com ênfase em Matemática Discreta e Combinatória

#### **CAMILA CAMPOS COLARES DAS DORES**

Possui graduação em Ciência da Computação pela Universidade Estadual do Ceará (2010), especialização em Gerência de Projetos de TI pela Faculdade Farias Brito (2014) e mestrado em Ciência da Computação pela Universidade Estadual do Ceará (2017). Atualmente é professora e coordenadora dos cursos de Redes de Computadores e Análise e Desenvolvimento de Sistemas da Faculdade CDL e é professora substituta do curso de Ciência da Computação da Universidade Estadual do Ceará. Tem experiência na área de Matemática Computacional, com ênfase em algoritmos e otimização combinatória.

#### **CHRISTIANE WENCK NOGUEIRA FERNANDES**

Possui Graduação em Matemática pela Universidade Federal de Santa Catarina (1999), Mestrado em Engenharia de Produção pela Universidade Federal de Santa Catarina (2002), Doutorado em Engenharia de Produção pela Universidade Federal de Santa Catarina (2010). Atualmente é Professora Adjunto IV do Departamento de Engenharias da Mobilidade da Universidade Federal de Santa Catarina. Tem experiência na área de Engenharia de Transportes, Pesquisa Operacional e Logística.

#### **DIOGO KENJI NOMURA**

Estudante de Engenharia de Produção do Centro Universitário Senac, possui noção básica em AutoCAD, experiência em projetos com temas relacionados ao curso de graduação como: controle de produção, controle de estoque, modelagem, protótipos da linha, entre outros.

#### **EDISON LUIZ BONOTTO**

Bacharel em Ciência da Computação pela Universidade Federal do Tocantins - UFT (2013). Mestre em Informática no Programa de Pós-Graduação em Informática da Universidade Federal da Paraíba - UFPB (2017), atuando na linha de pesquisa Computação Distribuída. Doutorando no Programa de Pós-Graduação em Ciência e Engenharia de Materiais da Universidade Federal da Paraíba - UFPB. Tem experiência na área de otimização combinatória, meta-heurísticas, computação gráfica e computação paralela.

#### **EDUARDO LUIS FESTA**

Graduado no curso de Bacharelado em Engenharia de Transportes e Logística pela Universidade Federal de Santa Catarina (UFSC) no ano de 2016. Trabalha no Departamento de Trânsito e Engenharia da Prefeitura Municipal de Balneário Camboriú, ocupando o cargo de Coordenador de Sinalização Viária. Empresário, atua como Diretor Técnico e Comercial de uma empresa do ramo de engenharia de tráfego, trânsito e transportes.



**EDUARDO VIEIRA QUEIROGA**

Graduado e Mestre em Ciência da Computação pela Universidade Federal da Paraíba. Atualmente é Doutorando em Ciência da Computação pela Universidade Federal Fluminense, com ênfase em Otimização Combinatória e Programação Linear Inteira.

**ELIZABETH F. G. GOLDBARG**

Possui graduação em Engenharia Industrial Mecânica pelo Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (1985), mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia (1993) e doutorado em Engenharia de Produção pela COPPE-Produção(UFRJ) (2001). Atualmente é professora na Universidade Federal do Rio Grande do Norte no Departamento de Informática e Matemática Aplicada (DIMAp) e atua no Programa de Pós-graduação em Sistemas e Computação (PPgSC) na linha de Algoritmos Experimentais. Os atuais interesses de pesquisa são Problemas de Otimização Combinatória e Meta-heurísticas.

**ELIZABETH FERREIRA GOLDBARG**

Possui graduação em Engenharia Industrial Mecânica pelo Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (1985), mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia (1993) e doutorado em Engenharia de Produção pela COPPE-Produção (UFRJ) (2001). Atualmente é professora na Universidade Federal do Rio Grande do Norte no Departamento de Informática e Matemática Aplicada (DIMAp) e atua no Programa de Pós-graduação em Sistemas e Computação (PPgSC) na linha de Algoritmos Experimentais. Os atuais interesses de pesquisa são Problemas de Otimização Combinatória e Metaheurísticas.

**EVERTON LUÍS GARCIA**

Possui graduação em Engenharia Civil pela Universidade Federal do Paraná (2009), está cursando mestrado em Métodos Numéricos em Engenharia (Estruturas) pela Universidade Federal do Paraná, está cursando MBA em Gerenciamento de Projetos pela Fundação Getúlio Vargas. Atuou como professor pela Universidade Tecnológica Federal do Paraná, no departamento de Construção Civil, subárea estruturas, onde ministrou as disciplinas de Resistência dos Materiais, Fundações, Informática e Programação aplicada, Mecânica dos Solos, Atuou com coordenação e gerencia de projetos voltados para engenharia civil. Atualmente atua como analista de projetos e pesquisador em projetos de Pesquisa e Desenvolvimento na Fundação Parque tecnológico Itaipu nas áreas de conhecimento: Concreto e Sustentabilidade, Engenharia de Estruturas, Método dos elementos Finitos, Análise numérico-computacional de estruturas, Análise de materiais cimentícios para reparos em estruturas hidráulicas, Modelagem tridimensional de estruturas de Engenharia, Segurança de barragens.

**FÁBIO LUIZ USBERTI**

Atua como Professor Doutor em regime de dedicação exclusiva no Instituto de Computação (IC) da Universidade Estadual de Campinas (UNICAMP). Possui graduação em Engenharia Agrícola (2004) e Ciência da Computação (2010) pela UNICAMP. Concluiu o mestrado em Engenharia Elétrica (2007) e o doutorado em Engenharia Elétrica (2012) na Faculdade de Engenharia Elétrica e de Computação (FEEC) da UNICAMP, onde também concluiu sua pesquisa de pós-doutoramento. Possui interesse nas áreas de Pesquisa Operacional, Otimização Combinatória e Otimização em Redes de Energia.

**FLAVIA HAWEROTH**

Graduada em Engenharia de Transportes em Logística pela Universidade Federal de Santa Catarina. Desenvolveu pesquisas nas áreas de Cadeia do Frio e Roteirização e Programação de Veículos em setores de saúde e transporte público.

**FLÁVIO VASCONCELOS DA SILVA**

Livre-Docente em Controle de Processos (UNICAMP-2012), Doutor em Engenharia de Alimentos (UNICAMP-2003), Mestre em Engenharia de Alimentos (UNICAMP-1998) e Graduado em Engenharia Química (UFS-1994). Atua como Professor Associado (MS5) no Departamento de Engenharia de Sistemas Químicos da Faculdade de Engenharia Química UNICAMP. Tem experiência na área de Engenharia Química, com ênfase em Automação de Processos Industriais, atuando principalmente nos seguintes temas: Controle de Processos, Inteligência Artificial (Controle Fuzzy e Redes Neurais), Instrumentação Industrial e Refrigeração Industrial.

**GABRIELA CORREIA COLAÇO**

Graduada em Ciência e Tecnologia (2018) pela Universidade Rural do Semi-Árido. Cursos na área de AutoCad pela Prepara Cursos (2016). Atualmente cursa Engenharia de Produção pela Universidade Rural do Semi-Árido.

**GEOVANA APARECIDA FRANÇA DOS SANTOS**

Bolsista no Programa de Pós-Graduação em Matemática Pura e Aplicada na Universidade Federal de Santa Catarina, licenciada em Matemática pela Universidade Estadual do Paraná e tem experiência na área de Matemática Aplicada, com ênfase em Otimização sob Incertezas.

**GERARDO VALDISIO RODRIGUES VIANA**

Possui Licenciatura em Matemática pela Universidade Estadual do Ceará (1976), Graduação em Engenharia Mecânica pela Universidade Federal do Ceará (1976), Mestrado em Ciência da Computação pela Universidade Federal do Ceará (1996) e Doutorado em Ciência da Computação pela Universidade Federal do Ceará (2007) com período Sandwich no IME/USP - Instituto de Matemática e Estatística da Universidade de São Paulo. Aposentou-se como Professor do Departamento de Computação do Centro de Ciências da UFC em Agosto de 2010. Atualmente é professor associado da Universidade Estadual do Ceará e vice-coordenador do Curso Bacharelado em Ciência da Computação. Tem experiência na área de Computação, com ênfase em Análise de Algoritmos e Pesquisa Operacional, atuando principalmente nos seguintes temas: otimização combinatória, programação matemática e meta-heurísticas.

**GISELE TESSARI SANTOS**

Doutora em Engenharia Agrícola pela Universidade Federal de Viçosa (2012). Mestre em Engenharia de Produção pela Universidade Federal de Minas Gerais (2008). Atualmente é professora adjunta do Mestrado em Engenharia e Gestão de Processos e Sistemas do Instituto de Educação Tecnológica - Ietec e professora adjunta do curso de Engenharia de Produção do Instituto Brasileiro de Mercado de Capitais- Ibmecc. Tem experiência nas seguintes áreas: modelagem e simulação de processos, pesquisa operacional, análise numérica aplicada, fenômenos de transporte, termodinâmica, secagem de grãos e engenharia financeira.

**GLÁUCIA REGINA DE OLIVEIRA ALMEIDA**

Graduada em Engenharia de Produção pela Faculdade de Administração e Negócios de Sergipe - Fanese. Mestre em Engenharia de Processos pela Universidade Tiradentes. Doutorado em Engenharia de Processos. Consultora em Ergonomia. Desenvolvimento de projetos integradores na área de Engenharia de Produção voltados para o desenvolvimento social. Membro de Colegiado do curso de Engenharia de Produção. Atualmente faz parte do Núcleo Docente Estruturante do curso de Engenharia de Produção da Universidade Tiradentes. Professora Adjunta dos cursos de Engenharia de Produção, Engenharia Civil, Engenharia Elétrica, Engenharia Mecatrônica e Engenharia de Petróleo. Possui certificação nível 1 nas ferramentas da G Suite for Education (Google Certified Educator Level 1).

**GUSTAVO HENRIQUE MORESCO**

Engenheiro de Transportes e Logística formado pela Universidade Federal de Santa Catarina e pós-graduando MBA em Logística e Supply Chain pela Fundação Getúlio Vargas. Atuando desde 2017 como Analista de Logística, no Grupo Tigre, em projetos logísticos e gestão de indicadores de desempenho com foco no aumento da eficiência e eficácia no processo logístico.

**HELLEN MARIANY SANTOS**

Graduanda em Engenharia de Produção pela Universidade Tiradentes (UNIT).

**JAIRO MARLON CÔRREA**

Possui graduação em Licenciatura Em Matemática pela Universidade Estadual do Oeste do Paraná (2004), Mestrado (2007) e Doutorado (2015) em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná, Pós-Doutoramento em Engenharia Elétrica e Computação pela UNIOESTE (2017). Atualmente é docente do Magistério Superior na UTFPR (Medianeira-Pr). Tem experiência na área de Previsões de Séries Temporais (Redes Neurais Artificiais, Modelos SARIMAX, Modelos GARCH, incluindo Decomposição Wavelet).

**JARBAS ANCELMO DA SILVA JÚNIOR**

Formado em Engenharia de Produção pela Universidade Federal de Goiás - Regional Catalão, foi estagiário da HPE Automotores do Brasil, integrante do Grupo de Estudos em Modelagem e Simulação (GEMS) 2014-2018, e do projeto de Iniciação Científica em Simulação de Sistemas de Coordenação de Ordens em ambientes de produção flow-shop.

**JESSICA INGRID DO NASCIMENTO SALES**

Cursando Engenharia de Produção no Centro Universitário SENAC com término em dezembro de 2018. Noção básica no Access e no AutoCAD. Onde foram realizados 8 projetos ao longo do curso com temas de economia, linha de produção, logística, sustentabilidade, entre outros. Experiência de dois anos na Pepsico do Brasil como aprendiz onde executava atividades de rotina de escritório. Atuação de dois anos na Editora Globo na área de planejamento e controle orçamentário.

**JOÃO PAULO DE MORAES PALUDO**

Possui graduação em Engenharia de Controle e Automação pela Universidade Federal de Santa Catarina (UFSC, 2012), e mestrado em Engenharia Aeronáutica pelo Instituto Tecnológico de Aeronáutica (ITA, 2016). Atualmente trabalhou como engenheiro de desenvolvimento de produto na EMBRAER S.A.

### **JOÃO PAULO PORDEUS GOMES**

Holds a bachelor's degree in Electrical Engineering from Universidade Federal do Ceará (UFC, 2004), Brazil; a master's (2006) degree on Aeronautical Engineering, and doctorate's (2011) degree in electronic engineering from Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, SP, Brazil. Dr. Gomes worked for EMBRAER S.A. between 2006 and 2013, as a Technology Development Engineer focusing on fault monitoring applications on aeronautical systems. He is currently an Assistant Professor at UFC.

### **JORDANA RAMALHO DE SOUSA**

Graduada recentemente em Engenharia de Produção, pelo Centro Universitário do Estado do Pará – Cesupa, atualmente exercendo o cargo de Analista de Planejamento da Agroindústria Biopalma da Amazônia, tendo como uma das principais funções controle orçamentário, geração e análise de relatórios de rendimentos e custos, elaboração de procedimentos, desenvolvimento e condução de reuniões com supervisores e gerentes de todas as áreas da empresa, tendo contato direto com indicadores de performance. Dedicou o período acadêmico a projetos sociais e a pesquisas relacionadas, principalmente, à Pesquisa Operacional e Métodos Multicritérios de Apoio à Decisão, sendo membro do grupo de pesquisa de Tomada de Decisão em Cenários Complexos.

### **JOSÉ HENRIQUE DE FREITAS GOMES**

Engenheiro de Produção Mecânica pela Universidade Federal de Itajubá (2003-2007), com mestrado (2008-2010) e doutorado (2010-2013) em Engenharia de Produção pela mesma instituição. Atualmente professor Adjunto I do Instituto de Engenharia de Produção e Gestão (IEPG) da Universidade Federal de Itajubá, nos cursos de graduação em Engenharia de Produção e Administração, e de pós-graduação em Engenharia de Produção. Atua nas linhas de pesquisa de modelagem, análise e otimização de sistemas de produção e operações de manufatura, cujos interesses são voltados para a aplicação e/ou o aprimoramento de métodos multi-objetivo (fronteiras de Pareto, funções de aglutinação, técnicas de priorização de funções), métodos estatísticos (Projeto e Análise de Experimentos - DOE, Metodologia de Superfície de Resposta, Análise de Componentes Principais, Metodologia Seis Sigma), métodos de programação matemática (otimização linear, não-linear e inteira) e técnicas de administração de materiais (gestão de estoques, previsão de demanda, MRP).

### **JOSÉ RICARDO MENEZES OLIVEIRA**

Mestre em Engenharia de Produção pela Universidade Federal de Pernambuco (UFPE) Conceito 6 (CAPES). Possui graduação em Desenho Industrial pela Universidade Federal da Paraíba (1995), é pós-graduado em Tecnologia Educacional, MBA em Marketing e Vendas e pós graduando em Gerenciamento de Projetos. Leciona desde 1996 em IES nos cursos de Engenharia de Produção, Administração, Petróleo e Gás, e Segurança e Saúde no Trabalho. Coordenador do Curso de Engenharia de Produção da Universidade Tiradentes - UNIT (2013 - atual).

### **JULIANO RESENDE BUCCHIANERI**

Possui graduação em Engenharia Elétrica pela Universidade de Uberaba (2004). Com especialização em Orçamentos e Finanças Públicas pela CESUC (Cuiabá). Com experiência em automação industrial em empresas nacionais e multinacionais e experiência em grupos de pesquisa e na docência. Atualmente é professor do Instituto Federal de Educação Ciência e Tecnologia de Rondônia. Mestrado em andamento em Engenharia de Produção e Sistemas (MEPROS) pela PUC - GO (2016). Está envolvido no Grupo de Pesquisas do IFRO em mecatrônica e pesquisas em Engenharia de Produção e modelagem Matemática.

**LARA CRISTINE PEREIRA DOS SANTOS**

Possui graduação em Engenharia de Produção pela Universidade Federal de Goiás (2018). Tem experiência na área de Engenharia de Produção.

**LAURA LUIZA PEREIRA DAVID**

Possui graduação em Engenharia de Produção pela faculdade IBMEC (2017), e cursa MBA em Finanças pela faculdade IBMEC. Atualmente atua em projetos de consultoria de gestão utilizando o método PDCA com foco para resultados financeiros em clientes do segmento público e privado. Tem experiência em planejamento estratégico e logístico, simulação de cenários e processos operacionais, pesquisa operacional, estatística, e modelagem de problemas de otimização.

**LEANDRO ANDRÉ CARDOSO DE SOUZA**

Mestre em Ciências Contábeis pela UFES (ano 2017). Possui graduação em Administração de Empresas pela Pontifícia Universidade Católica do Rio de Janeiro (2002). Atualmente é Auditor Fiscal do Tesouro Municipal da Prefeitura Municipal de Vitória.

**LEONARDO RAMOS RODRIGUES**

Holds a bachelor's degree in Electrical Engineering from Universidade Federal do Espírito Santo (UFES, 2003) Brazil, a Master's Degree (2008) and a D.Sc. Degree (2013) on Aeronautical Engineering from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is currently with Instituto de Aeronáutica e Espaço (IAE) in São José dos Campos - Brazil, where he performs research in the field of aerospace technologies. He was formerly with Embraer S.A. from 2006 to 2015, working with the PHM research group performing research in the fields of Prognostics and Health Management and Data Analytics solutions for the aviation industry.

**LUCAS SOUSA DE OLIVEIRA**

Holds a bachelor's degree in Mechatronics Engineering from Universidade de Brasília (UnB, 2013) Brazil and a Master's Degree (2017) on Aeronautical Engineering from Instituto Tecnológico de Aeronáutica (ITA), Brazil.

**LUCIDIO DOS ANJOS F. CABRAL**

Possui graduação em Ciência da Computação pela Universidade Federal do Ceará (1990), mestrado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1993) e doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (2001). Atualmente é professor Titular do Departamento de Computação Científica do Centro de Informática da Universidade Federal da Paraíba. Tem experiência na área de otimização combinatória, meta-heurísticas e computação paralela.

**LUCÍDIO DOS ANJOS FORMIGA CABRAL**

Possui graduação em Ciência da Computação pela Universidade Federal do Ceará (1990), mestrado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1993) e doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (2001). Atualmente é professor Titular do Departamento de Computação Científica do Centro de Informática da Universidade Federal da Paraíba. Tem experiência na área de otimização combinatória, meta-heurísticas e computação paralela.

### **LUIZ ALBINO TEIXEIRA JÚNIOR**

Possui curso técnico em Informática Industrial, CEFET (2003-2004); graduação em Matemática, FIC (2002-2005); especialização em Métodos Estatísticos Computacionais, UFJF (2006); mestrado e doutorado em Engenharia Elétrica, com Ênfase em Métodos de Apoio à Decisão, PUC-RIO (2007-2009 e 2009-2013, respectivamente); e pós-doutoramento em Métodos Numéricos Inteligentes em Segurança de Barragens, no Departamento de Engenharia Elétrica, na UNIOESTE (2014-2015). No Departamento de Matemática, Estatística e Computação, na The Open University (UK), em Milton Keynes, Inglaterra, fez parte de seu doutorado. Atualmente, é professor adjunto de Estatística, na Universidade Federal da Integração Latino Americana (UNILA), em Foz do Iguaçu, Paraná, Brasil; coordenador de um grupo de pesquisa multidisciplinar focado em Métodos Numéricos Inteligentes (MNIs) no Centro de Estudos Avançados em Segurança de Barragens (CEASB), na Fundação Parque Tecnológico Itaipu (FPTI); consultor em BI e em Segurança de Barragens; e revisor de periódicos. Tem interesse Business Intelligence (BI); Segurança de Barragens (hidrelétricas e de rejeitos); Simulação; e Realidade Aumentada em Barragens.

### **MAGNO DOS SANTOS NETO**

Mestre em Ciências Contábeis pela UFES (ano 2017). Possui graduação em Bacharelado em Estatística pela Universidade Federal do Espírito Santo (2010) e pós-graduação em Docência do Ensino Superior. Atualmente é estatístico no Tribunal de Justiça do Estado do Espírito Santo e professor nas faculdades FAVI (Faculdade de Vitória) e FACES (Faculdade do Espírito Santo). Foi professor substituto da Universidade Federal do Espírito Santo 2011-2012. Tem experiência na área de Estatística, com ênfase em Séries Temporais, Processos de Contagem, atuando principalmente nos seguintes temas: modelagem de processos de contagem GLARMA, cálculo de risco relativo, modelos GARCH e implementação na plataforma R

### **MARCO C. GOLDBARG**

Possui graduação na Academia Militar de Agulhas Negras (1975), graduação em Engenharia de Fortificação e Construção pelo Instituto Militar de Engenharia (1982), mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia (1987), doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1990), pós-doutorado pela Universidade Federal de Minas Gerais (1999). Atualmente é Professor Titular da Universidade Federal do Rio Grande do Norte. Tem experiência na área de Sistemas e Computação e Engenharia de Produção, com ênfase em Pesquisa Operacional. Atua principalmente nos seguintes temas: Algoritmos Experimentais e Meta-heurísticos, Modelagem em Otimização Combinatória, Aplicações da Pesquisa Operacional nos setores de: Petróleo e Gás, Roteamento de Veículos, Programação Multiobjetivo e Planejamento de Tratamentos com uso de Radioterapia.

### **MARCO CÉSAR GOLDBARG**

Possui graduação na Academia Militar de Agulhas Negras (1975), graduação em Engenharia de Fortificação e Construção pelo Instituto Militar de Engenharia (1982), mestrado em Sistemas e Computação pelo Instituto Militar de Engenharia (1987), doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (1990), pós-doutorado pela Universidade Federal de Minas Gerais (1999). Atualmente é Professor Titular da Universidade Federal do Rio Grande do Norte. Tem experiência na área de Sistemas e Computação e Engenharia de Produção, com ênfase em Pesquisa Operacional. Atua principalmente nos seguintes temas: Algoritmos Experimentais e Metaheurísticos, Modelagem em Otimização Combinatória, Aplicações da Pesquisa Operacional nos setores de: Petróleo e Gás, Roteamento de Veículos, Programação Multiobjetivo e Planejamento de Tratamentos com uso de Radioterapia.

**MARCOS WANDIR NERY LOBÃO**

Graduação em Engenharia Química pela Universidade Federal de Sergipe (1985), mestrado em Engenharia Química pela Universidade Federal do Rio de Janeiro (1990) e doutorado em Engenharia Industrial pela Universidade Federal da Bahia (2012). Atualmente é Diretor de Educação Corporativa da Universidade Tiradentes - UNIT/SE. Tem experiência na área de Engenharia Química, com ênfase em Operações Industriais e Equipamentos. Na área de gestão de ciência, tecnologia e inovação, tem experiência no setor público e privado, gerenciando: a Pró-Reitoria de Pós-Graduação e Pesquisa da UNIT/SE (1997-2002), a Fundação Estadual de Amparo à Pesquisa de Sergipe (2003-2005), o Instituto de Tecnologia e Pesquisa (2005-2008) e o Sergipe Parque Tecnológico (2009-2015).

**MARIA MADALENA GUERRA FERREIRA**

Engenheira de Produção, recém-formada, atualmente fazendo MBA em Controladoria e Processos Decisórios pela Pontifícia Universidade Católica de Minas Gerais. Apaixonada por desafios, e pela geração de seus respectivos resultados. Participante do grupo de pesquisa de tomada de decisão em cenários complexos, onde fora abordado algoritmos das escolas francesa e americana (AHP, Electre e Promethee). Possui domínio das ferramentas de qualidade: Fluxograma de Processos, Carta/Gráfico de Processos, PDCA/MASP, Ishikawa, Programa 5S e Kaizen. E experiência em mapeamento de processos, elaboração de procedimentos, escrita de relatórios gerenciais e contratuais, justificativas de estouros contábeis e assessoria do planejamento estratégico.

**MATHEUS DA SILVA MENEZES**

Possui graduação em Licenciatura em Matemática (2000) e graduação em Bacharelado em Ciências da Computação (2006), mestrado em Matemática pela Universidade Federal da Paraíba (2006) e Doutorado em Ciência da Computação pela UFRN (2014). Atualmente é professor adjunto da Universidade Federal Rural do Semi-Árido. Tem experiência na área de Matemática, com ênfase em Matemática Aplicada, atuando principalmente nos seguintes temas: otimização métodos numéricos, otimização combinatória e metaheurísticas.

**MICHELE MARTINS DA SILVA**

Graduanda em Engenharia de Produção no Centro Universitário Senac, realizou monitoria na disciplina Pesquisa Operacional - Modelos Estocásticos, e no momento, estagiando em Excelência Operacional na União Química Farmacêutica Nacional, no desenvolvimento de projetos relacionados às máquinas, definição de métodos, processos no controle e planejamento da produção.

**NAILSON DOS SANTOS CUNHA**

Tecnólogo em Desenvolvimento de Sistemas para Internet pelo Instituto Federal de Educação, Ciência e Tecnologia da Paraíba. Mestre em Informática pela Universidade Federal da Paraíba. Atualmente trabalha como desenvolvedor de sistemas na Universidade Federal da Paraíba.

**NATÁLIA CIBELE DE SOUSA SANTOS**

Graduanda em Engenharia de Produção com conclusão prevista para Dezembro de 2018, foi estagiária da HPE Automotores do Brasil na área de logística, tem experiência em 5S, controle de terceiros, projetos e processos logísticos. Foi Integrante do Grupo de Estudos em Modelagem e Simulação e do projeto de Iniciação Científica em Simulação de Sistemas de Coordenação de Ordens em ambientes de produção flow-shop.

**NAYARA FELÍCIO DE OLIVEIRA**

Tem experiência na área de Engenharia de Produção.

**NILSON CARVALHO DA MATA**

Engenheiro Civil pela PUC-GO (1999) e especialista em Gestão de Projetos (2010). Mestrado em andamento em Engenharia de Produção e Sistemas pela PUC - GO (2016) / MEPROS (Conceito CAPES 3). Experiência na área de Engenharia Civil, com ênfase em gerenciamento e coordenação planejamentos e logística desde 1997 em diversas frentes de trabalho tais como infraestrutura rodoviária, urbana, edificações horizontais e verticais, pch, saneamento e gerenciamentos de grandes obras com portfólio de mais de 250 empreendimentos. Atuação como professor de pós-graduação no Brasil através do Instituto de Pós Graduação IPOG em Lato Sensu (desde 2014) com aprovação superior a 97% em vários módulos voltados ao gerenciamento de projetos, planejamento e obras para formação de engenheiros e arquitetos. Realização de várias palestras no Brasil sobre gerenciamento e planejamento.

**PATRÍCIA AGNES PEREIRA DA SILVA**

Engenheira de Produção Mecânica com dupla diplomação pela Universidade Federal de Itajubá (UNIFEI) e pela universidade francesa Université de Technologie de Compiègne (UTC). Mestre em Engenharia de Produção na área de Modelagem, Otimização e Controle pela UNIFEI. Atuou como pesquisadora no projeto "Definição de Método de Modelagem Matemática para Criação de um Simulador de Consumo de Energia Elétrica de um Sistema de Refrigeração" e, posteriormente, coordenou o projeto de inovação tecnológica "Sistema de controle inteligente para ar condicionado", ambos realizados com o apoio do Programa de Pesquisa Inovativa em Pequenas Empresas (PIPE) da FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) no âmbito da RVT Energia, empresa de consultoria em gestão energética para sistemas de refrigeração.

**PAULO HENRIQUE DOS SANTOS**

Tecnólogo em Logística pelo IFG – Instituto Federal de Goiás (2014), com especialização em Docência Universitária pela Faculdade Católica de Anápolis e em Logística Empresarial pela FACEL/PR (2017), mestrando em Engenharia de Produção e Sistemas pela PUC/GO. Atua como Professor EBTT no IFES – Instituto Federal do Espírito Santo, pesquisa as relações interorganizacionais, especificamente as Redes de Cooperação Empresariais (RCEs) e Arranjos Produtivos Locais (APLs), em parceria com a universidade UNISINOS/RS. Tem experiência de sete anos na área de Logística com ênfase em análise de transportes e estoque.

**PEDRO PAULO GONDIM SIMMER**

Mestre em Gestão Pública pela Universidade Federal do Espírito Santo (UFES), Especialista em Finanças Cooperativas pela Fundação Getúlio Vargas (FGV), Administrador de Empresas. Atualmente é administrador na Assessoria de Planejamento, Orçamento e Gestão Estratégica do Tribunal de Justiça do Estado do ES e Coordenador e professor do curso superior de Administração das faculdades FAVI e FACES. Tem experiência na área de Administração Financeira, Orçamentária, Custos, Planejamento Estratégico e Gestão de Contratos, com ênfase em Organizações Públicas.

**PRISCILA HARUMI SOUZA**

Estudante de Engenharia de Produção do Centro Universitário Senac, possui experiência em controle de produção, estoque e modelagem. Atualmente, Assistente Administrativo na ALD Automotive Brasil no segmento de gestão e terceirização de frotas corporativas.



**RAFAEL KENDY ARAKAKI**

Mestre em ciência da computação (2016) pela universidade estadual de campinas. Concluiu graduação em ciência da computação (2014) pela universidade federal de mato grosso do sul com período sanduíche (2012-2013) em universidade do porto. Possui interesse na área de pesquisa de otimização combinatória, em especial problemas de roteamento e suas aplicações.

**RANMSÉS E. M. BASTOS**

Possui graduação em Engenharia de Computação pelo Instituto Militar de Engenharia (2010) e mestrado em Sistemas e Computação pela Universidade Federal do Rio Grande do Norte (2017). Atualmente, é doutorando na mesma Universidade e seus interesses de pesquisa incluem Otimização Combinatória, Modelagem Matemática e Algoritmos Experimentais. Atua como Analista de Sistemas - Infraestrutura na Petrobras e possui experiência em Computação de Alto Desempenho e gerenciamento de sistemas operacionais Red Hat.

**RENATA DE CÁSSIA FERREIRA SILVA**

Engenheira em Eficiência Energética na empresa GreenYellow, co-fundadora da empresa RVT Energia e mestranda em Engenharia Química na UNICAMP tendo como projeto o desenvolvimento de simulador para diagnóstico de eficiência energética em sistemas de refrigeração. Especialista em Administração de Empresa pela FGV (Fundação Getúlio Vargas) e bacharel em Engenharia Mecânica pela UNIFEI (Universidade Federal de Itajubá). Atuou como Especialista em Eficiência Energética no Centro de Engenharia da AMBEV e em outras atividades relacionadas ao tema por mais de 7 anos no Brasil e no Exterior.

**RICARDO LUIZ FERNANDES BELLA**

Professor Substituto do Magistério Superior no Departamento de Engenharia de Produção na Escola de Engenharia da Universidade Federal Fluminense (UFF). Qualificado em Engenharia de Produção pelo Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ). Mestre em Engenharia de Produção com ênfase em Gestão de Riscos e Projetos pela Universidade Federal Fluminense (UFF). Doutor em Engenharia de Produção com ênfase em Ergonomia Organizacional pela Universidade Federal Fluminense (UFF).

**RUY GOMES SILVA**

Mestre em Engenharia de Produção e Sistemas pela PUCPR, Graduado em Engenharia de Produção – CESUPA, Coordenador do grupo de pesquisa de Tomada de decisão e Simulação de Processos, Coordenador do curso de Especialização em Engenharia da Qualidade em Belém – PA, Professor de Pós-Graduação, Professor e membro efetivo do Núcleo Docente Estruturante – NDE dos cursos de Graduação de Engenharia de Produção e Engenharia Civil na FAMAZ – Belém. Experiência de 15 anos na indústria de beneficiamento de minério, na empresa HYDRO Alunorte, atuado nas funções de Coordenador de Produção, Gerente operacional e Especialista de Processo. Com experiência em SDCD (comando digital distribuído a distância) e extenso conhecimento nos processos e operações de “chão de fábrica”).

**SAMUEL BELLIDO RODRIGUES**

Possui graduação em Licenciatura Matemática pela Universidade Estadual do Oeste do Paraná (2004), mestrado em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná (2007), doutorado em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná (2015); e pós-doutoramento em Métodos Numéricos Híbridos de Previsão. Atualmente é professor adjunto da Universidade Tecnológica Federal do Paraná. Tem

experiência na área de Previsões de Séries Temporais (Support Vector Regression, Redes Neurais, Modelos SARIMA, Modelos GARCH, incluindo Decomposição Wavelet).

### **SAMUEL NATIVIDADE FERREIRA JUNIOR**

Engenheiro de Produção pelo CESUPA, foi membro voluntário do Grupo de Pesquisa Tomada de Decisão Gerencial, onde foram abordados algoritmos como AHP, PROMETHEE e ELECTRE. Fora isso, possui experiência no setor de Logística Fluvial, Manutenção e atualmente atuou no setor de Exportação. Durante minha trajetória pude desenvolver as habilidades de: Qualidade, Manutenção e Custos

### **SAULO ANDERSON FREITAS DE OLIVEIRA**

Holds a bachelor's (2013) degree in Computer Science and master's (2016) degree in Telecommunications Engineering from Federal Institute of Ceará (IFCE), Brazil. He is currently pursuing his doctorate in Computer Science from Universidade Federal do Ceará (UFC), Fortaleza, CE, Brazil.

### **SILVIA LOPES DE SENA TAGLIALENHA**

Professora da Universidade Federal de Santa Catarina, Centro Tecnológico de Joinville. Possui Licenciatura em Matemática, com habilitação em Física - UNESP - Faculdade de Ciências e Tecnologia - Júlio de Mesquita Filho, Mestrado em Matemática Pura (Análise) - ICMC/USP - Universidade de São Paulo, Doutorado em Engenharia Elétrica (Automação) - UNESP - Faculdade de Engenharia de Ilha Solteira. Atua principalmente nos seguintes temas: programação linear inteira, problemas de transporte e logística, meta heurísticas e planejamento da expansão de sistemas de transmissão de energia elétrica.

### **SIMONE MASSULINI ACOSTA**

Possui graduação em Engenharia Elétrica pela Universidade Federal de Santa Catarina (1992), mestrado em Engenharia Elétrica pela Universidade Federal de Santa Catarina (1997) e doutoranda do Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica do Paraná. Atualmente é professora da Universidade Tecnológica Federal do Paraná atuando principalmente nos temas: instrumentação industrial, automação de processos, modelagem e otimização de processos e engenharia da qualidade.

### **SOLANGE DA SILVA**

Graduada em Ciências - Habilitação em Matemática pela Universidade Católica de Goiás (1988), Especialista em Ciência da Computação pela Universidade Católica de Goiás (1995), Mestre em Engenharia Elétrica e de Computação pela Universidade Federal de Goiás (2000) e Doutora em Engenharia Elétrica pela Universidade Federal de Uberlândia (2005). Atualmente é professora adjunto da Pontifícia Universidade Católica de Goiás, na Escola de Ciências Exatas e de Computação. Além da graduação, também orienta dissertações no Programa de Mestrado em Engenharia de Produção e Sistemas da PUCGO e atualmente, pesquisa as relações interorganizacionais, especificamente as Redes de Cooperação Empresarial (RCEs) e Arranjos Produtivos Locais (APLs), em parceria com as universidades UNISINOS/RS e UNIP/SP. Tem experiência na área de Redes de Computadores, tecnologia da informação (TI), telecomunicações e tecnologia RFID.

**SOLANGE REGINA DOS SANTOS**

Possui graduação em Matemática pela Faculdade Estadual de Ciências e Letras de Campo Mourão/FECILCAM, atual Universidade Estadual do Paraná - UNESPAR/Campus de Campo Mourão (2004), mestrado em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná (2007) e doutorado em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná (2012). É professora da Colegiado de Matemática da Universidade Estadual do Paraná - UNESPAR/Campus de Campo Mourão, desde 2005. Realizou Pós-Doutorado pelo Programa de Pós-Graduação em Matemática - PPGM, Universidade Federal do Paraná no período de abril de 2015 a junho de 2016. Áreas de interesse: Otimização, Estatística e Confiabilidade.

**TAKASHI YONEYAMA**

Concluiu Engenharia Eletrônica pelo Instituto Tecnológico de Aeronáutica - ITA (1975), Medicina pela Universidade de Taubaté (1993)., Mestrado em Engenharia Eletrônica pelo ITA (1979) e Doutorado em Engenharia Elétrica pela Universidade de Londres (1983). Desde 1990 é Professor Titular do ITA. Foi Presidente (2004-2006) da Sociedade Brasileira de Automática. É co-autor de 4 livros. Possui mais de 99 artigos em periódicos e 350 em eventos científicos. Orientou, 31 doutores, 84 mestres e 67 bolsistas de iniciação científica, bem como 159 trabalhos de conclusão de curso.

**TÁSIA HICKMANN**

Possui graduação em Tecnologia em Manutenção Eletromecânica pela Universidade Tecnológica Federal do Paraná (UTFPR), graduação em Licenciatura em Matemática pela Universidade Estadual do Oeste do Paraná (UNIOESTE), mestrado em Matemática (Área de concentração: Álgebra) pela Universidade Estadual de Maringá (UEM), doutorado em Métodos Numéricos em Engenharia (Área de concentração: Programação Matemática) pela Universidade Federal do Paraná (UFPR) e Pós-Doutoramento em Engenharia Elétrica e Computação (UNIOESTE). Atualmente é docente do Magistério Superior na UTFPR - Campus Medianeira. Áreas de interesse: Análise Numérica; Modelagem Computacional; Cálculo Fracionário; Modelagem de Séries Temporais.

**TAYANE MAGALHÃES ALVAIA**

Graduanda em Engenharia de Produção pela Universidade Tiradentes (UNIT).

**THIAGO GOUVEIA DA SILVA**

Graduado e Mestre em Ciência da Computação pela Universidade Federal da Paraíba. Doutorando em Ciência da Computação pela Universidade Federal Fluminense, com ênfase em Grafos e Otimização Combinatória. Atualmente, é professor da Unidade Acadêmica de Informática do Instituto Federal da Paraíba.

**THIAGO MOURA DE MORAIS**

Graduado em Engenharia Civil pela PUC/GO (2015) e mestre em Engenharia de Produção e Sistemas pela PUC/GO (2018) com ênfase em climatologia, modelagem estatística de recursos hídricos e Geoestatística. É membro do grupo de pesquisa da PUC/GO que avalia eventos climáticos extremos na cidade de Goiânia/Goiás. Atuação na construção civil no planejamento e controle de obras oferecendo soluções para melhoria da produtividade. Também atua na infraestrutura urbana na elaboração e revisão de projetos de pavimentação e drenagem fluvial.

**THIAGO YUJI EGUCHI**

Graduado em Engenharia de Produção pelo CESUPA e fui membro de grupo de pesquisa de métodos de Tomada de Decisão, onde foram analisados algoritmos como: GUT, AHP, ELECTRE E PROMETHEE. Durante minhas experiências como estagiário na empresa Tramontina, adquiri domínios de ferramentas da qualidade: Diagrama de Ishikawa, Carta de Processos, PDCA. Além de conhecimentos em mapeamento de processos, controle de indicadores, planejamento de manutenções, implantação da Manutenção Produtiva Total (MPT), controle de gastos e elaboração de Procedimento Operacional Padrão.

**VANESSA ELIONARA SOUZA FERREIRA**

Mestre em Ambiente, Tecnologia e Sociedade (2016) pela Universidade Federal Rural do Semi-Árido. Graduada em Engenharia de Produção (2014) e Ciência e Tecnologia (2012) pela Universidade Federal Rural do Semi-Árido. Técnica em Informática pelo Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, campus Mossoró (2010). Atualmente é Assistente em Administração da Universidade Federal Rural do Semi-Árido.

**VANINA MACOWSKI DURSKI SILVA**

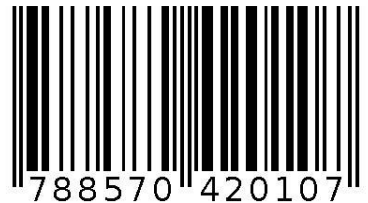
Engenharia de Produção Agroindustrial pela Faculdade Estadual de Ciências e Letras de Campo Mourão, especialista em Gestão de Agronegócios pela Fesurv. Mestre e doutora em Eng. de Produção na área de Logística e Transporte pela Universidade Federal de Santa Catarina. Atualmente é professora Adjunto III do Departamento de Engenharias da Mobilidade da Universidade Federal de Santa Catarina, sendo também Coordenadora do Grupo de Pesquisa em Logística Dinâmica e Autônoma voltado à logística autônoma e à dinamicidade dos sistemas logísticos. Atuou como engenheira de produção na área de Planejamento e Controle da Produção em empresas de grande porte como Perdigão Agroindustrial S.A., atual Brasilfoods e Macedo Agroindustrial S.A, atual Tyson Foods.

**WILLIAM DAITI KOBO TAKESHITA**

Graduando de Engenharia de Transportes e Logística pela Universidade Federal de Santa Catarina. Atuando na empresa Transmob Engenharia em projetos e planejamentos de transportes.

Agência Brasileira do ISBN

ISBN 978-85-7042-010-7



9 788570 420107