



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

REDUÇÃO DE DIMENSIONALIDADE ONLINE PARA PROBLEMA DE ROTEAMENTO DE VEÍCULOS COM TRANSPORTE REATIVO A DEMANDA

RENAN SANTOS MENDES

Orientadora: Prof^a. Dr^a. Elizabeth Fialho Wanner
CEFET-MG

Coorientadores: Prof. Dr. Flávio Vinícius Cruzeiro Martins
CEFET-MG

Prof. Dr. João Fernando Machry Sarubbi
CEFET-MG

BELO HORIZONTE
FEVEREIRO DE 2022

Mendes, Renan Santos
M538r Redução de dimensionalidade online para problema de roteamento de veículos com transporte reativo a demanda / Renan Santos Mendes. – 2022.
93 f.

Tese de doutorado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.

Orientadora: Elizabeth Fialho Wanner.

Coorientadores: Flávio Vinícius Cruzeiro Martins e João Fernando Machry Sarubbi.

Tese (doutorado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Levantamento de rotas – Teses. 2. Veículos – Frotas – Teses.
2. Transportes – Planejamento – Teses. 3. Otimização matemática – Teses.
4. Algoritmos – Teses. 5. Análise por agrupamento – Teses. I. Wanner, Elizabeth Fialho. II. Martins, Flávio Vinícius Cruzeiro. III. Sarubbi, João Fernando Machry. IV. Centro Federal de Educação Tecnológica de Minas Gerais. V. Título.

CDD 519.72

RENAN SANTOS MENDES

**REDUÇÃO DE DIMENSIONALIDADE ONLINE PARA
PROBLEMA DE ROTEAMENTO DE VEÍCULOS COM
TRANSPORTE REATIVO A DEMANDA**

Tese apresentada ao Programa de Pós-graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Doutor em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Sistemas Inteligentes

Orientadora: Prof^ª. Dr^ª. Elizabeth Fialho Wanner
CEFET-MG

Coorientadores: Prof. Dr. Flávio Vinícius Cruzeiro Martins
CEFET-MG

Prof. Dr. João Fernando Machry Sarubbi
CEFET-MG

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
FEVEREIRO DE 2022



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

REDUÇÃO DE DIMENSIONALIDADE ONLINE PARA PROBLEMA DE ROTEAMENTO DE VEÍCULOS COM TRANSPORTE REATIVO A DEMANDA.

Tese de Doutorado apresentada por **Renan Santos Mendes**, em 28 de março de 2022, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof.ª. Dr.ª. Elizabeth Fialho Wanner
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Flávio Vinícius Cruzeiro Martins (Coorientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. João Fernando Marchy Sarubbi (Coorientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Ivan Reinaldo Meneghini
O Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais

Prof.ª. Dr.ª. Carolina Gil Marcelino
Universidade Federal do Rio de Janeiro

Prof. Dr. Alisson Marques da Silva
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Gustavo Campos Menezes
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof.ª. Dr.ª. Elizabeth Fialho Wanner
Presidenta do Colegiado do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

Dedico este trabalho primeiramente a Deus,
a minha família, a minha esposa Gabi e aos
nossos filhos.

Agradecimentos

Olhando para tudo que se passou nestes mais de cinco anos, não há como ser comedido nesta parte do trabalho. Vou tomar a liberdade de usar este espaço sem me limitar a um número de parágrafos ou de páginas. Por isso, caro leitor, fique a vontade para ler ou para ir para próxima seção.

Agradeço primeiramente a Deus pelo dom da vida, por ser meu sustento em tantas horas de dificuldade e por nunca me abandonar em todas as situações. A presença e intercessão constante de Nossa Senhora, que sempre me ajudou e cobriu com seu manto sagrado.

Aos meus orientadores Beth e Flávio, pelos puxões de orelha, inúmeras reuniões e mudanças de horários e principalmente por todo apoio nessa fase. Somente vocês sabem como foi essa caminhada nesses últimos anos, todo cansaço, altos e baixos da vida. Em especial agradeço a Beth, por ser minha segunda mãe nessa jornada, que me chamou para fazer uma iniciação científica como voluntário em meados de 2009 enquanto era minha professora de Cálculo B lá no ICEB e depois me aceitou orientar no mestrado e finalizando agora a orientação no doutorado. Não tenho palavras para agradecer por toda força e apoio nesses anos todos (12 no total). Mesmo que nessa fase da minha carreira não esteja ligado diretamente ao ensino, tenho certeza que vou carregar pela vida todos ensinamentos sobre pesquisa, orientação e principalmente, como ser um excelente professor.

Agradeço à CAPES e ao CEFET-MG pela bolsa de estudos, a qual permitiu o desenvolvimento deste trabalho, assim como proporcionou minha formação. O presente trabalho foi realizado com apoio parcial da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

I would like to immensely thank Kalyanmoy Deb and Victoria Lush for all their partnership and contribution to the research, with ideas, knowledge and analysis. You were very important throughout the process.

Agradeço ao Dr Ismael Lopes que me acompanha há mais de trinta anos, cuidando da minha visão e ao Dr. Nassim Calixto (*in memoriam*) que, em meados de 1989, foram essenciais para mim. Para os que não sabem, as minhas duas graduações, mestrado, agora o doutorado, tudo mais pude ver nessa vida e se hoje consigo ler esse texto que escrevo de agradecimento foi por causa desses dois grandes médicos que diagnosticaram e me curaram de um glaucoma congênito o qual poderia ter me impedido de fazer o que fiz. Chego ao final de mais essa etapa com todo auxílio de vocês.

A toda Take Blip e especial ao time de Data & Analytics Research (antigo BI) que

me apoia desde 2018, nas pessoas da Paty, Tiagão, Fábio Lacerda e Arthur. Aos colegas de trabalho Gabriel Oliveira, Ramón, Milo e Rogers que sempre me escutaram, aguentaram o cansaço, deram todo apoio todas as vezes que precisei. Saibam que sem vocês, terminar essa fase da minha vida jamais seria possível.

Á todos os amigos que, por inúmeras vezes, não consegui dar atenção que precisavam, peço minhas sinceras desculpas pela ausência em diversos momentos na vida de cada um e agradeço por toda compreensão, apoio e oração de vocês. Aos amigos do Centro Cultural Itacolomy por todas as partilhas e direções na vida.

A minha família, aos meus pais e ao meu irmão por toda força nesses anos todos, pelo amor incondicional, por toda dedicação, por lutarem tanto por mim e por me compreenderem em todas as ausências. Vocês foram e sempre serão a base de tudo.

Agora sim, um agradecimento de todo tamanho desse mundo para a minha companheira de vida, amada esposa Gabi que tanto me segurou nas mãos nesses últimos anos, me acompanha e me sustenta desde 2011. Obrigado por ser meu sustento em tantos momentos difíceis e por não me deixar desistir todas as vezes que o cansaço batia na porta. Ao nosso pequeno José (*in memoriam*) que já nos espera junto do Pai, um dia estaremos juntos na eternidade e ao Antônio que em breve chegará ao mundo.

“Se você parar de aprender, logo vai começar a esquecer as coisas que já sabe” (Provérbios 19:27)

Resumo

Neste trabalho é abordada uma formulação com muitos objetivos para o Problema de Roteamento de Veículos com Transporte Reativo a Demanda (PRVTRD). O problema pode ser considerado como um modo de transporte aproximado ao dos serviços de transporte sob demanda com compartilhamento de corridas. Este trabalho propõe o uso da análise de *cluster online* baseada nos coeficientes de correlação de Pearson e τ de Kendall para realizar a redução de dimensionalidade a cada geração do algoritmo MOEA/D. Inicialmente, foram comparadas as seguintes abordagens: (i) *offline* usando o coeficiente de correlação de Pearson; (ii) *offline* usando τ de Kendall; (iii) *online* usando o coeficiente de correlação de Pearson; (iv) *online* usando τ de Kendall; e (v) uma versão de baseline, o MOEA/D. Os algoritmos foram testados usando um conjunto de dados para a cidade de Belo Horizonte. Para avaliar a dispersão das soluções obtidas pelos algoritmos, foi aplicado o indicador de hipervolume. Os resultados destes experimentos mostraram que (i) não há diferença na formulação obtida para as abordagens *offline*; (ii) os algoritmos *online* são estatisticamente melhores do que o algoritmo *offline*; (iii) não é possível afirmar que há diferença estatística entre os algoritmos *online* e o MOEA/D. Os diagramas de cordas foram aplicados sobre as soluções obtidas e indicaram que a diversidade das soluções é diferente. A segunda sequência de experimentos realizados compara as abordagens de redução de dimensionalidade baseada em agregação *online* e outra baseada em seleção de atributos, que usa duas técnicas: Taxa de Dispersão (DR) e Máxima Variância (MV), as quais são usadas para selecionar o representante do *cluster*. O impacto da frequência da redução da dimensionalidade no desempenho do algoritmo também foi analisado. As abordagens foram acopladas ao algoritmo MOEA/D. Três algoritmos foram testados (i) *cluster online* usando o coeficiente de correlação Pearson com MOEA/D (OnCL ρ_g -MOEA/D); (ii) *cluster online* usando o coeficiente de correlação de Pearson e DR com MOEA/D (OnDR ρ_g -MOEA/D); (iii) *cluster online* usando coeficiente de Pearson e MV com MOEA/D (OnMV ρ_g -MOEA/D), sendo $g = 1, 25, 50, 100$. Os resultados mostraram que, independentemente da frequência da abordagem de redução de dimensionalidade, os algoritmos MOEA/D e OnCL ρ_g -MOEA/D são estatisticamente superiores aos OnDR ρ_g -MOEA/D e OnMV ρ_g -MOEA/D. Não é possível afirmar que haja diferença estatística entre os resultados dos algoritmos baseados na seleção de atributos *online*. Os diagramas de cordas mostraram que há uma maior diversidade das soluções obtidas quando todas as funções objetivo são utilizadas na forma reduzida do problema e também quando os algoritmos permanecem por mais gerações em uma determinada formulação reduzida.

Palavras-chave: Problema de Roteamento de Veículos com Transporte Reativo a Demanda. Otimização com Muitos Objetivos. Redução de Dimensionalidade. Análise de *Cluster*. Seleção de Atributos.

Abstract

In this work, a many-objective formulation for the Vehicle Routing Problem with Demand Responsive Transport (VRPDRT) is addressed. The problem can be considered as a mode of transport similar to that of ride-sharing on-demand transport services. This work proposes the use of online cluster analysis based on Pearson's and Kendall's τ correlation coefficients to perform dimensionality reduction at each generation of the MOEA/D algorithm. Initially, the following approaches were compared: (i) offline using Pearson's correlation coefficient; (ii) offline using τ of Kendall; (iii) online using Pearson's correlation coefficient; (iv) online using τ of Kendall; and (v) a baseline version, the MOEA/D. The algorithms were tested using a dataset for the city of Belo Horizonte. To evaluate the dispersion of the solutions obtained by the algorithms, the hypervolume indicator was applied. The results of these experiments showed that (i) there is no difference in the formulation obtained for the offline approaches; (ii) the online algorithms are statistically better than the offline algorithm; (iii) it is not possible to state that there is a statistical difference between the online algorithms and the MOEA/D. The chord diagram was applied to the solutions obtained and indicated that the diversity of solutions is different. The second sequence of experiments performed compares the dimensionality reduction approaches based on online aggregation and another one based on feature selection, which uses two techniques: Dispersion Ratio (DR) and Maximum Variance (MV), which are used to select the cluster representative. The impact of the frequency of dimensionality reduction on algorithm performance was also analyzed. The approaches were coupled to the MOEA/D algorithm. Three algorithms were tested (i) online cluster using the Pearson correlation coefficient with MOEA/D (OnCL ρ_g -MOEA/D); (ii) online cluster using Pearson's correlation coefficient and DR with MOEA/D (OnDR ρ_g -MOEA/D); (iii) online cluster using Pearson's coefficient and MV with MOEA/D (OnMV ρ_g -MOEA/D), where $g = 1, 25, 50, 100$. The results showed that, regardless of the frequency of the dimensionality reduction approach, the MOEA/D and OnCL ρ_g -MOEA/D algorithms are statistically superior to OnDR ρ_g -MOEA/D and OnMV ρ_g -MOEA/D. It is not possible to state that there is a statistical difference between the results of algorithms based on online attribute selection. The chord diagram showed that there is a greater diversity of solutions obtained when all objective functions are used in the reduced form of the problem and also when the algorithms remain for more generations in a given reduced formulation.

Keywords: Vehicle Routing Problem with Demand Responsive Transport. Many Objective Optimization. Dimensionality Reduction. Cluster Analysis. Feature Selection.

Lista de Figuras

Figura 1 – Estado inicial e o deslocamento de solicitações para o problema. Os mapas são feitos usando o Google Maps [®]	10
Figura 2 – Rotas para atender as solicitações. Considera-se que cada passageiro solicitou um táxi do depósito para transportá-lo entre a origem e o destino. Os mapas são feitos usando Google Maps [®]	11
Figura 3 – Duas soluções TRD diferentes para atender às solicitações. Os mapas são feitos usando Google Maps [®]	12
Figura 4 – Representação da solução, baseada na Figure 3b.	12
Figura 5 – Rota R_2 é concatenada após a rota R_1	13
Figura 6 – Rotas concatenadas.	13
Figura 7 – Mapeamento do espaço de parâmetros no espaço de objetivos. Adaptado de Deb (2008).	21
Figura 8 – Exemplo do conceito de Dominância. Adaptado de Deb (2008).	22
Figura 9 – Aumento das regiões não dominadas. Adaptado de Farina e Amato (2002).	24
Figura 10 – Exemplo de soluções resistentes à dominância para o problema DTLZ2. Resultados obtidos usando <i>MOEA Framework</i> Hadka (2017), versão 2.12. Adaptado de Jaimes (2011).	25
Figura 11 – Exemplos de aplicação do algoritmo NSGA-II para o problema DTLZ2. Resultados obtidos usando <i>MOEA Framework</i> Hadka (2017), versão 2.12.	27
Figura 12 – Gráfico de bolhas apresentando visualização para soluções em \mathbb{R}^5	28
Figura 13 – Exemplos de ferramentas para visualização de soluções em \mathbb{R}^{12}	29
Figura 14 – Exemplo de conflito e suporte entre funções. A funções $f_1(x)$ suporta $f_2(x)$ no intervalo de x_1 até x_2 . funções $f_3(x)$ e $f_4(x)$ apresentam conflito no intervalo de x_3 até x_4 . Adaptado de Jaimes (2011).	33
Figura 15 – Dendrograma para um conjunto de dados com nove pontos distintos usado o método de ligação única.	38
Figura 16 – Dendrograma do <i>cluster</i> hierárquico usando coeficiente de correlação de Pearson em 2.500 soluções aleatórias.	43
Figura 17 – Operador de cruzamento de dois pontos.	47
Figura 18 – Operador de mutação <i>2-Shuffle</i>	48
Figura 19 – Dendrogramas para as 12.500 soluções aleatórias usando correlação de Pearson e τ de Kendall como medidas de similaridade entre as funções objetivo. As figuras foram feitas utilizando o pacote <i>hclust</i> da linguagem R.	51
Figura 20 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos a partir das 30 execuções dos algoritmos OffCL-MOEA/D, OnCL $_{\tau}$ -MOEA/D, OnCL $_{\rho}$ -MOEA/D e MOEA/D usando 50, 100, 150, 200 e 250 requisições.	53

Figura 21 – Os diagramas de corda mostram a dispersão do conjunto Pareto combinado para as instâncias com 50 e 100 solicitações.	57
Figura 22 – Os diagramas de corda mostram a dispersão do conjunto Pareto combinado para as instâncias com 150, 200 e 250 solicitações.	58
Figura 23 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos a partir das 30 execuções dos algoritmos MOEA/D, OnCL ρ_g -MOEA/D e OnMV ρ_g -MOEA/D com $g = 1, 100$ para as instâncias com 50, 100, 150, 200 e 250 requisições.	62
Figura 24 – Instância com 50 solicitações.	64
Figura 25 – Instância com 100 solicitações.	65
Figura 26 – Instância com 150 solicitações.	66
Figura 27 – Instância com 200 solicitações.	67
Figura 28 – Instância com 250 solicitações.	68
Figura 29 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D e MOEA/D usando 50 solicitações com $g = 1, 25, 50, e 100$	74
Figura 30 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D e MOEA/D usando 100 solicitações com $g = 1, 25, 50, e 100$	74
Figura 31 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 150 solicitações com $g = 1, 25, 50, e 100$	75
Figura 32 – <i>Boxplots</i> dos valores do indicador de hipervolume obtido amos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 200 solicitações com $g = 1, 25, 50, e 100$	75
Figura 33 – <i>Boxplots</i> dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 250 solicitações com $g = 1, 25, 50, e 100$	76
Figura 34 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 50 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50 e 100$).	79
Figura 35 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 100 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50 e 100$).	80

Figura 36 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 150 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).	81
Figura 37 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 200 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).	82
Figura 38 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 250 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).	83

Lista de Tabelas

Tabela 1 – Para ilustrar uma pequena instância de PRVTRD, quatro solicitações e seus respectivos pontos de embarque e desembarque, janela de tempo de embarque e janela de tempo de desembarque são fornecidos. PL, PU, DL e DU representam, respectivamente, os limites inferiores e superiores para as janelas de tempo de embarque e desembarque.	10
Tabela 2 – Relações entre o limite superior do número de soluções e dimensão do problema.	26
Tabela 4 – p -valores para os testes de aleatoriedade para todos os pares de algoritmos do grupo (a) usando instâncias com 50, 100, 150, 200 e 250 requisições. O nível de significância é definido como 5%.	77
Tabela 5 – p -valores para os testes de aleatoriedade para todos os pares de algoritmos do grupo (b) usando instâncias com 50, 100, 150, 200 e 250 requisições. O nível de significância é definido como 5%.	78

Lista de Algoritmos

Algoritmo 1 – Construção de uma solução gulosa para o PRVTRD	14
Algoritmo 2 – <i>Cluster</i> hierárquico aglomerativo	37
Algoritmo 3 – On-MOEA/D	45

Lista de Abreviaturas e Siglas

CRL	<i>Time-window Rank List</i>
DR	<i>Dispersion Ratio</i>
DRL	<i>Delivery time Rank List</i>
GRASP	<i>Greedy Randomized Adaptive Search Procedure</i>
HI	<i>Hypervolume Indicator</i>
IGD	<i>Inverted Generational Distance</i>
IGD ⁺	<i>Inverted Generational Distance Plus</i>
MOEA/D	<i>Multiobjective Evolutionary Algorithm Based on Decomposition</i>
MV	<i>Maximum Variance</i>
NRF	<i>Node-Ranking Function</i>
NRL	<i>Number of passengers Rank List</i>
NSGA-II	<i>Non-Dominated Sorting Genetic Algorithm II</i>
NSGA-III	<i>Non-Dominated Sorting Genetic Algorithm III</i>
OffCL _{ρ} -MOEA/D	<i>Offline cluster MOEA/D com correlação de Pearson</i>
OffCL _{τ} -MOEA/D	<i>Offline cluster MOEA/D com τ de Kendall</i>
OnCL _{ρ} -MOEA/D	<i>Online cluster MOEA/D com correlação de Pearson</i>
OnCL _{τ} -MOEA/D	<i>Online cluster MOEA/D com τ de Kendall</i>
OnDR _{ρ} -MOEA/D	<i>Online cluster MOEA/D com taxa de dispersão</i>
OnMV _{ρ} -MOEA/D	<i>Online cluster MOEA/D com máxima variância</i>
PRVTRD	<i>Problema de Roteamento de Veículos com Transporte Reativo a Demanda</i>
SPEA 2	<i>Strength Pareto Evolutionary Algorithm 2</i>
TRD	<i>Transporte Reativo a Demanda</i>
TRL	<i>Delivery time Rank List</i>

Lista de Símbolos

\uparrow	Suporte entre funções
\downarrow	Conflito entre funções
α_1	Peso do fator distância
α_2	Peso do número de passageiros
α_3	Peso do fator janela de tempo de desembarque
α_4	Peso do fator janela de tempo de embarque
ρ_{ij}	Coefficiente de correlação de Pearson, entre a i -ésima e j -ésima variável do vetor X
$\tau(\mathcal{P}_1, \mathcal{P}_2)$	Coefficiente de correlação de Kendall entre conjuntos de pares ordenados \mathcal{P}_1 e \mathcal{P}_2
$d_{i,j}$	Distância associada ao arco $a_{i,j}$
$d(X_l, X_k)$	Distância entre os vetores coluna X_l e X_k do vetor X
\preceq	Relação de dominância
\prec	Relação de dominância estrita
$\mathbf{f}(\cdot)$	Funções vetoriais que mapeiam pontos de \mathbb{R}^n em \mathbb{R}^m
$\mathbf{g}(\cdot)$	Restrições de desigualdade que mapeiam pontos de \mathbb{R}^n em \mathbb{R}^r
$\mathbf{h}(\cdot)$	Restrições de igualdade que mapeiam pontos de \mathbb{R}^n em \mathbb{R}^p
$q_{k,i}$	Taxa de ocupação do veículo k quando deixa o ponto de parada i
m_v	Índice de capacidade do ponto de parada v
p^+	Vértice de embarque
p^-	Vértice de desembarque
$t_{i,j}$	Tempo de viagem associado ao arco $a_{i,j}$
DL_l	Limite inferior da janela de tempo de desembarque da solicitação l
DT_l^k	Momento em que um passageiro, com sua respectiva solicitação l , desembarca do veículo k

DU_l	Limite superior da janela de tempo de desembarque da solicitação l
K	Conjunto de veículos
P_{in_v}	Conjunto com as solicitações de embarque no ponto de parada v
P_{out_v}	Conjunto com as solicitações de desembarque no ponto de parada v
$P_{p \times p}$	Matriz de correlação de dimensão $p \times p$
PL_l	Limite inferior da janela de tempo de embarque da solicitação l
PT_l^k	Momento em que um passageiro, com sua respectiva solicitação l , embarca no veículo k
PU_l	Limite superior da janela de tempo de embarque da solicitação
R_{MAX}	Rota com a maior duração
R_{MIN}	Rota com a duração mais curta
R	Conjunto com as solicitações de transporte dos passageiros
S	Solução para o PRVTRD
V	Conjunto de pontos de parada
U	Conjunto de solicitações não atendidas
X	Vetor de variáveis aleatórias

Sumário

1 – Introdução	1
1.1 Motivação	5
1.2 Objetivos	5
1.3 Contribuições desta Tese	6
1.4 Publicações Realizadas Durante o Desenvolvimento desta Tese	6
1.5 Organização do Trabalho	7
2 – O Problema de Roteamento de Veículos com Transporte Reativo a Demanda	8
2.1 Descrição do Problema	8
2.2 Representação da Solução do Problema	12
2.3 Algoritmo de Construção Gulosa	13
2.4 Formulação com Muitos Objetivos	15
2.5 Revisão da Literatura para o PRVTRD	16
3 – Otimização Multiobjetivo	20
3.1 Principais Conceitos	20
3.1.1 Problema de Otimização Multiobjetivo	20
3.2 Otimização com Muitos Objetivos	22
3.2.1 Principal Diferença entre Otimização Multiobjetivo e com Muitos Objetivos	23
3.3 Principais Dificuldades para Resolver os Problemas com Muitos Objetivos	23
3.3.1 Deterioração da Habilidade de Busca	23
3.3.2 Soluções Resistentes à Dominância	24
3.3.3 Aumento no Número de Soluções Não Dominadas	25
3.3.4 Dificuldade de Visualização das Soluções	27
3.3.5 Custo Computacional dos Indicadores de Performance	29
3.4 Principais Classes de Algoritmos	30
4 – Algoritmo de Redução de Dimensionalidade <i>Online</i>	32
4.1 Conflito entre Funções	32
4.2 Análise de <i>Cluster</i>	33
4.2.1 Notação Utilizada	33
4.2.2 Principais Conceitos	33
4.2.3 Medidas de Similaridade	34
4.2.3.1 Distância de Manhattan	34
4.2.3.2 Distância Euclidiana	34

4.2.3.3	Distância de Minkowsky	34
4.2.4	Coeficiente de Correlação	34
4.2.4.1	Coeficiente de Correlação de Pearson	35
4.2.4.2	Coeficiente de Correlação de Kendall	36
4.2.5	Classificação de um <i>Cluster</i>	36
4.2.6	Construção de um <i>Cluster</i>	37
4.2.7	Métodos para Calcular a Distância entre <i>Clusters</i>	37
4.2.8	Dendrograma	38
4.2.9	Interpretação de um Dendrograma	38
4.3	Redução de Dimensionalidade Usando <i>Clustering</i>	39
4.4	Técnicas de Seleção de Atributos	41
4.4.1	Taxa de Dispersão	41
4.4.2	Máxima Variância	42
4.4.3	Exemplos de Aplicação das Técnicas de Seleção de Atributos para Escolha de Objetivos	42
4.5	Redução de Dimensionalidade <i>Online</i>	44
4.5.1	Detalhes do Algoritmo de Redução de Dimensionalidade <i>Online</i> Proposto	45
4.5.2	Geração dos Pesos para os Objetivos	46
4.6	Operadores Genéticos	46
4.6.1	Operador de Seleção	46
4.6.2	Operador de Cruzamento	47
4.6.3	Operador de Mutação	47
5	Resultados Computacionais	49
5.1	Geração de Instâncias	49
5.2	Configuração dos Parâmetros dos Algoritmos	50
5.3	Redução de Dimensionalidade <i>Online</i> com Análise de <i>Cluster</i>	51
5.4	Redução de Dimensionalidade <i>Online</i> Usando Seleção de Atributos	60
6	Conclusões e Trabalhos Futuros	69
	Apêndices	72
	APÊNDICE A – <i>Boxplots</i> e <i>p</i>-valores dos Testes de Aleatoriedade	73
A.1	<i>Boxplots</i>	73
A.1.1	<i>Boxplots</i> para Instância com 50 Solicitações	73
A.1.2	<i>Boxplots</i> para Instância com 100 Solicitações	74
A.1.3	<i>Boxplots</i> para Instância com 150 Solicitações	74
A.1.4	<i>Boxplots</i> para Instância com 200 Solicitações	75

A.1.5	<i>Boxplots</i> para Instância com 250 Solicitações	75
A.2	Tabelas com Testes de Aleatoriedade	77
APÊNDICE B	Diagramas de Corda	79
B.1	Diagramas de Corda para Resultados da Instância com 50 solicitações . . .	79
B.2	Diagramas de Corda para Resultados da Instância com 100 solicitações . .	80
B.3	Diagramas de Corda para Resultados da Instância com 150 solicitações . .	81
B.4	Diagramas de Corda para Resultados da Instância com 200 solicitações . .	82
B.5	Diagramas de Corda para Resultados da Instância com 250 solicitações . .	83
Referências	84

1 Introdução

O termo Transporte Reativo a Demanda (TRD) tem sido cada vez mais utilizado nos últimos anos para se referenciar uma parcela do mercado que substitui (normalmente por meio de ônibus de pequena capacidade ou de táxis) o transporte convencional onde a demanda é baixa e muitas vezes espalhada por uma grande área [Mulley e Nelson \(2009\)](#). Para ser considerado um TRD o serviço deve possuir as seguintes características: estar disponível ao público em geral sem estar restrito a grupos específicos; ser fornecido por veículos de pequena capacidade; e possuir rotas com horários flexíveis, sendo alterados sob demanda [Zografos, Androutsopoulos e Sihvola \(2008\)](#). O TRD pode ser visto como uma solução alternativa aos sistemas de transporte mais tradicionais [Mulley e Nelson \(2009\)](#), [Zografos, Androutsopoulos e Sihvola \(2008\)](#), [Velaga et al. \(2012\)](#) e tem sido argumentado que, se bem implementado, pode fornecer um sistema de transporte mais econômico.

Esse trabalho tem como objetivo estudar um problema denominado Problema de Roteamento de Veículos com o Transporte Reativo a Demanda (PRVTRD). O problema apareceu pela primeira vez em [Gomes, Souza e Dias \(2010\)](#) e apresenta características encontradas nos problemas clássicos de roteamento, serviços de compartilhamento de carros e gerenciamento de sistemas de transporte reativos a demanda. Nesse tipo de sistema, os veículos devem seguir rotas e horários com o objetivo de atender aos pedidos e levar o maior número possível de usuários no mesmo veículo, mantendo a qualidade do serviço o mais alta possível. A qualidade é medida em termos de tempos de embarque/desembarque e duração da viagem. O transporte de passageiros é feito sob demanda, tentando atender às necessidades dos passageiros por meio de rotas e horários flexíveis.

O PRVTRD pode ser considerado como uma extensão da abordagem de serviço para mobilidade urbana oferecida por aplicativos como Uber, Bridj e GrabCoach [Uber \(2009\)](#), [Bridj \(2014\)](#), [Grabcoach \(2017\)](#). Um exemplo já conhecido é o UberPool, serviço de *car-pooling* que a Uber oferece nas grandes cidades e combina os clientes individuais com os passageiros que seguem na mesma direção, compartilhando não só a viagem, mas também o custo. Neste trabalho, o problema levará em conta não apenas as características de compartilhamento e demanda, mas também uma janela de tempo de chegada de passageiros pré-determinada. O tipo de transporte reativo à demanda tem excelente potencial de crescimento, desde o desenvolvimento conjunto de uma rede de veículos autônomos sob demanda que integra os fabricantes de automóveis e os serviços bem-sucedidos de compartilhamento de viagens. Alguns exemplos de tais associações são a General Motors e a Lyft, a Toyota com a Uber e a Volkswagen que lançou uma divisão interna para o serviço de transporte sob demanda chamado MOIA no final de 2016 [Solomon \(2017\)](#), [Moia \(2017\)](#).

A ideia principal do transporte reativo a demanda é gerar um meio de transporte intermediário entre o transporte privado e o público, para que vários passageiros possam compartilhar o mesmo veículo, levando em conta as necessidades de embarque e desembarque. Assim, será possível oferecer um serviço de mobilidade urbana com flexibilidade sobre as rotas e horários quando comparado ao transporte público como os sistemas de TRD [Mageean e Nelson \(2003\)](#).

Neste trabalho, os passageiros especificam os pontos de embarque/desembarque e uma janela de tempo para ambos. Os pontos de parada dos veículos são pré-definidos de acordo com fluxo de passageiros dentro da cidade de Belo Horizonte. Os dados de deslocamento (tempo e distância) foram coletados pela API do Google Maps que usa uma série histórica para os valores. A frota de veículos considerada é homogênea, ou seja, todos os veículos possuem a mesma capacidade de cinco pessoas: um motorista e quatro passageiros. As janelas de tempo embarque dos passageiros devem ser respeitadas e as janelas de desembarque podem ser violadas mas com alguma forma de penalização. Os veículos devem começar e terminar as rotas no depósito.

Segundo diversos trabalhos encontrados [Zitzler e Thiele \(1999a\)](#), [Zitzler, Deb e Thiele \(2000\)](#), [Deb \(2008\)](#), problemas de otimização que buscam resolver mais de uma função objetivo ao mesmo tempo, são denominados de problemas de otimização multiobjetivo. Uma segunda denominação, mais específica, também pode ser encontrada, de tal forma que problemas de otimização que possuem uma quantidade maior ou igual a quatro funções objetivo são chamados de problemas com muitos objetivos, tradução literal de *Many Objective Optimization Problem* (MaOOP) [Ishibuchi et al. \(2009\)](#), [Duro et al. \(2014\)](#), [Lucken, Barán e Brizuela \(2014\)](#)

Problemas do mundo real são multiobjetivos por natureza e sua otimização pode ser vista em diversas áreas [Safi, Ucan e Bayat \(2018\)](#) como por exemplo: engenharia aeronáutica e aeroespacial [Jet \(2003\)](#), [Che e Tang \(2008\)](#), [Chiba, Makino e Takatoya \(2008\)](#), [Li et al. \(2018\)](#), distribuição de energia elétrica, água e gás [Kou et al. \(2017\)](#), [Marques, Cunha e Savić \(2018\)](#), [Sekizaki, Nishizaki e Hayashida \(2018\)](#), robótica [Liu et al. \(2019\)](#), [Fang et al. \(2019\)](#), indústria [Raja, Jhala e Patel \(2017\)](#), [Cao et al. \(2020\)](#). Muitas dessas aplicações apresentam objetivos conflitantes e podem conter de 6 [Fu et al. \(2013\)](#) a 14 objetivos [Chikumbo, Goodman e Deb \(2012\)](#).

A solução de problemas multiobjetivo (como no caso das aplicações anteriores) ou com muitos objetivos usando-se os algoritmos tradicionais, baseados em Pareto dominância, não é trivial [Safi, Ucan e Bayat \(2018\)](#) e a razão é que os algoritmos apresentam dificuldades de trabalhar em um espaço de objetivos muito grande, podendo apresentar diversos problemas como: alto custo computacional para se avaliar as métricas de performance, número de soluções eficientes e visualização das soluções [Duro et al. \(2014\)](#), [Li et al. \(2015\)](#), [Coello Coello et al. \(2020\)](#).

O Problema de Roteamento de Veículos com Transporte Reativo a Demanda, abordado neste trabalho, utiliza oito funções objetivo. Foram realizadas algumas adaptações das funções objetivo utilizadas em trabalhos anteriores, [Mendes et al. \(2016\)](#), [Mendes et al. \(2017a\)](#). A função que mede a quantidade de requisições não atendidas não é utilizada mais como um objetivo e foi transformada em uma restrição que avalia a qualidade de serviço do transporte. Dessa forma, oito funções objetivo diferentes foram consideradas de modo que três diferentes perspectivas do problema pudessem ser avaliadas: passageiro, empresa e motorista. Com base nas definições citadas anteriormente, pode-se dizer que o problema em questão é considerado um problema com muitos objetivos (MaOOP).

Como forma de contornar as dificuldades encontradas nos problemas com muitos objetivos, algumas abordagens foram feitas durante toda a pesquisa realizada, com o foco em reduzir a dimensionalidade do problema para transformá-lo em uma formulação bi-objetivo, sendo esta mais simples de ser resolvida pelos algoritmos. Esta abordagem é uma extensão de trabalhos anteriores [Mendes et al. \(2016\)](#), [Mendes et al. \(2016\)](#), [Mendes et al. \(2017a\)](#), em que, para resolver as formulações com muitos objetivos do PRVTRD, algumas técnicas, Árvore de Agregação [Freitas, Fleming e Guimarães \(2015\)](#) e Análise de *Cluster* [Mingoti \(2005\)](#), foram utilizadas para se realizar a redução da dimensionalidade do problema antes de se aplicar o processo evolutivo no algoritmo NSGA-II [Deb et al. \(2002\)](#). O termo agregação, utilizado neste texto, refere-se à soma ponderada de objetivos. No entanto, nesta tese, diversos avanços foram realizados adicionalmente em relação aos trabalhos já desenvolvidos, os quais devem ser destacados:

- A análise de *cluster* também é aplicada, mas a tarefa de redução de dimensionalidade é realizada a cada geração de um Algoritmo Evolucionário Multiobjetivo (MOEA). A redução da dimensionalidade *online* proposta transforma o problema de muitos objetivos em um bi-objetivo usando um procedimento de agregação;
- Para realizar a análise de *cluster*, uma medida de correlação não paramétrica, τ de Kendall, também é usada para obter a similaridade entre os *clusters*. Em [Mendes et al. \(2017a\)](#), apenas a correlação de Pearson é usada, que mede a força da relação linear entre variáveis normalmente distribuídas;
- O conjunto de funções objetivo é diferente do apresentado em [Mendes et al. \(2016\)](#), [Mendes et al. \(2016\)](#), [Mendes et al. \(2017a\)](#). Duas novas funções objetivo para o problema são adicionadas à formulação apresentada em [Mendes et al. \(2017a\)](#). Como uma janela de tempo para a entrega foi incluída na formulação do problema, a primeira função objetivo adicionada representa a janela de tempo de entrega total antecipada. A segunda função objetivo agregada visa melhorar a ocupação do veículo levando em consideração o número médio de assentos livres nos veículos. Além disso, a função objetivo apresentada em [Mendes et al. \(2016\)](#), [Mendes et al. \(2016\)](#), [Mendes et al. \(2017a\)](#) que representa o número de solicitações não atendidas agora está incluída na restrição do problema definida como uma meta de qualidade de serviço;

- Além das reduções de dimensionalidade usando a agregação dos objetivos, esse trabalho também utiliza duas técnicas não supervisionadas de seleção de atributos para se escolher um objetivo que seja mais significativo dentro de cada *cluster* encontrado. A escolha da função também é realizada dentro das gerações dos algoritmos evolutivos;
- As abordagens propostas são acopladas ao Algoritmo Evolutivo Multiobjetivo baseado em Decomposição, do inglês *Multiobjective Evolutionary Algorithm Based on Decomposition* (MOEA/D) Zhang e Li (2007). Nos trabalhos anteriores, o algoritmo de *baseline* foi o NSGA-II Deb et al. (2002). A escolha do MOEA/D foi com o propósito de aplicar um algoritmo para lidar melhor com um aumento no número de funções objetivo fornecendo uma comparação de algoritmo significativa no espaço objetivo original.

Um ponto importante que vale ser ressaltado é o fato de que o algoritmo apresentado nesta tese e aplicado ao PRVTRD não se limita apenas ao problema de roteamentos em questão. Pode-se considerar como uma metodologia que de um modo geral, poderia ser aplicado a qualquer problema com muitos objetivos.

Os testes para avaliar as abordagens propostas foram feitos em duas partes: na primeira, foram feitos os experimentos com as duas medidas de similaridade e os resultados foram comparados com o *baseline* MOEA/D, sendo que a redução de dimensionalidade com agregação das funções objetivo foi realizada em todas as gerações. Na segunda parte, os experimentos testaram a eliminação das funções objetivo como também a frequência em que as reduções de dimensionalidade ocorriam. Em ambas as partes, as soluções obtidas pelos algoritmos foram avaliadas qualitativamente com base dos diagramas de corda Koochaksaraei et al. (2017).

Após as execuções de todos os experimentos, os resultados mostraram que:

- A abordagem *online*, proposta neste trabalho, atinge melhores resultados em todas as instâncias testadas quando comparada com a abordagem *offline*;
- Não é possível afirmar que haja diferença estatística entre os algoritmos baseados em agregação *online* reduzindo o problema de 8 para 2 objetivos e o MOEA/D resolvendo o problema em 8 objetivos;
- Não há diferença no conjunto Pareto combinado resultante quando uma medida de similaridade paramétrica é comparada com uma não paramétrica em todas as instâncias testadas em relação ao indicador de hipervolume;
- Quando comparamos os resultados no espaço objetivo original usando o indicador de hipervolume, não há diferença no desempenho quando a abordagem *online* é comparada com o algoritmo MOEA/D;
- As abordagens *online* e o algoritmo *baseline* são estatisticamente superiores às abordagens baseadas em seleção de atributos;

- Não é possível afirmar que haja diferença estatística entre os resultados dos algoritmos baseados na seleção de atributos *online* em termos do indicador de hipervolume.

1.1 Motivação

Pode ser visto que em alguns trabalhos encontrados na literatura que usam a redução de dimensionalidade *online* buscam, de uma forma geral, encontrar algum conjunto mínimo de funções objetivo para representar o problema Lopez, Coello e Chakraborty (2008), Jaimes, Coello e Chakraborty (2008), Jaimes, Coello e Barrientos (2009), Breaban e Iftene (2015). A eliminação das funções objetivo pode gerar, de alguma forma, a perda de informação no processo evolutivo, visto que a relações entre esses objetivos pode ser perdida com sua eliminação. Outro detalhe observado é que, estes trabalhos sempre utilizam problemas teste para se avaliar os métodos e algoritmos propostos. Há poucos trabalhos que utilizam alguma técnica para reduzir a dimensionalidade dos problemas de natureza combinatória, seja por agregação ou eliminação das funções objetivo.

Uma das grandes dificuldades ao se resolver os problemas com muitos objetivos é a questão relacionada à escalabilidade dos algoritmos, que perdem seu desempenho à medida que o número de funções objetivo cresce Li et al. (2015). Segundo Coello et al. (2019), esse ponto ainda é uma das questões abertas sobre problemas com muitos objetivos e com potencial para ser explorado. Nesse sentido, os focos deste trabalho são: (i) resolver o PRVTRD na forma original para se obter um *baseline* de comparação, (ii) aplicar técnicas de redução de dimensionalidade para descobrir uma forma mais simples de se resolver o problema, gastando menos recursos computacionais, obtendo soluções tão boas quando o *baseline* e (iii) utilizar uma técnica de visualização para se comparar a qualidade das soluções em termos de convergência e dispersão.

1.2 Objetivos

O objetivo geral deste trabalho é acoplar um método de redução de dimensionalidade *online* dentro do ciclo evolutivo de um algoritmo evolutivo multiobjetivo para resolver o problema PRVTRD com muitos objetivos. Os objetivos específicos deste trabalho são:

- Reformular as funções objetivo utilizadas no PRVTRD encontradas em trabalhos anteriores;
- Fazer um levantamento dos principais trabalhos encontrados que utilizam a redução da dimensionalidade em problemas de otimização com muitos objetivos;
- Gerar instâncias com dados de distância e tempos de deslocamento reais para a cidade de Belo Horizonte;
- Resolver o PRVTRD através da abordagem de redução de dimensionalidade *online*;

- Utilizar dois tipos de técnicas de redução de dimensionalidade, uma baseada em agregação e outra baseada na eliminação de objetivos;
- Usar uma técnica de visualização para analisar qualitativamente as soluções;
- Apresentar e analisar os resultados obtidos.

1.3 Contribuições desta Tese

Diversas contribuições podem ser destacadas, dentre elas, as mais significativas são:

- Aplicação das técnicas de redução de dimensionalidade de forma *online* em problema combinatório explorando melhor o espaço de busca;
- Mostrar que a remoção de funções objetivo não conflitantes pode ocasionar perda significativa de informações do problema a ser resolvido;
- Resolver um problema de dimensão elevada de maneira mais simples e menos complexa, obtendo-se resultados sem diferença estatística.

1.4 Publicações Realizadas Durante o Desenvolvimento desta Tese

Durante o desenvolvimento desta tese, foi possível realizar diversas parcerias e ter contato com grupos de pesquisa na Universidade Estadual de Michigan (MSU) e na Universidade de Aston. Como produção científica, foram elaborados e publicados os seguintes trabalhos:

- MENDES, R. S.; WANNER, E. F. ; SARUBBI, J. F. M. ; MARTINS, F. V. C. *Optimization of the Vehicle Routing Problem with Demand Responsive Transport Using the NSGA-II Algorithm*. IEEE Intelligent Transportation Systems Society Conference Management System, Rio de Janeiro, 2016.
- MENDES, R. S.; WANNER, E. F.; MARTINS, F. V. C.; SARUBBI, J. F. M. *Dimensionality Reduction Approach for Many-Objective Vehicle Routing Problem with Demand Responsive Transport*. Lecture Notes in Computer Science, 2017. ISSN: 0302-9743
- MENDES, R. S.; LUSH, V.; WANNER, E. F.; MARTINS, F.V.C.; SARUBBI, J. F.M.; DEB, K. *Online clustering reduction based on parametric and non-parametric correlation for a many-objective vehicle routing problem with demand responsive transport*. Expert Systems With Applications, 2020. ISSN: 0957-4174
- MENDES, R. S.; WANNER, E. F.; MARTINS, F. V. C.; DEB, K. *Aggregation or Selection? Clustering Many Objectives for Vehicle Routing Problem with Demand Responsive Transport*. IEEE Congress on Evolutionary Computation, Krakow, 2021.

1.5 Organização do Trabalho

Este trabalho está organizado da seguinte forma:

- O Capítulo 2 apresenta uma descrição detalhada para o problema de roteamento de veículos com transporte reativo a demanda (PRVTRD), a forma como as soluções são representadas e codificadas. A formulação com oito objetivos é apresentada e ao final do capítulo é realizada a revisão da literatura para o problema em estudo neste trabalho.
- O Capítulo 3 descreve o problema de otimização multiobjetivo, alguns conceitos da área, as principais diferenças em relação à otimização com muitos objetivos e as maiores dificuldades encontradas. As principais classes de algoritmos para resolver problemas com muitos objetivos encontradas na literatura são apresentadas no final do capítulo.
- O Capítulo 4 apresenta as técnicas para redução de objetivos utilizadas neste trabalho: Análise de *Cluster* e Seleção de Atributos. São apresentadas os principais conceitos e características de cada técnica utilizada. Após a apresentação das técnicas, o algoritmo de redução de dimensionalidade *online* dentro do ciclo evolutivo do algoritmo é apresentado. Ao final do capítulo, alguns pontos adicionais são apresentados: como os pesos das funções objetivo foram gerados e a construção dos operadores genéticos utilizados.
- No Capítulo 5 são apresentados os principais resultados obtidos através da redução do número de objetivos *online* usando duas medidas de similaridade, da redução de objetivos *offline* e com o uso da seleção de atributos não supervisionada. São apresentados os *boxplots* para os valores do indicador de hipervolume para cada conjunto Pareto obtido para os algoritmos. São feitos testes estatísticos para descobrir qual dos algoritmos obteve melhor desempenho em cada instância testada. O diagrama de cordas também é utilizado neste capítulo para avaliar os resultados qualitativamente.
- No Capítulo 6 são apresentadas as conclusões obtidas após as execuções dos experimentos propostos e novas perspectivas de trabalhos futuros.
- Por fim, os apêndices apresentam todos os *boxplots*, tabelas dos testes de aleatoriedade e todos os diagramas de corda para os experimentos feitos com as técnicas de seleção de atributos. As referências bibliográficas são apresentadas ao final do trabalho.

2 O Problema de Roteamento de Veículos com Transporte Reativo a Demanda

Este capítulo apresenta o problema de otimização denominado Problema de Roteamento de Veículos com Transporte Reativo a Demanda (PRVTRD), o qual foi usado para se testar as técnicas de redução de dimensionalidade. Inicialmente são apresentadas as definições e características do problema com um exemplo de solução. Em seguida é mostrado o algoritmo de construção de uma solução e ao final é apresentada a formulação com muitos objetivos utilizada neste trabalho.

2.1 Descrição do Problema

No trabalho de [Gomes, Souza e Dias \(2010\)](#) é feita a primeira definição do denominado Problema de Roteamento de Veículos com o Transporte Reativo a Demanda (PRVTRD). O problema lida com as características encontradas nos problemas amplamente conhecidos de roteamento de veículos, serviços de compartilhamento de carros e gerenciamento de sistemas de transporte reativos a demanda. Para atender às solicitações dos passageiros, os veículos possuem horários e rotas flexíveis, os quais podem ser alterados de acordo com a demanda com o intuito de atender o maior número de usuários e manter a qualidade do serviço. A qualidade do serviço é medida em relação a duração das viagens e dos horários de embarque/desembarque dos passageiros.

Nos últimos anos, alguns pesquisadores adotaram abordagens diferentes para resolver o PRVTRD, desde modelos exatos [Viana e Santos \(2015a\)](#), [Gomes \(2012\)](#) a heurísticas e meta-heurísticas [Gomes, Souza e Dia \(2010\)](#), [Gomes, Souza e Dias \(2011\)](#), [Gomes, Souza e Dias \(2014\)](#). O PRVTRD lida com perspectivas diferentes [Gomes, Souza e Dias \(2010\)](#) e diversas funções objetivo conflitantes, as quais devem ser otimizadas ao mesmo tempo. Dessa forma, o problema pode ser considerado um problema com muitos objetivos.

A versão do PRVTRD adotada neste trabalho é baseada em [Gomes, Souza e Dias \(2010\)](#). Em resumo, as principais características do problema são: veículos começam e terminam a rota em um único depósito; as janelas de tempo de embarque e desembarque são especificadas pelos usuários; em cada ponto de parada, o veículo pode estar realizando o embarque ou desembarque de passageiros; uma frota de veículos com a mesma capacidade (frota homogênea); múltiplas janelas de tempo em cada ponto de parada; as

solicitações de transporte podem ser de qualquer lugar para qualquer lugar (considerando os pontos de parada definidos); janelas de tempo de embarque devem ser respeitadas e janelas de tempo de desembarque podem ser violadas, no entanto, sendo penalizadas.

Usando a notação baseada em [Gomes, Souza e Dias \(2010\)](#) para o PRVTRD, um grafo $G = (V, E)$ em que V é o conjunto de vértices e E é o conjunto de arestas pode ser usado para modelar o problema. O conjunto V representa o conjunto dos pontos de embarque, desembarque e o depósito. O vértice 0 indica o depósito. A distância d_{ij} e o tempo de viagem t_{ij} são associado a cada arco a_{ij} conectando o vértice i ao j em que $i \neq j$ e $i, j \in V$. Uma frota homogênea é representada pelo conjunto de veículos $K = \{K_1, K_2, \dots, K_{|K|}\}$. R_k representa a rota do veículo k . A taxa de ocupação do veículo k quando deixa o ponto de embarque/desembarque i é representado por q_{ki} . As solicitações de transporte dos passageiros é representada pelo conjunto $R = \{r_1, r_2, \dots, r_{|R|}\}$. Cada passageiro corresponde a uma e somente uma solicitação.

No depósito, os passageiros não podem embarcar no veículo. Um vértice pode ser visitado pelo veículo mais de uma vez se as visitas não forem consecutivas. Ao chegar em um ponto de embarque, cada veículo pode esperar por um intervalo de tempo específico permitindo o embarque de passageiros. Se não houver mais passageiros para embarque, não há necessidade de o veículo aguardar no ponto de parada.

Os conjuntos P_{in_v} e P_{out_v} armazenam, respectivamente, as solicitações de embarque e desembarque em v para cada $v \in V - \{0\}$. Como no depósito os passageiros não podem embarcar no veículo, então $P_{in_0} = P_{out_0} = \emptyset$. O índice de carga m_v calculado como o número de solicitações de embarque menos o número de solicitações de desembarque em um vértice. Sua equação é dada da seguinte forma:

$$m_v = |P_{in_v}| - |P_{out_v}| \text{ com } v \in V \quad (1)$$

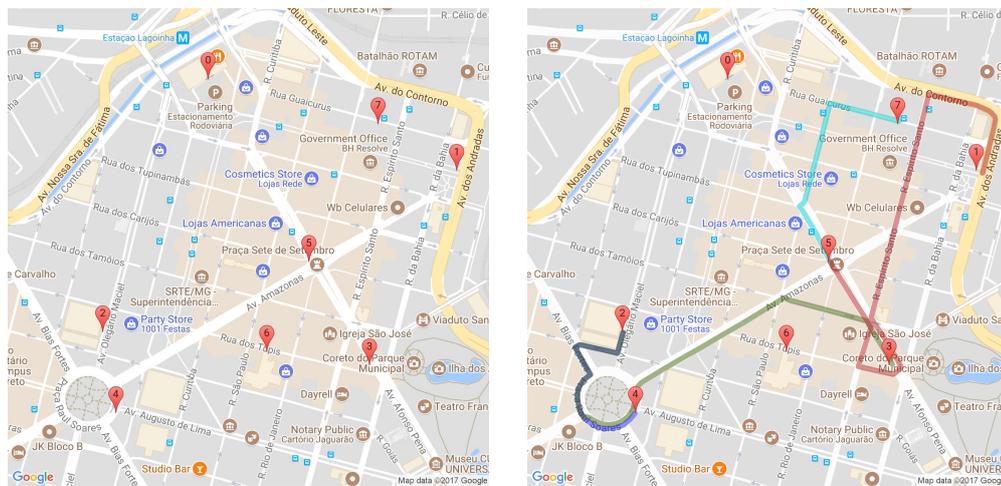
As informações a seguir são encontradas em todas as solicitações $r_l \in R$ com $l = 1, 2, \dots, |R|$: identificação (id); vértice de embarque (p^+); vértice de desembarque (p^-); janela de tempo de embarque ($[PL_l, PU_l]$); janela de tempo de desembarque ($[DL_l, DU_l]$); o momento em que um passageiro (solicitação) l embarca no veículo k (PT_l^k) e desembarca do veículo k (DT_l^k). O conjunto U representa as solicitações inviáveis, ou seja, aquelas que a frota de veículos não consegue atender. O tamanho deste conjunto é definido como um parâmetro de entrada para o problema.

Para ilustrar, considere um pequeno exemplo do problema. A Figura 1 mostra o mapa da parte central da cidade de Belo Horizonte, no qual são mostrados oito pontos de parada (marcadores vermelhos). O ponto com o rótulo 0 representa o depósito, onde os veículos começam e terminam suas rotas. Os veículos aguardam na garagem até que uma nova solicitação seja feita. As solicitações estão descritas na Tabela 1.

Tabela 1 – Para ilustrar uma pequena instância de PRVTRD, quatro solicitações e seus respectivos pontos de embarque e desembarque, janela de tempo de embarque e janela de tempo de desembarque são fornecidos. PL, PU, DL e DU representam, respectivamente, os limites inferiores e superiores para as janelas de tempo de embarque e desembarque.

Id	Embarque	Desembarque	Janela de Embarque		Janela de Desembarque	
			PL	PU	DL	DU
A	2	4	08h15	08h25	08h20	08h30
B	2	3	08h15	08h25	08h20	08h30
C	7	5	08h25	08h35	08h40	08h50
D	5	1	08h30	08h40	08h40	08h50

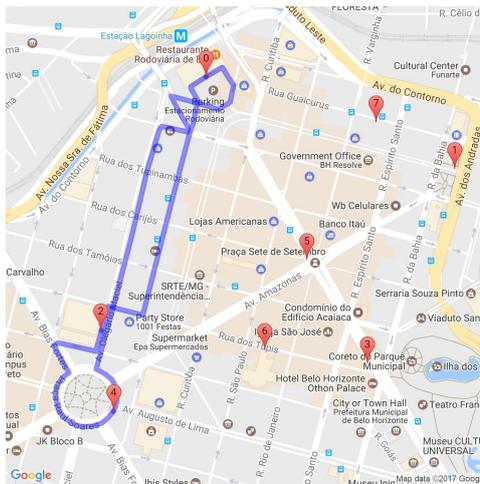
O deslocamento das solicitações pode ser visto na Figura 1. Na Figura 1a, oito pontos de parada (marcadores vermelhos) são mostrados. A rota entre a origem e o destino de cada passageiro é mostrada e cada rota é representada em uma cor diferente na Figura 1b.



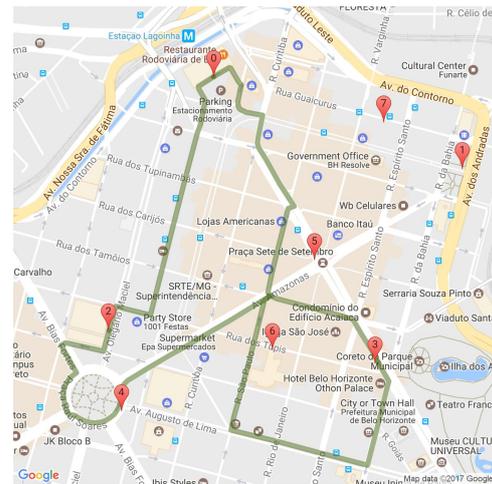
(a) Exemplo do estado inicial do problema. Marcadores em vermelho representam os pontos de parada dos veículos. (b) Exemplo de solicitações de transporte. O deslocamento entre a origem e o destino de cada passageiro é mostrado pelas linhas coloridas.

Figura 1 – Estado inicial e o deslocamento de solicitações para o problema. Os mapas são feitos usando o Google Maps®.

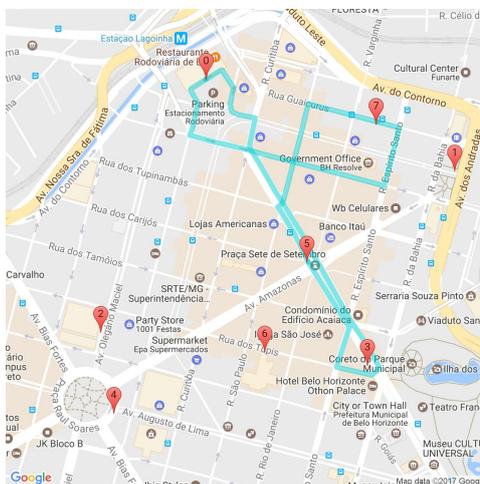
A Figura 2 mostra cada uma das rotas para cada solicitação do exemplo. Considera-se que cada passageiro é atendido por um táxi que parte do depósito. Em todas as rotas, é considerado o menor caminho para se chegar do depósito ao ponto de embarque, entre o ponto de embarque e o ponto de desembarque, e do ponto de desembarque ao depósito. Em todos os casos, considera-se o sentido das ruas.



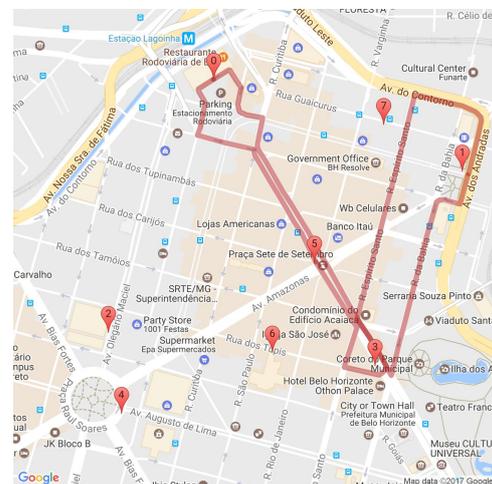
(a) Rota para solicitação A.



(b) Rota para solicitação B.



(c) Rota para solicitação C.



(d) Rota para solicitação D.

Figura 2 – Rotas para atender as solicitações. Considera-se que cada passageiro solicitou um táxi do depósito para transportá-lo entre a origem e o destino. Os mapas são feitos usando Google Maps®.

Duas soluções possíveis para atender todos os passageiros também são mostradas na Figura 3. A primeira solução, Figura 3a, mostra todas as rotas realizadas pelo transporte convencional e individual, como no caso de um serviço de táxi, o qual utilizaria quatro veículos para atender todos os passageiros. Na Figura 3b, uma segunda solução é apresentada, e, neste caso, a abordagem de transporte reativo a demanda é usada para resolver o problema com apenas dois veículos. As rotas possíveis para atender todos os passageiros são $R_1 : (0,2,4,3,0)$ e $R_2 : (0,5,7,5,1,0)$.

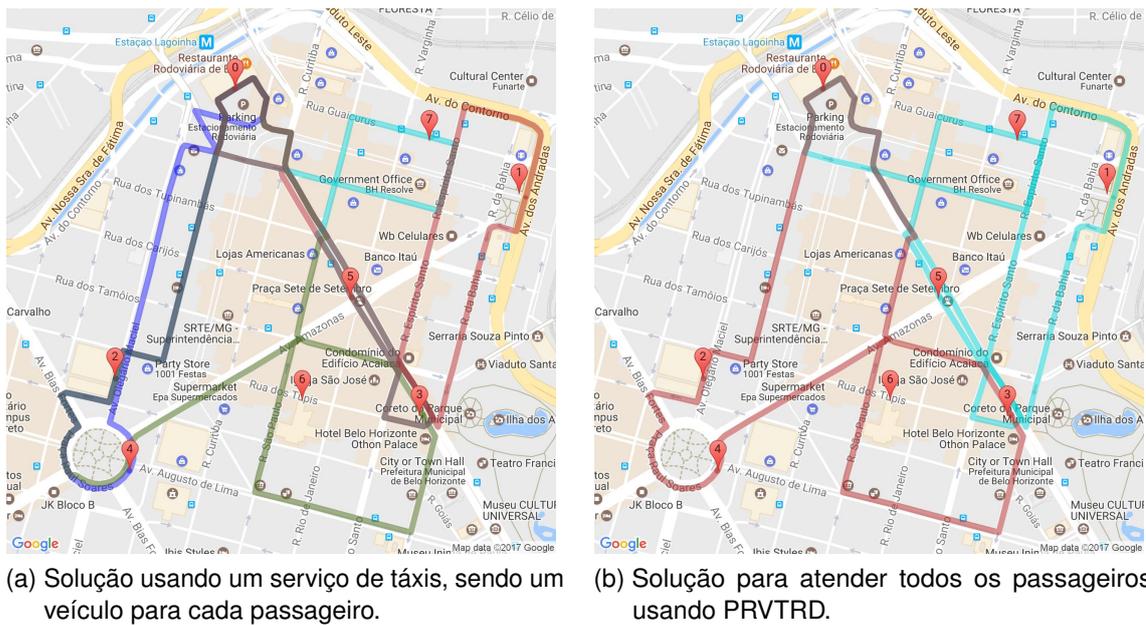


Figura 3 – Duas soluções TRD diferentes para atender às solicitações. Os mapas são feitos usando Google Maps®.

2.2 Representação da Solução do Problema

Um subconjunto de n rotas R_i compõe uma solução para o PRVTRD, por exemplo, $S = \{R_1, \dots, R_n\}$, $n \leq k$ no qual k é o número de veículos disponíveis na frota K . Cada rota é uma lista de vértices que o veículo deve cobrir.

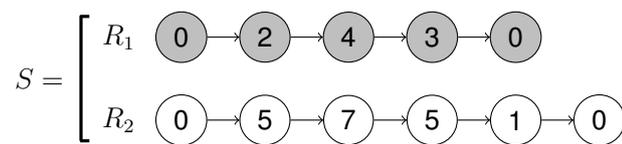
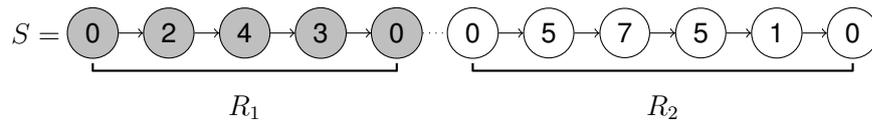


Figura 4 – Representação da solução, baseada na Figure 3b.

A Figura 4 mostra um exemplo de solução. Uma rota pode conter vértices que já foram visitados para que outras solicitações possam ser atendidas por um único veículo, desde que a visita não seja consecutiva. Considerando o exemplo das rotas nas Figuras 3b e 4, uma solução para atender todas as solicitações seria $S = \{R_1, R_2\}$.

Um movimento é uma modificação na solução que permitem a exploração de um conjunto de soluções, denominada vizinhança Miranda (2012). Para realizar algum movimento em uma solução do PRVTRD, um vetor de rotas concatenadas é criado e um movimento pode ser executado neste vetor. O objetivo da criação do vetor de rotas concatenadas é realizar alterações na sequência dos pontos de parada buscando melhorar a qualidade de uma solução existente. A Figura 5 mostra o processo no qual uma rota é concatenada a outra. Nesta primeira etapa, o depósito ainda não foi removido.

Figura 5 – Rota R_2 é concatenada após a rota R_1 .

Após a remoção do depósito, a rota concatenada pode ser vista na Figura 6. O movimento pode realizar trocas inter ou intra-rotas, ou seja, pontos de parada contidos em uma rota podem ser trocados com outra rota ou podem ter a posição alterada na rota que pertence.

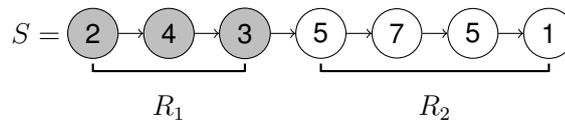


Figura 6 – Rotas concatenadas.

Depois de aplicar algum movimento no vetor de rotas concatenadas, é necessário reconstruir a solução para que ela possa ser avaliada. O algoritmo de reconstrução da solução, o qual é baseado no Algoritmo 1 (descrito na Seção 2.3), usa o vetor de rotas concatenadas para tentar inserir cada nó em uma nova rota. Se alguma restrição for violada no nó atual, o algoritmo tentará inserir o próximo nó e assim por diante.

2.3 Algoritmo de Construção Gulosa

No trabalho de [Gomes, Souza e Dias \(2010\)](#) é proposto o algoritmo de construção gulosa para o PRVTRD. Os objetivos são classificados em duas perspectivas: veículo/empresa e passageiro. Na perspectiva do veículo é considerado o custo das rotas (em termos de distância) e o fluxo de passageiros por nó. A perspectiva do passageiro leva em conta o tempo total de espera dos passageiros no ponto de parada e a duração de todas as viagens.

Uma vez que é feito um algoritmo guloso para construir uma solução do problema, é necessário definir uma função que avalie cada elemento candidato a ser inserido na solução. No caso do PRVTRD, [Gomes, Souza e Dias \(2010\)](#) propõem a Função de Ranqueamento de Nó, do inglês *Node Ranking Function* (NRF) que calcula qual o melhor ponto de parada a ser inserido em uma determinada iteração. O cálculo da NRF é feito por meio da soma ponderada das avaliações de cada perspectiva descrita anteriormente. Avalia-se cada elemento e em seguida, são criadas listas com os valores de cada critério para todos os pontos de parada. O valor de NRF para um ponto de parada v é calculado de acordo com a Equação 2.

$$NRF_v = (\alpha_1 \times CRL_v + \alpha_2 \times NRL_v) + (\alpha_3 \times DRL_v + \alpha_4 \times TRL_v), \forall v \in V - \{0\} \quad (2)$$

em que α_1 é o peso do fator distância; CRL_v é o custo de viagem (distância) normalizado; α_2 é o peso do número de passageiros; NRL_v é o índice de lotação m_v normalizado, calculado de acordo com a Equação (1); α_3 é o peso do fator janela de tempo de desembarque; DRL_v é o limite inferior da janela de tempo de desembarque normalizado; α_4 é o peso do fator janela de tempo de embarque e TRL_v é o limite inferior da janela de tempo de embarque normalizado. A única restrição para os valores $\alpha_1, \alpha_2, \alpha_3$ e α_4 é que devem respeitar a Equação (3).

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1 \quad (3)$$

O processo de construção gulosa da solução é apresentado no Algoritmo 1. Define-se como entrada do algoritmo um conjunto de solicitações e pesos para o problema, e como saída, uma solução.

Algoritmo 1: Construção de uma solução gulosa para o PRVTRD

Input: conjunto de solicitações R e conjunto de pesos α

Output: solução S

Inicializa solução e conjunto de solicitações inviáveis

while existir solicitações e disponibilidade na frota **do**

 Inicializa nova rota e nó corrente

 Aloca um veículo para a rota

while existir solicitações **do**

 Encontrar nós viáveis a partir do nó corrente

for cada nó viável **do**

 | Avaliar e Normalizar: CRL, NRL, DRL e TRL

end

 Calcular NRF para cada nó e adicionar na rota aquele com maior valor

if não existir mais solicitações para serem atendidas **then**

 | Finalizar rota

 | Adicionar rota na solução

end

if existir solicitações inviáveis **then**

 | Atualizar conjunto de solicitações

 | Esvaziar conjunto de solicitações inviáveis

end

end

end

O funcionamento do algoritmo de construção gulosa para o PRVTRD ocorre da seguinte forma: a solução S e o conjunto de solicitações não atendidas são inicializados. Em seguida, o laço principal do algoritmo é iniciado, que ocorre enquanto houver veículos na frota e solicitações viáveis. A rota e o nó corrente são inicializados tendo como ponto de partida o depósito. Logo após, aloca-se um veículo para a rota em construção. Um segundo laço é iniciado e enquanto existir solicitações factíveis o algoritmo tenta inserir outros nós na rota. Esse passo é feito analisando a viabilidade de cada nó e para aqueles

factíveis os valores de CRL , NRL , DRL e TRL são calculados e normalizados. O valor de NRF é obtido em seguida e o nó que apresentar maior valor de NRF é inserido na rota. Atualiza-se o conjunto de solicitações e o conjunto de solicitações não atendidas, retirando ou inserindo solicitações que desembarcaram/embarcaram na rota. No caso de não existir mais alguma solicitação viável, a rota é finalizada inserindo novamente o depósito como ponto final e é inserida no conjunto de rotas da solução. Confere-se a existência de alguma outra solicitação que ainda possa ser atendida, caso afirmativo, o laço principal é novamente executado.

2.4 Formulação com Muitos Objetivos

Esta seção apresenta os objetivos usados na abordagem proposta de PRVTRD. Os objetivos foram escolhidos para que o problema possa ser avaliado de diferentes perspectivas. As funções f_1 a f_6 , encontradas em [Mendes et al. \(2017a\)](#), foram reescritas em [Mendes et al. \(2021\)](#) e serão apresentadas a seguir. Considere que:

- $D(RT_k)$ é a distância total da rota associada ao veículo k .
- DT_l é o momento em que o passageiro associado à solicitação l desembarca do veículo.
- DU_l é o limite superior da janela de tempo de desembarque associada à solicitação l .
- R_{MAX} é a rota com a maior duração.
- R_{MIN} é a rota com a duração mais curta.
- PT_l é o momento em que o passageiro associado à solicitação l embarca no veículo.
- PL_l é o limite inferior da janela de tempo de embarque associada à solicitação l .
- q_{ki} é a taxa média de ocupação do veículo k saindo do ponto de coleta i .

As oito funções objetivo podem ser definidas como:

$$\text{Minimize } f_1 = \sum_{k=1}^{|S|} D(RT_k), \quad (4)$$

$$\text{Minimize } f_2 = \sum_{l=1}^{|R|} \max(DT_l - DU_l, 0), \quad (5)$$

$$\text{Minimize } f_3 = R_{MAX} - R_{MIN}, \quad (6)$$

$$\text{Minimize } f_4 = |K|, \quad (7)$$

$$\text{Minimize } f_5 = \sum_{l=1}^{|R|} (DT_l - PT_l), \quad (8)$$

$$\text{Minimize } f_6 = \sum_{l=1}^{|R|} \max(PT_l - PL_l, 0), \quad (9)$$

$$\text{Minimize } f_7 = \sum_{l=1}^{|R|} |\min(DT_l - DU_l, 0)|, \quad (10)$$

$$\text{Minimize } f_8 = 1 - \frac{1}{|K|} \sum_{k=1}^{|K|} \sum_{i \in R_k / \{0\}} \frac{q_{ki}}{|R_k / \{0\}|}. \quad (11)$$

As funções objetivo acima indicam respectivamente: (i) a distância total percorrida, (ii) o atraso no desembarque dos passageiros, (iii) a diferença entre a rota mais longa e a mais curta (levando em consideração o tempo total), (iv) a quantidade de veículos usados, (v) o tempo total de viagem, (vi) o tempo de espera dos passageiros, (vii) a minimização da antecipação da janela de tempo total de desembarque, e (viii) a média de assentos livres nos veículos.

Vale ressaltar que, em [Mendes et al. \(2017a\)](#), o número de solicitações não atendidas é incluído na formulação do problema como uma função objetivo a ser minimizada. Neste trabalho, a porcentagem de solicitações não atendidas é incluída como uma restrição de qualidade de serviço e é um parâmetro de entrada do usuário. Considera-se que o passageiro chega no primeiro minuto da janela de horário de embarque e, a partir desse momento, o tempo de espera é calculado como a diferença entre o tempo que o veículo passa pelo ponto de embarque e o início da janela de horário de embarque. Se o veículo chegar ao ponto de embarque exatamente no início da janela de tempo de embarque e o passageiro ainda não estiver lá, não haverá embarque. O tempo de espera é utilizado no caso de se atender mais solicitações e é um valor predefinido pelo usuário.

2.5 Revisão da Literatura para o PRVTRD

O trabalho de [Gomes, Souza e Dias \(2010\)](#) resolve o PRVTRD por meio de duas metaheurísticas: *GRASP-like* e *GRASP-Reativo*, sendo estas duas diferentes versões do GRASP (*Greedy Randomized Adaptive Search Procedure*), proposto por [Feo e Resende \(1995\)](#). Na primeira versão, o algoritmo *GRASP-like*, executa apenas uma construção gulosa da solução. Na segunda versão, o *GRASP-Reativo*, é executado as duas fases do GRASP original e uma lista de parâmetros que é recalculada quando boas soluções são encontradas.

[Miranda \(2012\)](#) resolve o PRVTRD na forma estática e dinâmica usando os algoritmos *GRASP-like*, *GRASP-Reativo* e também com uma outra versão *GRASP-Tabu*. O algoritmo *GRASP-Tabu* usa a metaheurística *Busca Tabu* como busca local. O algoritmo *GRASP-Tabu* obteve os melhores resultados em relação aos outros algoritmos utilizados.

No trabalho de [Chevrier et al. \(2012\)](#) é proposto a utilização de três funções objetivo ao mesmo tempo para resolver um problema de TRD. As funções minimizam: a quantidade

de veículos utilizados, tempo total das jornadas e o atraso total. Os autores resolveram o problema usando uma abordagem evolucionária multiobjetivo não fazendo nenhuma redução dos objetivos ou simplificação. Os algoritmos NSGA-II, SPEA2 e IBEA foram aplicados com uma metaheurística (ILS) como operador de mutação. A utilização do ILS melhorou significativamente os resultados obtidos.

Gomes propõe um modelo matemático para o PRVTRD em sua tese no ano de 2012. O modelo pode ser encontrado em [Gomes \(2012\)](#). Considera-se quatro funções objetivo: custo total, tempo de espera dos passageiros, tempo total das viagens e número de passageiros atendidos. Apesar de Gomes apresentar uma formulação multiobjetivo, o problema é resolvido aplicando-se o algoritmo *GRASP-Reativo* para uma formulação mono-objetivo em que é feita a soma ponderada dos objetivos. O autor apresenta um modelo de sistema de apoio à decisão para o TRD a ser implementado.

Em [Gomes, Souza e Dias \(2014\)](#), os autores utilizam uma versão paralelizada do *GRASP-Reativo* com o objetivo de obter uma solução em um menor tempo para o PRVTRD. O algoritmo utilizado foi integrado a um sistema de apoio à decisão, descrito em [Gomes, Souza e Dias \(2011\)](#) e [Gomes \(2012\)](#). Esse sistema possui um módulo de simulação que utiliza modelos que levam em consideração a área de serviço, solicitações de viagem, veículos e tempo real dos eventos.

[Schyns \(2015\)](#) desenvolve um *framework* para resolver o roteamento dos caminhões de abastecimento no aeroporto de Liège na Bélgica. A frota de caminhões abastece todos os aviões do aeroporto. O autor considera o problema em questão correlato a uma abordagem de TRD uma vez que o abastecimento deve ser de tal forma que seja capaz de ser responsivo a imprevistos. O objetivo do trabalho do Schyns é atender toda demanda de combustível o mais rápido possível, visto o alto custo de um avião parado no aeroporto. O *framework* utiliza o algoritmo de otimização por colônia de formigas (*Ant Colony Optimization - ACO*).

Em [Viana e Santos \(2015b\)](#) os autores propõem dois modelos matemáticos para um serviço de transporte reativo a demanda. Os modelos propostos são baseados no modelo matemático do DARP. Cada um dos modelos propostos apresentam objetivos diferentes: minimizar o número de veículos e minimizar o tempo total de duração das rotas. O problema foi resolvido em duas etapas, encontrando-se o número mínimo de veículos que possam atender todos os passageiros e, em seguida, reduzindo o tempo gasto por cada veículo utilizado na primeira etapa.

O trabalho de [Viana, Santos e Arroyo \(2015\)](#) usa uma abordagem multiobjetivo para um serviço baseado em TRD. Os autores propõem um modelo matemático para minimizar três funções objetivo: o número de veículos utilizados, tempo total de viagem e o atraso total no desembarque. A representação da solução que os autores utilizaram é a mesma encontrada em [Chevrier et al. \(2012\)](#). O algoritmo NSGA-II foi utilizado para resolver o

problema e os operadores genéticos foram feitos de modo que a viabilidade das soluções fosse garantida.

Em [Mendes et al. \(2016\)](#) os autores fazem uma redução do número de objetivos do PRVTRD. Essa redução, também chamada de redução de dimensionalidade, é feita para tornar o problema mais simples, uma vez que algoritmos multiobjetivos possuem dificuldades para resolver problemas com mais de três objetivos. Tais dificuldades serão detalhadas no Capítulo 3.2. O problema original é composto por cinco objetivos e a redução de dimensionalidade é feita usando a Árvore de Agregação, de forma que o problema resultante seja bi-objetivo. De acordo com a agregação feita pela árvore, os objetivos foram divididos em duas perspectivas: humana e empresa. Os algoritmos utilizados para resolver o problema foram o NSGA-II e SPEA 2, sendo que o NSGA-II conseguiu resultados significativamente melhores.

Um problema correlato ao PRVTRD pode ser encontrado em [Huang et al. \(2020\)](#). Neste trabalho, os autores abordam o problema de projeto de rede reativo a demanda para uma frota de ônibus customizados. O problema é resolvido em duas fases: inserindo-se solicitações dinamicamente de forma interativa e, em seguida, otimizando o atendimento rede com base na demanda geral.

No trabalho de [Xiong et al. \(2020\)](#) o problema de roteamento busca gerar rotas e horários de partida dos veículos para viagens entre pontos de ônibus e estações de metrô. Um modelo de otimização inteira mista é criado buscando minimizar o custo total do sistema, incluindo o custo de operação e custo no veículo do passageiro. As restrições utilizadas são referente à duração da rota, janela de tempo, desvios e capacidade dos veículos. Os autores resolvem o problema usando duas metaheurísticas: Busca Tabu e VNS (*Variable Neighborhood Search*).

Uma abordagem baseada em dados (do inglês *data driven*) é feita em [Liu e Wang \(2022\)](#). A análise de *cluster* é utilizada para agrupar os passageiros de acordo com as informações relacionadas aos pontos de embarque/desembarque e seus respectivos horários de embarque/desembarque. Apesar dos autores buscarem minimizar o custo de operação dos ônibus e do transporte dos passageiros, a função objetivo é a agregação dos dois custos mencionados.

No trabalho de [Melis e Sorensen \(2022\)](#) é proposto o problema de roteamento de ônibus por demanda para o transporte de passageiros. A principal diferença entre este problema e o PRVTRD é o fato dos usuários escolherem um conjunto de pontos de parada para embarque e desembarque. Os autores apresentam uma heurística como forma de resolver o problema. Os resultados são comparados aos de uma simulação de um sistema de transporte público tradicional com linhas e horários fixos. Para o problema em questão, os resultados mostraram que para uma frota maior e veículos com menor capacidade, o tempo total de deslocamento dos usuários pode ser significativamente reduzido.

Como contribuições desta tese, foram publicados quatro trabalhos relacionados a redução de dimensionalidade do PRVTRD com muitos objetivos. Os trabalhos desenvolvidos serão descritos a seguir.

Em [Mendes et al. \(2016\)](#), a formulação do PRVTRD utiliza a função objetivo que calcula o número de solicitações não atendidas como um terceiro objetivo a ser minimizado. O algoritmo NSGA-II é aplicado para resolver o problema e os resultados mostram que o algoritmo pode melhorar as soluções não-dominadas no espaço objetivo original.

A análise de *cluster* também é usada para reduzir a dimensão do PRVTRD com sete funções objetivo em [Mendes et al. \(2017a\)](#). A correlação entre as funções objetivo é usada como uma medida de similaridade para determinar as distâncias entre funções. Apesar dos bons resultados obtidos, a aplicação da técnica de *cluster* é feita apenas na formulação de sete objetivos e em apenas uma instância do problema. Além disso, correlação é uma medida linear de distância entre dois vetores e uma medida linear pode não ser a melhor maneira de calcular o distância entre funções em um problema com muitos objetivos. Um ponto favorável que pode ser observado nessa abordagem está relacionado à inclusão de mais funções objetivo na formulação original.

Recentemente, em [Mendes et al. \(2021\)](#), é feita uma abordagem de redução de dimensionalidade baseada em agregação, em que a análise de cluster foi aplicada, e a redução de dimensionalidade foi realizada em cada geração do MOEA baseado em decomposição (MOEA/D) [Zhang e Li \(2007\)](#). Acoplando a abordagem de clustering online com os coeficientes de correlação de Pearson e Kendall dois algoritmos foram utilizados: OnCL $_{\rho}$ -MOEA/D e OnCL $_{\tau}$ -MOEA/D. O desempenho dos algoritmos foi comparado no espaço de objetivo original por meio do cálculo do hipervolume. Os resultados mostraram que a abordagem de cluster online é um método eficaz como uma técnica de redução de dimensionalidade transformando uma formulação com muitos objetivos do PRVTRD em uma formulação bi-objetivo.

Uma análise comparativa entre o uso de técnicas de seleção e agregação de objetivos foi feita sobre o PRVTRD em [Mendes et al. \(2021\)](#). Inicialmente o algoritmo de clusterização é aplicado obtendo-se dois centróides e duas técnicas (Dispersion Ratio [Liu, Chen e Wu \(2018\)](#), Roy et al. (2019) e Maximum Variance [Dy e Brodley \(2004\)](#)) foram usadas para escolher qual função objetivo representaria cada cluster. É obtida uma formulação bi-objetivo e em seguida o algoritmo MOEA/D [Zhang e Li \(2007\)](#) é utilizado para resolver o problema. Assim, mostrou-se que eliminar objetivos pode acarretar na perda de informação significativa do problema. Além disso, pode ser visto que ao se executar a redução de dimensionalidade entre um número pré-definido de gerações pode ser uma boa estratégia para se obter conjuntos Pareto de mesma qualidade em um menor custo computacional.

3 Otimização Multiobjetivo

Vários problemas encontrados na indústria, engenharia, como em diversas outras áreas, envolvem a otimização de mais de um objetivo ao mesmo tempo. Tais objetivos apresentam um certo grau de conflito ou estão em unidades diferentes [Jaimes \(2011\)](#). Problemas que pertencem a esse campo são chamados de problemas de otimização multiobjetivo.

Este capítulo apresenta os principais conceitos de otimização multiobjetivo utilizados neste trabalho. Inicialmente, será apresentada a definição de otimização multiobjetivo. Em seguida, será apresentada a definição de otimização com muitos objetivos, suas principais dificuldades e as classes de algoritmos evolutivos para resolver os problemas com essa formulação.

3.1 Principais Conceitos

A otimização de problemas do mundo real comumente representa situações em que vários objetivos devem ser otimizados ao mesmo tempo, porém com um certo grau de conflito entre eles. Assim, não há solução única em relação a todos os objetivos. Em vez disso, existe um conjunto de soluções, chamadas soluções Pareto-ótimas, que formam o conjunto Pareto-ótimo, representando o melhor compromisso entre os objetivos. Cada objetivo no conjunto Pareto-ótimo não pode ser refinado sem a deterioração de outro objetivo [Cruz et al. \(2007\)](#).

3.1.1 Problema de Otimização Multiobjetivo

O problema de otimização multiobjetivo pode ser formulado da seguinte maneira:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \mathbf{f}(\mathbf{x}),$$

$$\text{sujeito a: } \begin{cases} g_i(\mathbf{x}) \leq 0; & i = 1, 2, \dots, r \\ h_j(\mathbf{x}) = 0; & j = 1, 2, \dots, p \end{cases} \quad (12)$$

no qual $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{f}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{g}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ e $\mathbf{h}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^p$. As funções g_i e h_j são, respectivamente, funções de restrições de desigualdade e de igualdade. Os vetores $\mathbf{x} \in \mathbb{R}^n$ são chamados *vetores de parâmetros* do problema multiobjetivo e formam o *espaço de parâmetros*. Os vetores $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$ encontram-se em um espaço vetorial chamado *espaço de objetivos*.¹

¹ Para diferenciar o uso de um escalar para vetor, é utilizada a notação em negrito, como o caso do vetor \mathbf{x} .

O espaço de variáveis, denotado por x , é onde se faz a busca pelas soluções do problema e o espaço de objetivos, denotado por $f(x)$, é o espaço formado pelas funções objetivo do problema. A Figura 7 mostra um exemplo de mapeamento feito pelas funções objetivo entre o espaço de variáveis e o espaço de objetivos.

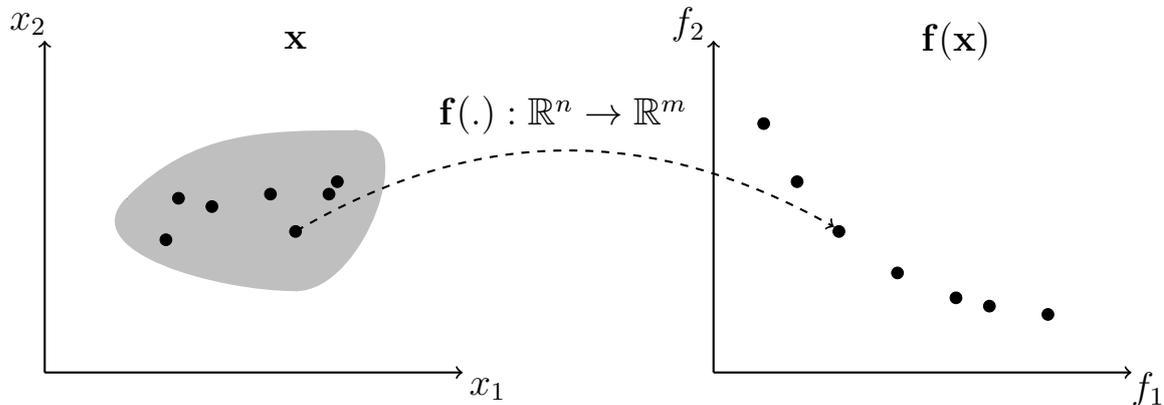


Figura 7 – Mapeamento do espaço de parâmetros no espaço de objetivos. Adaptado de Deb (2008).

A seguir, algumas definições, encontradas em Deb (2008), serão apresentadas:

Definição 1 (Dominância). *Uma solução $x' \in x$ domina outra solução $x'' \in x$ se $f(x') \leq f(x'')$ e $f(x') \neq f(x'')$. Da mesma forma, diz-se que $f(x') \in f(x)$ domina $f(x'') \in f(x)$, nessas mesmas condições. Esta relação de dominância é representada pela notação $f(x') \preceq f(x'')$.²*

Definição 2 (Dominância Estrita). *Se uma solução $x' \in x$ é estritamente melhor que a solução x'' para todos os objetivos, pode-se dizer que x' domina fortemente x'' . Esta relação pode ser representada pela notação $f(x') \prec f(x'')$.³*

Para exemplificar as relações de dominância nas Definições 1 e 2 vamos considerar a Figura 8 que mostra cinco soluções diferentes no espaço de objetivos Y . Vamos considerar também que ambas as funções f_1 e f_2 devem ser minimizadas.

Comparando-se as soluções A e C, pode ser visto que A possui menor valor de f_1 e f_2 em relação a C logo, pode-se dizer que A domina estritamente C. Usando a notação anterior, $A \prec C$. Ao se comparar a solução A com a solução D, temos que para o valor de f_1 a solução A possui menor valor e para f_2 , as duas soluções possuem o mesmo valor. Assim, a solução A domina a solução D ($A \preceq D$).

Outra situação importante acontece ao se comparar as soluções A e B. Percebe-se que a solução A possui menor valor f_1 em relação a B porém, possui maior valor de f_2 .

² A relação de ordem acontece nas coordenadas de $f(x')$ e $f(x'')$.

³ Ver nota de rodapé 2.

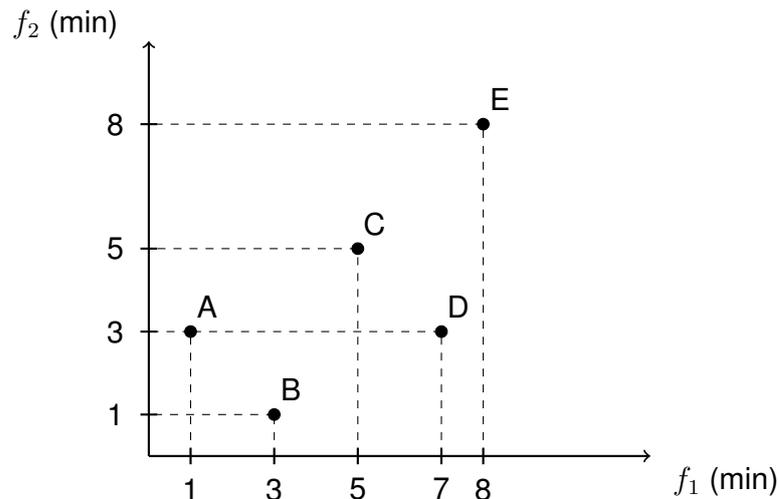


Figura 8 – Exemplo do conceito de Dominância. Adaptado de Deb (2008).

Neste caso, como as relações de dominância $A \prec B$ e $B \prec A$ ou $A \preceq B$ e $B \preceq A$ não puderam ser verificadas, as soluções são chamadas incomparáveis.

Definição 3 (Solução Pareto-Ótima). *Diz-se que $x^* \in F_x$ é uma solução Pareto-Ótima se não existe qualquer outra solução $x^* \in F_x$ tal que $f(x) \prec f(x^*)$, ou seja, se x^* não é dominado por nenhum outro ponto factível.*

Pode ser encontrado em diversos trabalhos uma outra nomenclatura para solução Pareto-ótima, sendo referida como solução eficiente ou solução não-dominada. Considerando a Figura 8, dizemos que as soluções A e B são soluções não-dominadas.

Definição 4 (Conjunto Pareto-Ótimo). *Diz-se que $X^* \subset X$ é um conjunto Pareto-Ótimo se todas as soluções que o compõem são soluções Pareto-Ótimas. O conjunto-imagem $Y^* \subset Y$ associado ao conjunto Pareto-Ótimo é chamado de fronteira Pareto-Ótima.*

3.2 Otimização com Muitos Objetivos

A otimização de muitos objetivos (do inglês *Many Objective Optimization*) é uma área importante, e desafiadora devido a baixa escalabilidade com o número de objetivos dos algoritmos multiobjetivo baseados em Pareto dominância, uma vez que há uma relação de crescimento exponencial entre o número de soluções não dominadas e o aumento de dimensões do problema. Coello et al. (2019). Nesta seção, serão apresentadas a definição de Otimização com Muitos Objetivos, suas principais dificuldades e as classes de algoritmos evolutivos para resolver os problemas com essa formulação.

3.2.1 Principal Diferença entre Otimização Multiobjetivo e com Muitos Objetivos

Um problema de otimização multiobjetivo com mais de três objetivos é chamado de problema de otimização com muitos objetivos e pode ser formulado da seguinte maneira:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} \mathbf{f}(\mathbf{x}), \\ \text{sujeito a: } &\begin{cases} g_i(\mathbf{x}) \leq 0; & i = 1, 2, \dots, r \\ h_j(\mathbf{x}) = 0; & j = 1, 2, \dots, p \end{cases} \end{aligned} \quad (13)$$

em que $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{f}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 4$, $\mathbf{g}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ e $\mathbf{h}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^p$. As funções g_i e h_j são, respectivamente, funções de restrição de desigualdade e igualdade. Vale ressaltar que a principal diferença em relação à formulação apresentada nas Equações 12 da Seção 3.1.1, é o intervalo para os valores de m .

3.3 Principais Dificuldades para Resolver os Problemas com Muitos Objetivos

Considerando o aumento no número de funções objetivo, algumas dificuldades podem surgir quando se utiliza um algoritmo multiobjetivo baseado em Pareto dominância para obter um conjunto representativo de soluções do conjunto Pareto [Deb e Saxena \(2005\)](#). Alguns problemas encontrados são listados a seguir e serão detalhados nesta seção.

1. Deterioração da habilidade de busca;
2. Soluções resistentes à dominância;
3. Aumento no número de soluções eficientes;
4. Dificuldade visualização das soluções;
5. Custo computacional dos indicadores de performance.

3.3.1 Deterioração da Habilidade de Busca

A primeira dificuldade para se resolver problemas com muitos objetivos surgem em relação à deterioração da habilidade de busca do algoritmo. Uma das razões para que isso ocorra é o fato da proporção de soluções não dominadas crescer em função do número de objetivos [Goulart e Campelo \(2016\)](#). Dessa forma, a pressão seletiva para encontrar soluções melhores é reduzida e o processo de busca ocorre mais lentamente [Deb e Saxena \(2005\)](#).

O trabalho de [Mostaghim e Schmeck \(2008\)](#) mostra que uma busca aleatória consegue ser melhor que o algoritmo NSGA-II para um problema com dez objetivos. Um exemplo

deste aumento da proporção de soluções não dominadas é encontrado em [Farina e Amato \(2002\)](#) e será apresentado a seguir.

Considerando a Figura 9, é apresentado um problema de minimização multiobjetivo em duas formulações: em \mathbb{R}^2 e \mathbb{R}^3 . Uma vez que o problema em questão é de minimização das funções objetivo, o quadrado e o cubo azul representam as regiões que possuem soluções dominadas pela solução x . O quadrado e o cubo amarelo são as regiões que possuem soluções que dominam x . Já os quadrados e cubos em branco (linhas tracejadas), são as regiões que possuem soluções não-dominadas.

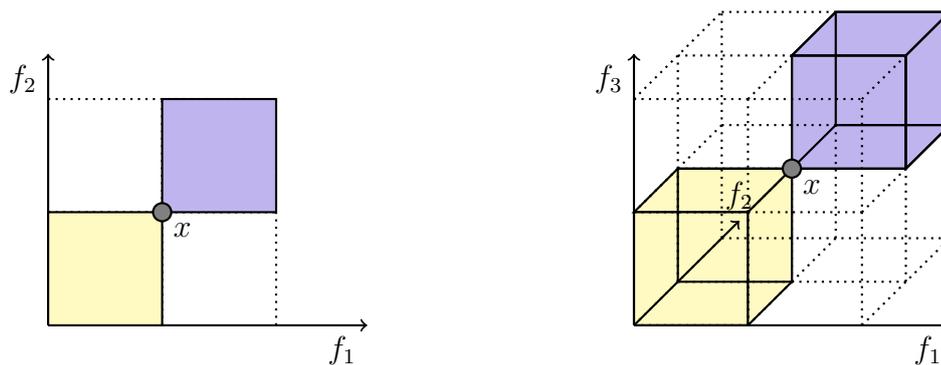


Figura 9 – Aumento das regiões não dominadas. Adaptado de [Farina e Amato \(2002\)](#).

Fazendo uma comparação das regiões não dominadas, para a formulação com dois objetivos tem-se que 1/2 das regiões (quadrados tracejados) apresentadas possuem soluções não dominadas. Ao se adicionar uma função objetivo, a proporção de regiões com soluções não dominadas aumenta para 6/8, simplificando, tem-se 3/4 que pode ser visto pelos cubos feitos por linhas tracejadas na formulação com três objetivos.

3.3.2 Soluções Resistentes à Dominância

As soluções encontradas pelos algoritmos podem ser classificadas em dois tipos: as que foram ou serão dominadas por outras soluções, e soluções que dificilmente são dominadas, e pertencem (quase) ao conjunto Pareto [Ikeda, Kita e Kobayashi \(2001\)](#). Dependendo do problema, essas soluções podem permanecer na população por muitas gerações embora sejam inferiores qualitativamente. Isso acontece porque as soluções que podem dominar essas soluções problemáticas não são fáceis de encontrar [Kita et al. \(1996\)](#).

De acordo com [Jaimes \(2011\)](#) a definição de Solução Resistente a Dominância (SRD), ou *Dominance Resistance Solution* - DRS, é uma solução que é inferior a outras soluções em pelo menos um objetivo porém, próxima dos valores ótimos para os outros objetivos. Soluções que dominam uma SRD são difíceis de serem encontradas.

Se as SRDs ocuparem uma maior área no espaço de busca do que o verdadeiro conjunto Pareto, os algoritmos baseados em Pareto dominância podem falhar em obter

a boa aproximação desse conjunto. Dessa forma, todas as soluções não dominadas são tratadas como sobreviventes, e as SRDs se espalham pelo espaço de busca e permanecem várias gerações. Assim, ocorre uma baixa intensidade na busca de soluções em direção ao conjunto Pareto Ikeda, Kita e Kobayashi (2001). Um exemplo de soluções resistentes à dominância podem ser vistos na Figura 10.

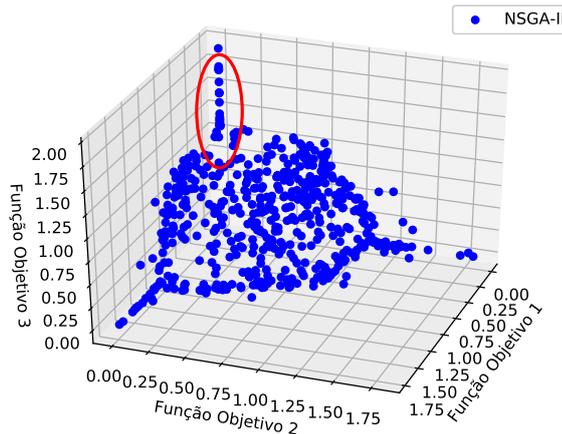


Figura 10 – Exemplo de soluções resistentes à dominância para o problema DTLZ2. Resultados obtidos usando *MOEA Framework Hadka* (2017), versão 2.12. Adaptado de Jaimes (2011).

No exemplo mostrado na figura anterior, são apresentadas diversas soluções para o problema-teste DTLZ2. As soluções dentro da elipse vermelha são SRD pois possuem valores para os objetivos 1 e 2 porém apresentam valores ruins em relação ao objetivo 3.

3.3.3 Aumento no Número de Soluções Não Dominadas

Uma outra dificuldade surge quando a dimensão de um problema multiobjetivo aumenta: para se ter uma boa aproximação da fronteira Pareto, em termos de convergência e cobertura, o número de soluções não dominadas cresce exponencialmente em relação a dimensão. Ao se pensar em recursos computacionais e tempo de execução, obter essa aproximação para o conjunto Pareto pode se tornar difícil Duro et al. (2014). Um outro caso acontece quando se usa indicadores de desempenho para avaliar os algoritmos como por exemplo, o hipervolume Cheung, Gu e Liu (2016).

De acordo com Miller (1955), Nisbett e Wilson (1977) apud Duro et al. (2014), considerando o ponto de vista psicológico, as pessoas não são boas em lidar com um volume grande de dados, nem podem lidar os fatores 7 ± 2 de uma só vez, que seria um intervalo de 5 a 9 fatores. Em outras palavras, um ser humano teria certa dificuldade de

expressar suas preferências e tomar uma decisão quando o número de opções aumenta. No contexto da otimização com muitos objetivos, uma vez que o número de soluções na fronteira Pareto aumenta, uma pessoa teria dificuldade de comparar e escolher uma solução do conjunto obtido como também haveria problemas ao definir quais seriam melhores para cada funções objetivo.

Uma relação que descreve o comportamento exponencial entre o número de soluções e a dimensão do problema é encontrada em [Deb e Saxena \(2005\)](#). Sua definição é feita da seguinte forma: sendo N o número de soluções para representar adequadamente a fronteira Pareto em uma dimensão e M o valor da dimensão, são necessários $O(N^M)$ soluções para representar a fronteira de dimensão M . Neste caso, $O(N^M)$ é o limite superior para o número de soluções necessárias para representar a fronteira.

Apesar de serem similares, [Sen e Yang \(1998\)](#) apud [Jaimes, Coello e Barrientos \(2009\)](#) apresentam uma outra relação, inserindo um valor r , chamado de resolução que é análogo ao valor de N definido em [Deb e Saxena \(2005\)](#). A expressão do número de pontos necessários para representar uma fronteira Pareto com k objetivos e resolução r é limitada por $O(kr^{k-1})$. Esta expressão é derivada assumindo que cada solução é contida dentro de um hipercubo para preservar uma distribuição uniforme das soluções em cada dimensão. Para mais detalhes da definição do hipercubo, ver [Jaimes, Coello e Barrientos \(2009\)](#).

A Tabela 2 apresenta um exemplo para os limites superiores encontrados em [Deb e Saxena \(2005\)](#) e [Sen e Yang \(1998\)](#) apud [Jaimes, Coello e Barrientos \(2009\)](#) para o número de soluções em função de diferentes números de objetivos. Para fazer os cálculos, foi considerado que o número de soluções para ter uma boa aproximação da fronteira Pareto em uma dimensão é igual a dez, ou seja, $N = r = 10$.

Tabela 2 – Relações entre o limite superior do número de soluções e dimensão do problema.

Dimensão	$O(N^M)$	$O(kr^{k-1})$
2	100	20
3	1.000	300
4	10.000	4.000
5	100.000	50.000
6	1.000.000	600.000
7	10.000.000	7.000.000

Com base nos valores encontrados, percebe-se o aumento significativo do limite superior para ambas as relações. Deste modo, reforça o fato do aumento do uso de recursos computacionais uma vez que a quantidade de soluções será grande [Duro et al. \(2014\)](#).

Para exemplificar graficamente a relação entre o aumento do número de soluções não dominadas e o aumento do número de funções objetivo, considere a Figura 11. Ela apresenta os conjuntos Pareto combinado após a aplicação do algoritmo NSGA-II. O algoritmo foi executado dez vezes com um máximo de duas mil avaliações de função

objetivo no problema DTLZ2 para as formulações com dois e três objetivos. O problema DTLZ2 pertence a um conjunto de testes de problemas teste denominado DTLZ. Este conjunto de problemas foi criado para problemas multiobjetivo e leva o nome de seus autores Deb, Thiele, Laumanns e Zitzler [Deb et al. \(2005\)](#).

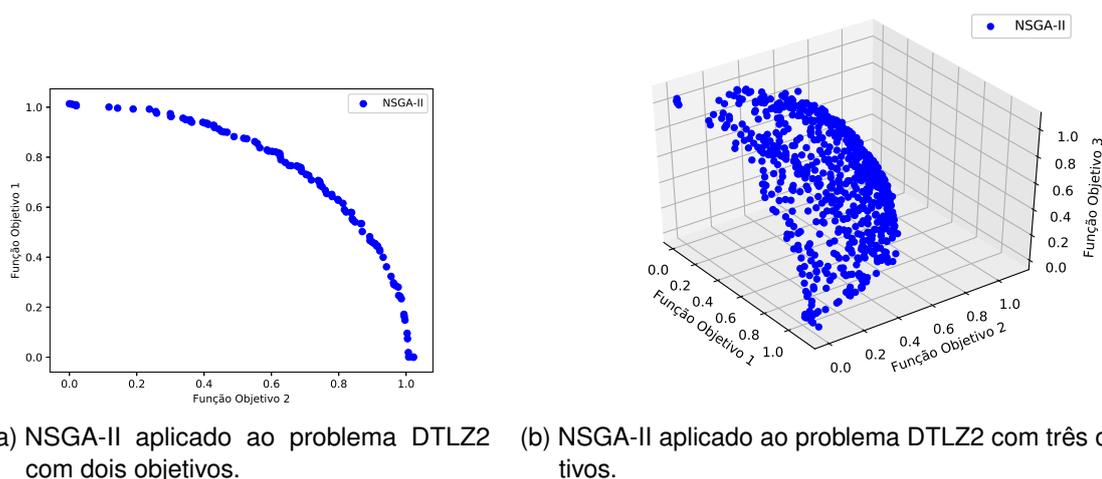


Figura 11 – Exemplos de aplicação do algoritmo NSGA-II para o problema DTLZ2. Resultados obtidos usando *MOEA Framework* [Hadka \(2017\)](#), versão 2.12.

Na Figura 11a o conjunto Pareto combinado obtido apresenta 100 soluções não dominadas enquanto que na Figura 11b o conjunto Pareto combinado obtido apresenta 730 soluções não dominadas. Considerando o exemplo anterior, pode-se perceber que a relação encontrada em [Deb e Saxena \(2005\)](#) foi verificada. Vale ressaltar que não foi feito nenhum experimento para determinar a quantidade ideal de soluções para se ter uma boa aproximação do conjunto Pareto-ótimo em uma dimensão para o exemplo anterior.

3.3.4 Dificuldade de Visualização das Soluções

Uma forma de comparar soluções é por meio da análise visual. Porém, à medida que a dimensão cresce, torna-se mais difícil ter a visualização e principalmente uma interpretação da solução. Visualizar pontos em dimensões maiores que \mathbb{R}^3 não é algo trivial [Jaimes, Coello e Barrientos \(2009\)](#).

Alguns autores buscam contornar essa situação de diversos modos. No trabalho de [Wegman \(1990\)](#), é proposto o gráfico de coordenadas paralelas, em que o eixo y representa as dimensões de um ponto no espaço e o eixo x os valores de cada coordenada. Uma solução é representada por um linha conectando os respectivos valores de cada coordenada.

Em [He e Yen \(2016\)](#) é apresentado um método de visualização que mapeia as soluções de um espaço de dimensão elevada em um gráfico polar de duas dimensões.

Nessa visualização é mantida a relação de dominância, localização e formato da fronteira Pareto, além de manter a distribuição dos indivíduos. Características também buscadas em [Ibrahim et al. \(2018\)](#). Paralelamente, em [He e Yen \(2016\)](#), os autores propõem uma métrica (chamada *polar metric*) para avaliar a convergência e diversidade.

Os gráficos de bolha em três dimensões podem representar soluções em \mathbb{R}^5 ao se utilizar alguns atributos adicionais como tamanho e cor das bolhas [Koochaksaraei et al. \(2017\)](#). A Figura 12 apresenta 2.500 soluções aleatórias do PRVTRD em um gráfico de bolhas. Foram consideradas as cinco primeiras funções objetivo descritas na Seção 2.4. As funções objetivo f_1 , f_2 e f_3 são representadas, respectivamente, nos eixos x , y e z . A função objetivo f_4 é representada no tamanho das bolhas e a função f_5 pode ser vista na barra de cores localizada na parte de baixo da figura.

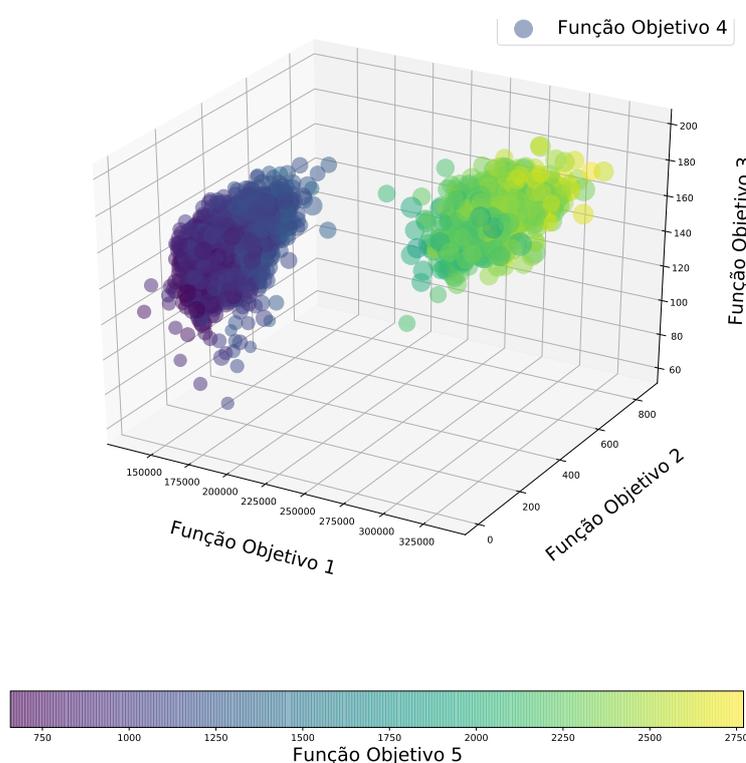


Figura 12 – Gráfico de bolhas apresentando visualização para soluções em \mathbb{R}^5 .

[Koochaksaraei et al. \(2017\)](#) propõem uma ferramenta chamada Diagrama de Cordas, uma outra forma de visualização de soluções em problemas com dimensões elevadas. Nela, os pontos são representados em um formato circular e cada coordenada é representada por um arco. Um exemplo do diagrama pode ser visto na Figura 13b

Mais recentemente, em [He, Yen e Ding \(2021\)](#) é desenvolvido um método para obter várias soluções de interesse de um grande número de soluções na fronteira Pareto. Além disso, a fim de reduzir as dificuldades no processo de tomada de decisão nos problemas com muitos objetivos, uma nova visualização em duas dimensões é desenvolvida, forne-

cendo informações sobre a forma e localização da fronteira Pareto, bem como o grau de convergência e distribuição das soluções.

Como forma de exemplificar a dificuldade de se visualizar a relação entre as funções objetivo, a Figura 13 apresenta gráficos com as soluções para um problema em \mathbb{R}^{12} . Fica claro que, em problemas com dimensões mais elevadas, há uma dificuldade em se entender as relações entre as coordenadas de cada solução na Figura 13a. Apenas de ser possível extrair as muitas informações na Figura 13b, pode ser visto apenas a distribuição dos valores dentro de cada arco.

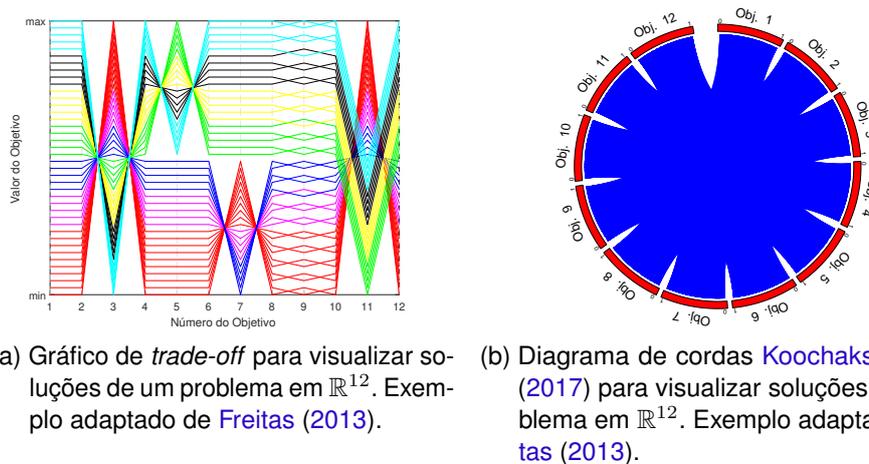


Figura 13 – Exemplos de ferramentas para visualização de soluções em \mathbb{R}^{12} .

3.3.5 Custo Computacional dos Indicadores de Performance

Outra dificuldade encontrada ao se resolver os problemas com muitos objetivos é o alto custo computacional necessário para medir o desempenho de um algoritmo. Para muitas métricas de desempenho, avaliar um conjunto aproximado de uma fronteira Pareto resultante de um algoritmo muitas vezes pode requerer um grande número de soluções para o problema em questão, caso o mesmo possua uma dimensão elevada Maltese, Ombuki-Berman e Engelbrecht (2018), Zhou, Wang e Zhang (2020).

Como consequência, torna-se praticamente infactível obter o valor de algumas métricas sobre um certo número de objetivos, como, por exemplo, o hipervolume Li e Yao (2019). O cálculo do hipervolume exato para um conjunto de soluções possui relação exponencial em relação ao número de objetivos While et al. (2006).

Em problemas com mais de quatro objetivos, o cálculo do hipervolume é frequentemente estimado por meio de uma técnica de amostragem de Monte Carlo Bader e Zitzler (2011). Uma alternativa para a amostragem seria a divisão do conjunto de soluções de não dominadas em várias frações, calculando-se o valor do hipervolume de cada fração

separadamente, e depois, somando o valor de cada fração para obter o valor final para o conjunto todo [Watanabe, Tatsukawa e Oyama \(2015\)](#).

Uma outra forma de se avaliar os algoritmos é por meio de métricas como IGD [Coello e Sierra \(2004\)](#), IGD⁺ [Ishibuchi et al. \(2015\)](#) que medem a distância entre as soluções de um conjunto não dominado e o conjunto Pareto para um problema. Porém, para essas métricas serem calculadas, é necessário ter o conjunto Pareto real do problema. Já nos problemas em que não se tem o conjunto Pareto real, não é possível utilizar métricas que necessitem desse conjunto. Especificamente no caso do cálculo do hipervolume, não é necessário conhecimento prévio da fronteira Pareto.

3.4 Principais Classes de Algoritmos

De acordo com [Li et al. \(2015\)](#), vários algoritmos evolutivos multiobjetivo (MOEA) tentam lidar com o problema de muitos objetivos de diferentes maneiras, usando: relação de dominância relaxada, pontos de referência, agregação, preferências, indicador e redução de dimensionalidade. Enquanto os MOEA de última geração perdem sua eficácia ao trabalhar em um espaço objetivos de dimensão alta, funciona muito bem em um espaço objetivos de baixa dimensão.

- Baseados em Dominância Relaxada - A dominância relaxada é uma forma de se diferenciar soluções não dominadas e assim, conseguir o aumento da pressão seletiva. Essa relaxação é uma forma de contornar o aumento das regiões não dominadas, apresentadas na Seção 3.3.1. Existem duas formas de realizar a relaxação na dominância:
 - Relaxação baseada no número de funções objetivo dominantes: busca comparar duas soluções usando a contagem do número de objetivos em que uma solução é maior, igual ou menor que a outra;
 - Relaxação baseada no valor das funções objetivo: ocorre uma modificação na dominância (Definição 1) em que os valores de cada objetivo são alterados. Esse tipo de alteração busca aumentar a região de soluções dominadas para as soluções não dominadas.
- Baseados em Diversidade - São algoritmos que utilizam mecanismos para controle da diversidade. Como vários trabalhos encontrados na literatura buscam aumentar a pressão seletiva dos algoritmos na direção da fronteira Pareto, a pressão de seleção pode ser muito alta, e nesses casos, a população se concentra em uma região específica do espaço [Ishibuchi et al. \(2011\)](#).
- Baseados em Indicadores - Os algoritmos avaliam os conjuntos de soluções obtidos com indicadores que guiam o processo de busca. Um exemplo de indicador que pode ser utilizados é o hipervolume. O algoritmo HypE, encontrado em [Bader e](#)

Zitzler (2011), utiliza um procedimento para obter uma aproximação para o cálculo do hipervolume.

- Baseados em Pontos de Referência - Os algoritmos baseados em pontos de referência utilizam um conjunto de soluções, tidas como referência, para se avaliar a qualidade de outras soluções. Alguns das dificuldades encontradas nesses métodos está em definir o conjunto de soluções de referência e a forma de medir a qualidade das soluções em relação a esse conjunto. Um exemplo de algoritmo que utiliza pontos de referência para guiar a busca é o algoritmo NSGA-III proposto em Deb e Jain (2014).
- Baseados em Preferência - São algoritmos que buscam reduzir o número de soluções da fronteira Pareto utilizando preferências dos usuários. Um dos critérios importantes a ser definido nesse tipo de algoritmo é quando se deve inserir a preferência. Esse processo pode ser feito antes, durante ou após o processo de otimização. O trabalho de Goulart e Campelo (2016) apresenta um método para definição de uma região de interesse (também conhecida como ROI - *Region Of Interest*) que é obtida com base na preferência do usuário. Esse método é feito durante o processo de otimização e não precisa de definir nenhum parâmetro extra, podendo ser inserido em qualquer algoritmo evolutivo.
- Baseados em Agregação - Os algoritmos utilizam funções de agregação que podem ser baseadas no indivíduo, como a uma soma ponderada, ou por comparações por pares, como valores máximos e mínimos. Um exemplo de algoritmo que usa agregação é o MOEA/D, o qual realiza a decomposição do problema em vários subproblemas de otimização escalar, realizando as agregações, e resolve os subproblemas simultaneamente. Outros exemplos de algoritmos podem ser encontrados em Li et al. (2015).
- Baseados em Redução de Dimensionalidade - São algoritmos que buscam resolver os problemas com muitos objetivos reduzindo o número de funções do problema, seja pela seleção de um conjunto representativo de funções ou pela soma de funções ditas redundantes. O trabalho de Jaimes, Coello e Barrientos (2009) apresenta o algoritmo KOSSA que busca reduzir a dimensionalidade do problema durante o processo de otimização. Esta tese tem o foco nessa classe de algoritmos.

4 Algoritmo de Redução de Dimensionalidade *Online*

Neste capítulo serão apresentadas outras definições utilizadas durante a execução do trabalho bem como o algoritmo utilizado para a realização de todos os testes. Inicialmente será apresentada a definição de conflito entre funções, análise de aglomerados (mais conhecida como análise de *cluster*) para redução de objetivos e as técnicas de seleção de atributos. Ao final do capítulo é apresentado o algoritmo de redução de dimensionalidade *online* e seus detalhes juntamente com os operadores genéticos utilizados para o PRVTRD.

4.1 Conflito entre Funções

Uma das importantes características dos problemas de otimização multiobjetivo é a presença de conflito entre os objetivos do problema em análise [Deb \(2008\)](#). Caso não existisse nenhuma forma de conflito, os objetivos poderiam ser otimizados separadamente. Uma definição formal do conceito de conflito pode ser encontrada em [Carlsson e Fuller \(1996\)](#) e é apresentada a seguir:

Definição 5 (Conflito). *Sejam f_i e f_j duas funções objetivo nas Equações 13 e X o espaço de variáveis, diz-se que:*

- *f_i suporta f_j em X (denotado por $f_i \uparrow f_j$) se $f_i(x') \geq f_i(x)$ implica $f_j(x') \geq f_j(x)$, para todo $x', x \in X$;*
- *f_i possui conflito com f_j em X (denotado por $f_i \downarrow f_j$) se $f_i(x') \geq f_i(x)$ implica $f_j(x') \leq f_j(x)$, para todo $x', x \in X$.*

A Figura 14 exemplifica a Definição 5 apresentada anteriormente. Entre as funções f_1 e f_2 percebe-se que a função f_1 suporta f_2 enquanto que há conflito entre as funções f_3 e f_4 . Segundo [Jaimes \(2011\)](#), em diversos problemas de otimização multiobjetivo as relações de conflito e suporte entre duas funções pode se alterar dependendo do domínio das funções.

No trabalho de [Freitas, Fleming e Guimarães \(2015\)](#) pode ser visto uma forma de calcular analiticamente a localidade do conflito entre duas funções. Descobrir a região onde há maiores valores de conflito trás informações importantes do ponto de vista da otimização. Dada duas funções, se o conflito é localizado em valores ruins, pode-se dizer que ocorre a inexistência de conflito inerente entre as funções [Freitas \(2013\)](#). Uma outra forma de interpretar o conceito de conflito é se pensar que a melhoria (aumento ou diminuição,

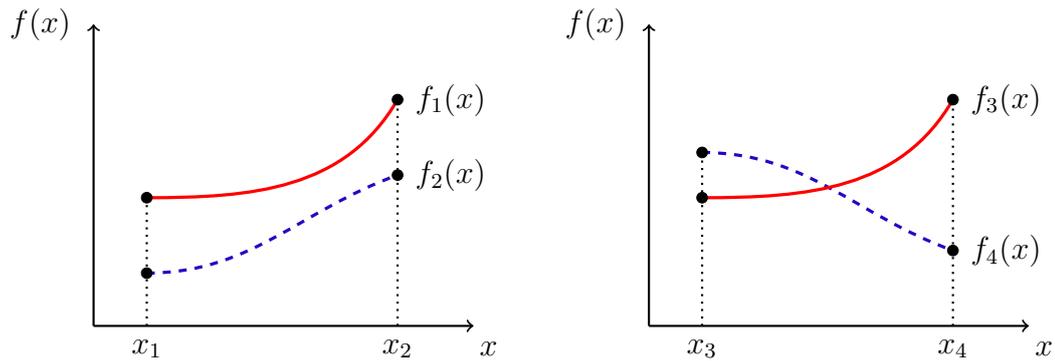


Figura 14 – Exemplo de conflito e suporte entre funções. A funções $f_1(x)$ suporta $f_2(x)$ no intervalo de x_1 até x_2 . funções $f_3(x)$ e $f_4(x)$ apresentam conflito no intervalo de x_3 até x_4 . Adaptado de [Jaimes \(2011\)](#).

dependendo do problema) dos valores de uma função implicam na piora (respectivamente diminuição ou aumento) dos valores da outra função.

4.2 Análise de *Cluster*

Nesta seção será apresentada a principal técnica utilizada neste trabalho para a redução de objetivos: Análise de *Cluster* (*Clustering*). Os principais conceitos, definições e classificações serão mostrados a seguir.

4.2.1 Notação Utilizada

Para melhor entendimento dos conceitos desta seção, a seguinte notação, encontrada em [Mingoti \(2005\)](#), será introduzida: considera-se a existência de um conjunto de dados com n elementos e em cada um há p -variáveis aleatórias. Um vetor de medidas X_j é dado para cada elemento da amostra j e pode ser definido como:

$$X_j = [x_{1j}, x_{2j} \cdots x_{pj}]', \quad j = 1, 2, \dots, n \quad (14)$$

no qual x_{ij} representa o valor da i -ésima variável no j -ésimo elemento. A seguir serão apresentados os principais conceitos utilizados durante o processo de análise de aglomerados.

4.2.2 Principais Conceitos

A técnica conhecida como análise de *cluster* permite decompor um conjunto de dados em grupos significativos. Se o objetivo é obter grupos significativos, o *cluster* deve ser capaz de manter a estrutura natural dos dados [Tan, Steinbach e Kumar \(2005\)](#). Existem várias aplicações da análise de *cluster* em diferentes áreas, como pesquisa de mercado, educação, mineração de dados, seguro de vida e aprendizado de máquina.

A análise de *cluster* usa o princípio de que objetos pertencentes ao mesmo grupo têm maior similaridade e valores mais baixos em relação a outros grupos. Para membros de um grupo, quanto maior a sua similaridade, maior a diferença entre os grupos [Tan, Steinbach e Kumar \(2005\)](#). A semelhança é uma medida de correspondência ou afinidade entre os objetos a serem agrupados. A seguir serão descritas algumas das principais medidas de similaridade usadas na literatura.

4.2.3 Medidas de Similaridade

Uma das medidas possíveis usadas para calcular a similaridade é a distância, por exemplo, as distâncias de Manhattan, Euclidiana ou Minkowski [Soler et al. \(2013\)](#). Medidas de correlação e associação também são outras maneiras de calcular a similaridade entre objetos [Balcan, Blum e Vempala \(2008\)](#). A seguir, serão apresentadas as definições das distâncias citadas porém este trabalho se limita ao uso das medidas de correlação.

4.2.3.1 Distância de Manhattan

A distância de Manhattan entre dois elementos X_l e X_k , $l \neq k$, é definida por:

$$d(X_l, X_k) = \left[\sum_{i=1}^p |X_{il} - X_{ik}| \right] \quad (15)$$

4.2.3.2 Distância Euclidiana

A distância Euclidiana entre dois elementos X_l e X_k , $l \neq k$, é definida por:

$$d(X_l, X_k) = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{\frac{1}{2}} \quad (16)$$

4.2.3.3 Distância de Minkowsky

A distância de Minkowsky entre dois elementos X_l e X_k , $l \neq k$, é definida por:

$$d(X_l, X_k) = \left[\sum_{i=1}^p w_i |X_{il} - X_{ik}|^\lambda \right]^{\frac{1}{\lambda}} \quad (17)$$

em que w_i é o peso para se ponderar as variáveis. Fazendo $\lambda = 1$, temos que a distância de Manhattan, definida em [4.2.3.1](#). Caso $\lambda = 2$, a distância calculada é a Euclidiana ([4.2.3.2](#)). A distância de Minkowsky é normalmente utilizada quando se pretende reduzir o efeito de *outliers* na amostra.

4.2.4 Coeficiente de Correlação

Um coeficiente de correlação pode ser considerado como forma de se mensurar a relação entre duas variáveis, buscando compreender o comportamento de uma variável

quando há variação na outra (pensando no caso da correlação entre duas variáveis). Apesar do coeficiente de correlação linear ser amplamente utilizado, ele não é a única forma de se calcular as relações entre duas variáveis. Dependendo dos dados, um coeficiente de pode ser mais adequado que outro. A seguir, serão apresentados alguns dos principais coeficientes de correlação.

4.2.4.1 Coeficiente de Correlação de Pearson

O coeficiente de correlação de Pearson, simbolizado pela letra ρ mede o grau de relação linear entre duas variáveis. É também chamado de coeficiente de correlação linear e seu valor pertence ao intervalo $[-1, +1]$.

Quando mais próximo de +1 for o valor do coeficiente, significa que o aumento nos valores de uma variável implicam o aumento nos valores da outra e neste caso, há a existência de uma relação linear positiva forte. No caso do valor do coeficiente ser próximo de -1, pode-se dizer que há também uma relação forte porém, essa relação é negativa. Um outro caso acontece quando o coeficiente possui valor próximo de zero, diz-se então que não há relação entre as variáveis.

O coeficiente de correlação de Pearson, ρ_{ij} , entre a i -ésima e j -ésima variável do vetor X é definido como:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i\sigma_j} \quad (18)$$

em que σ_{ij} é a covariância entre a i -ésima e a j -ésima variável do vetor X , σ_i e σ_{ii} representam, respectivamente, o desvio padrão e a variância do i -ésimo componente do vetor X .

A covariância entre os valores da i -ésima e a j -ésima variável do vetor X é definida como:

$$Cov(X_i, X_j) = \sigma_{ij} = E(X_i - \mu_i)(X_j - \mu_j) \quad (19)$$

em que $\mu_i = E(X_i)$ é a média do vetor X , também chamada de esperança.

Uma outra definição importante a ser apresentada é a Matriz de Correlação $P_{p \times p}$, do vetor aleatório X , sendo denotada da seguinte forma:

$$P_{p \times p} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2p} \\ \rho_{31} & \rho_{32} & 1 & \cdots & \rho_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \rho_{p3} & \cdots & 1 \end{bmatrix} \quad (20)$$

em que cada elemento ρ_{ij} denota a correlação entre o i -ésimo e o j -ésimo componente do vetor X .

4.2.4.2 Coeficiente de Correlação de Kendall

Sabe-se que o coeficiente de correlação de Pearson é uma medida linear da distância entre dois vetores e uma medida linear pode não ser a melhor maneira de mensurar a similaridade entre funções em um problema de muitos objetivos. Uma outra forma para se medir a similaridade, porém de maneira não-paramétrica, é com o uso do Coeficiente de Correlação de Kendall, também conhecido como τ de Kendall [Kendall \(1938\)](#).

O τ de Kendall é uma estatística usada para medir a associação ordinal entre duas grandezas medidas. É uma medida de correlação de ranqueamento: a similaridade dos ordenamentos dos dados quando classificados por cada uma das quantidades [Abdi \(2007\)](#). O valor é calculado da seguinte forma:

$$\tau(\mathcal{P}_1, \mathcal{P}_2) = 1 - \frac{2 \times d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)}{N(N-1)} \quad (21)$$

em que N é o número de objetos e $d_{\Delta}(\mathcal{P}_1, \mathcal{P}_2)$ é a distância simétrica da diferença entre dois conjuntos de pares ordenados \mathcal{P}_1 e \mathcal{P}_2 .

A normalização da diferença simétrica resulta em τ de Kendall. Essa correlação levará valores entre -1 e $+1$ e a interpretação é a mesma da correlação de Pearson, isto é, -1 representa a distância máxima (relação inversa) e $+1$ representa a distância mínima (relação direta). A análise de *cluster* avalia a distância (ou similaridade) entre dois vetores e, desta forma, a seguinte transformação linear é feita: $d(\mathcal{P}_1, \mathcal{P}_2) = 1 - \tau(\mathcal{P}_1, \mathcal{P}_2)$.

4.2.5 Classificação de um *Cluster*

Existem diversas formas de se fazer um *cluster*, alguns deles podem ser encontrados em [Mingoti \(2005\)](#) e são apresentados à seguir:

- Hierárquico Aglomerativo - Começa o processo de agrupamento fazendo com que o número de *clusters* seja igual o número de elementos do conjunto e em cada iteração agrega-se os os elementos desse conjunto, formando um novo *cluster*. O processo termina quando todos os elementos são agrupados em um único *cluster*.
- Hierárquico Divisivo - Ocorre o contrário do *cluster* hierárquico aglomerativo: começa o processo com um grande conjunto e vai quebrando-o em partes ate chegar a elementos isolados. No final, tem-se que o número de *clusters* é igual ao número de elementos do conjunto.
- K-Means - É um processo não hierárquico de agrupamento em que cada elemento é agrupado ao *cluster* com centróide mais próximo. O centróide é calculado como a média do agrupamento.

4.2.6 Construção de um *Cluster*

Este trabalho usa métodos de *cluster* hierárquico aglomerativo, que consistem em criar uma hierarquia de relacionamentos entre objetos e trabalhar criando conjuntos a partir de elementos isolados. Além disso, eles não possuem nenhuma justificativa teórica baseada em estatística ou teoria da informação [Balcan, Blum e Vempala \(2008\)](#).

Existem diferentes maneiras de medir a distância entre dois *clusters*, tais como: ligação única, ligação completa, ligação média, centroide e *ward* [Tan, Steinbach e Kumar \(2005\)](#). Este trabalho adotou o método de ligação única, no qual agrupa os elementos com maior similaridade. O Algoritmo 2 apresenta o algoritmo de agrupamento hierárquico aglomerativo.

Algoritmo 2: *Cluster* hierárquico aglomerativo

Input: matriz de pontos X
Output: conjunto de *clusters* C
 Inicializa iteração corrente
 Inicializa conjunto de *clusters*
while existir mais de um *cluster* **do**
 Encontrar os *clusters* mais próximos
 Agrupar os *clusters*
 Atualizar valor mínimo dos *clusters* agrupados
 Atualizar iteração corrente
end

4.2.7 Métodos para Calcular a Distância entre *Clusters*

- Ligação única: a distância entre dois *clusters* p e q é definida como a distância entre os pontos mais próximos. Pode ser calculada da seguinte forma:

$$d(C_p, C_q) = \min d(x_i \in C_p, x_j \in C_q) \quad (22)$$

- Ligação completa: a distância entre dois *clusters* p e q é definida como a distância entre os pontos mais distantes. Pode ser calculada da seguinte forma:

$$d(C_p, C_q) = \max d(x_i \in C_p, x_j \in C_q) \quad (23)$$

- Ligação média: a distância é calculada como a média das distâncias entre todos os pares de observações:

$$d(C_p, C_q) = \frac{1}{N_p N_q} \sum_{i=1}^{N_p} \sum_{j=1}^{N_q} d(x_i \in C_p, x_j \in C_q) \quad (24)$$

em que N_p e N_q são respectivamente o número de observações de C_p e C_q .

- Centroide: a distância entre os *clusters* é definida como a distância entre os centroides:

$$d(C_p, C_q) = d(\bar{c}_p, \bar{c}_q) \quad (25)$$

- *Ward*: a distância é calculada aplicando-se um peso nos valores de distância. O peso é referente ao número de observações. O cálculo é feito da seguinte forma:

$$d(C_p, C_q) = \frac{N_p N_q}{N_p + N_q} d(\bar{c}_p, \bar{c}_q) \quad (26)$$

4.2.8 Dendrograma

O dendrograma é uma ferramenta visual em formato de árvore que mostra os *clusters* formados por agrupamento de observações em cada iteração do algoritmo em relação à similaridade. A similaridade é medida no eixo vertical e as observações agrupadas são apresentadas no eixo horizontal.

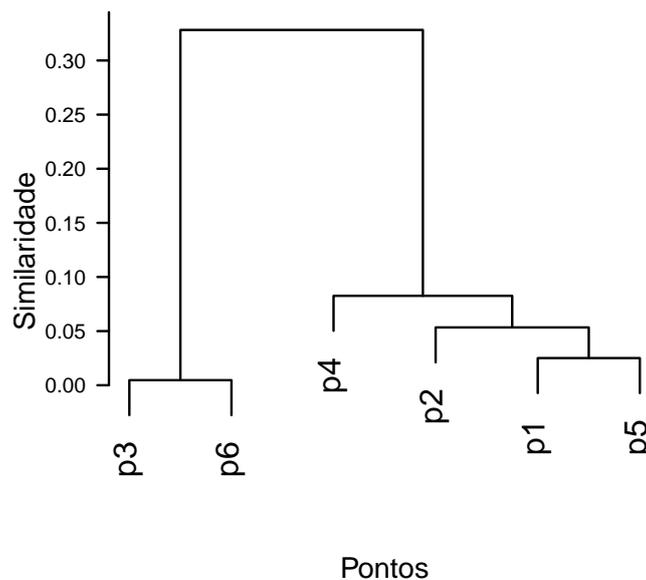


Figura 15 – Dendrograma para um conjunto de dados com nove pontos distintos usado o método de ligação única.

A Figura 15 apresenta um exemplo de dendrograma com mostra seis pontos que são agrupados hierarquicamente até que um *cluster* com todos os pontos seja obtidos. A interpretação e forma de analisá-lo serão descritos a seguir.

4.2.9 Interpretação de um Dendrograma

O dendrograma deve ser usado para se visualizar a formação de *clusters* a cada iteração e, em paralelo, identificar o valor da similaridade. O nível de similaridade entre

dois elementos (pontos ou *clusters*) pode ser encontrado por meio de uma linha traçada horizontalmente. A forma como os valores de similaridade se alteram à medida que caminha-se pelo dendograma pode fornecer uma visão do agrupamento final dos dados.

Para se obter o agrupamento final é feito um corte no dendograma. Esse corte é feito com o desenho de uma linha horizontal percorrendo todo dendograma. Podem ser usadas diversas formas de se calcular a similaridade entre os dados e, dessa forma, os dendogramas podem chegar a agrupamentos finais distintos. A escolha do dendograma é feita de modo que seja escolhido aquele que faz mais sentido para o problema em questão.

4.3 Redução de Dimensionalidade Usando *Clustering*

Conforme levantado na Seção 3.3, lidar com um problema que possua muitos objetivos não é uma tarefa fácil. Diversas dificuldades podem surgir na busca por soluções como por exemplo redução do desempenho de algoritmos, dificuldade de visualização das soluções, aumento no número de soluções do conjunto Pareto e o aumento no custo computacional dos indicadores de performance.

Uma das formas de se contornar tais dificuldades é fazendo a redução do número de funções objetivo. Algumas das principais técnicas utilizadas são: seleção de atributos Lopez, Coello e Chakraborty (2008), análise de componentes principais Deb e Saxena (2005) e análise de cluster Denysiuk, Costa e Santo (2014). De acordo com Li et al. (2015), essas técnicas podem ser divididas em duas classes, métodos *offline* e *online*.

No contexto abordado neste trabalho, os métodos chamados de *offline* reduzem a dimensionalidade do problema antes da execução de um algoritmo evolutivo. Já os métodos *online* reduzem a dimensionalidade do problema gradualmente a cada iteração do algoritmo. Podem ser encontrados vários trabalhos na literatura que reduz o tamanho dos problemas de otimização com muitos objetivos usando a redução *offline* como *online*. Em Jaimes, Coello e Barrientos (2009) o algoritmo KOSSA e duas estratégias para reduzir a dimensão do problema dentro do ciclo evolutivo são apresentadas. No entanto, quando a dimensão do problema é reduzida, somente o conjunto mais representativo de funções é levado em consideração e as funções com o maior conflito são removidas.

Outra redução *online* é feita em Guo et al. (2012), onde a correlação entre os objetivos é calculada com base no conceito de Informação Mútua como uma nova maneira de medir a relação entre os objetivos. Os autores propõem o algoritmo MICA-NORMOEA que usa o *Partitioning Around Medoids* (PAM) para reduzir o tamanho do problema removendo os objetivos mais semelhantes ou redundantes. Em Guo, Wang e Wang (2013), o conceito de Coeficiente de Interdependência é inserido como uma forma de melhoria na medição das relações entre os objetivos.

Em [Denysiuk, Costa e Santo \(2014\)](#), a aplicação da técnica de *clustering* é feita de forma diferente. Nesse caso, a técnica é usada no algoritmo EMyO/C para agrupar soluções em espaços de alta dimensão, de modo que as soluções mais distantes do espaço original sejam consideradas o grupo mais representativo. Um ponto de referência é usado durante a busca e a distância é minimizada até esse ponto.

Alguns trabalhos encontrados na literatura aplicam o *cluster* hierárquico *online* para obter um grupo representativo de funções, eliminando aquelas que são redundantes [Jaimes, Coello e Barrientos \(2009\)](#), [Guo et al. \(2012\)](#), [Guo, Wang e Wang \(2013\)](#), [Breaban e Iftene \(2015\)](#). Nestes trabalhos, a seleção do número de funções é feita com base em um procedimento padrão não supervisionado. O algoritmo NSGA-II é usado para testar os agrupamentos obtidos.

A integração do método de *clustering* do PAM no algoritmo IDEMO, que é o *Improved Differential Algorithm* para Otimização Multiobjetivo, é feita em [Pal, Saha e Bandyopadhyay \(2018\)](#). Esta integração resulta no algoritmo DECOR: Evolução Diferencial usando Redução de *Cluster* para otimização de muitos objetivos. Embora o algoritmo proposto utilize o coeficiente de correlação de Pearson, é feita a alternância entre os tamanhos dos problemas e, assim, permitindo uma busca global e local. Vale ressaltar que em todos os estudos citados, metodologias e algoritmos foram aplicados com base nos problemas encontrados na literatura e não em problemas reais ou combinatórios.

Embora o principal objetivo das técnicas de redução da dimensionalidade seja identificar e eliminar os objetivos não conflitantes buscando obter a fronteira Pareto original [Lopez, Coello e Chakraborty \(2008\)](#), o método de redução de dimensionalidade proposto nesta tese é feito através da identificação, por meio da análise hierárquica de *cluster* dos objetivos não conflitantes e em seguida é feita sua agregação, sem eliminar nenhuma função objetivo. A similaridade é calculada com base nos coeficientes de correlação de Pearson e τ de Kendall. A cada geração do algoritmo, o problema original é reduzido a uma formulação bi-objetivo, e em seguida, os operadores evolutivos são aplicados, buscando encontrar eficientemente a fronteira Pareto do problema.

Abordagens semelhantes podem ser encontradas em [Mendes et al. \(2016\)](#), [Mendes et al. \(2016\)](#), [Mendes et al. \(2017b\)](#). Em todos os casos, os autores usam um grande conjunto de soluções geradas *a priori* e a redução é realizada antes da execução do algoritmo, evitando a eliminação dos objetivos não conflitantes. Nesses trabalhos citados, é utilizado um conjunto de soluções geradas aleatoriamente, como forma de se fazer uma amostragem do espaço de busca. Além disso, a mesma formulação reduzida do problema é aplicada em todo o processo. A ideia proposta nesta tese é dar a liberdade para que a redução seja feita de acordo com a similaridade atual no espaço objetivo ao longo da busca.

4.4 Técnicas de Seleção de Atributos

A seleção de atributos, do inglês *Feature Selection* (FS), [Min et al. \(2020\)](#) é conhecida como o método para escolha dos atributos mais relevantes em problemas de classificação, regressão e de agrupamento. FS é aplicada para reduzir a dimensionalidade dos dados, facilitando assim sua compreensão e visualização [Pal e Mitra \(2004\)](#).

As técnicas de seleção de atributos podem ser classificadas em três categorias: métodos de filtro, *wrapper* e os intrínsecos [Pal e Mitra \(2004\)](#). Métodos de filtro usam algumas métricas com base nas características dos atributos para estabelecer quais atributos têm mais peso no conjunto de dados. Esses métodos, avaliando independentemente cada atributo sem considerar a interação entre eles, podem eliminar atributos irrelevantes e com ruído. No entanto, seu desempenho é reduzido, uma vez que os métodos de filtro não são sensíveis a atributos redundantes. Por outro lado, aplicando um algoritmo de aprendizagem, no processo de treinamento os métodos de *wrapper* obtêm um nível mais alto desempenho enquanto exige um custo computacional maior. Já os métodos intrínsecos combinam as qualidades dos métodos de filtro e *wrapper* para criar o melhor subconjunto de atributos [Honest \(2020\)](#).

Este trabalho tem o foco na seleção de atributos não-supervisionados, ou também conhecidos como *unsupervised feature selection* (UFS), uma vez que não há rótulos para os dados em análise [Hastie, Tibshirani e Friedman \(2001\)](#). As técnicas não supervisionadas de seleção de atributos serão usadas para selecionar os atributos (no caso deste trabalho, as funções objetivo) de acordo com a estrutura de informação local [Min et al. \(2020\)](#).

Além disso, a análise também se limita aos métodos de filtro evitando a sobrecarga de computação dos métodos baseados em *wrapper*. Observe que as técnicas UFS serão aplicadas dentro o ciclo evolutivo do MOEA/D para escolher o *cluster* representativo na redução da dimensionalidade. Os métodos descritos abaixo foram escolhidos, considerando que a sobrecarga do cálculo incorrerá no tempo de execução geral do algoritmo. A escolha do método de seleção foi baseada no fato de (i) ser intuitivo, (ii) fácil de se reproduzir e (iii) ter um curto tempo de execução. Os dois exemplos de métodos de UFS são descritos a seguir.

4.4.1 Taxa de Dispersão

Dado um atributo/objetivo m , a taxa de dispersão (DR_m) pode ser definida como a proporção entre as médias aritmética e geométrica (AM_m e GM_m) dos valores do atributo/objetivo. Vale ressaltar que a forma utilizada para o cálculo da taxa de dispersão neste trabalho é uma simplificação do cálculo encontrado em [Liu, Chen e Wu \(2018\)](#), [Roy](#)

et al. (2019). A taxa de dispersão é calculada da seguinte forma:

$$DR_m(P) = \frac{AM_m(P)}{GM_m(P)}, \quad (27)$$

em que AM_m é obtido como:

$$AM_m(P) = \bar{F}_m = \frac{1}{N} \sum_{n=1}^N F_m^n, \quad (28)$$

e GM_m é calculado da seguinte forma:

$$GM_m(P) = \left(\prod_{n=1}^N F_m^n \right)^{\frac{1}{N}}, \quad (29)$$

em que $P = \{S_1, S_2, \dots, S_N\}$, é um conjunto de N soluções, \bar{F}_m é a média dos valores do atributo/objetivo m , F_m^n é o valor do atributo/objetivo m na observação/solução n . Dado um conjunto de atributos/objetivos, a taxa de dispersão de cada atributo/objetivo é calculada e aquela com maior valor é escolhida como o representante do cluster.

4.4.2 Máxima Variância

Uma das maneiras mais simples de usar a seleção de atributos não-supervisionado é por meio de Máxima Variância Dy e Brodley (2004). Dado um conjunto de atributos, a variância de cada atributo é calculada e aquele com o valor mais alto é escolhido como o representante do conjunto. O valor da variância é calculado como:

$$Var_m(P) = \frac{\sum_{n=1}^N (F_m^n - \bar{F}_m)^2}{n - 1}, \quad (30)$$

em que $Var_m(P)$ é a variância do atributo/objetivo m no conjunto de N soluções P , \bar{F}_m é a média dos valores do atributo/objetivo m , F_m^n é o valor do atributo/objetivo m na observação/solução n .

Na seção a seguir serão apresentados dois exemplos do uso dessas técnicas não-supervisionadas de seleção de atributos. Os exemplos foram feitos considerando um conjunto de dados previamente agrupado por meio do algoritmo hierárquico aglomerativo.

4.4.3 Exemplos de Aplicação das Técnicas de Seleção de Atributos para Escolha de Objetivos

Como exemplo, foram geradas 2.500 soluções aleatórias e todas foram avaliadas na formulação original com muitos objetivos. Essas soluções foram utilizadas como entrada para a redução da dimensionalidade. Foi aplicado o algoritmo de agrupamento hierárquico às soluções aleatórias usando o coeficiente de correlação de Pearson e foram obtidos dois

clusters. A Figura 16 mostra o dendrograma do *cluster* hierárquico usando o coeficiente de Pearson obtido usando 2.500 soluções aleatórias. Neste exemplo, dois *clusters* são obtidos:

$$\begin{aligned} \text{Cluster 1} &= \{f_1, f_2, f_6, f_7, f_8\} \\ \text{Cluster 2} &= \{f_3, f_4, f_5\} \end{aligned} \tag{31}$$

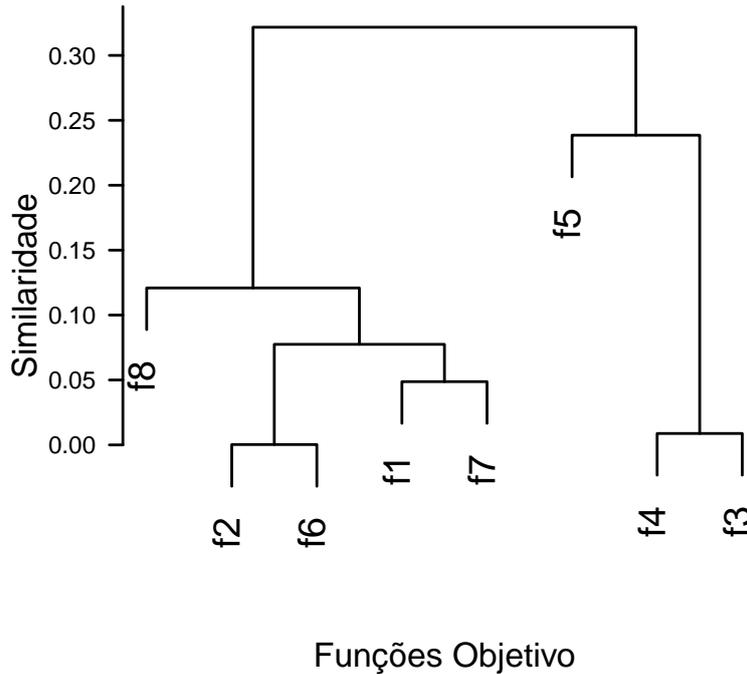


Figura 16 – Dendrograma do *cluster* hierárquico usando coeficiente de correlação de Pearson em 2.500 soluções aleatórias.

Para o exemplo dado anteriormente, usando os dois *clusters* obtidos na Equação 31, a formulação bi-objetivo para o problema é dada por:

$$\begin{aligned} \text{Min } F_1 &= \lambda_1 f_1 + \lambda_2 f_2 + \lambda_6 f_6 + \lambda_7 f_7 + \lambda_8 f_8, \\ \text{Min } F_2 &= \lambda_3 f_3 + \lambda_4 f_4 + \lambda_5 f_5, \end{aligned} \tag{32}$$

- Taxa de Dispersão: usando o mesmo exemplo dado anteriormente e levando em consideração os dois *clusters* obtidos pela Equação 31, a formulação bi-objetivo quando é utilizada a taxa de dispersão é dada por:

$$\begin{aligned} \text{Min } F_1(S) &= f_6, \\ \text{Min } F_2(S) &= f_3, \end{aligned} \tag{33}$$

em que F_1 é a função objetivo no *Cluster 1* que tem o maior valor da taxa de dispersão e F_2 é a função objetivo no *Cluster 2* com o maior valor da taxa de dispersão.

- Máxima Variância: como um exemplo de aplicação, ao se utilizar os *clusters* obtidos na Equação 31, a formulação bi-objetivo para o problema quando a máxima variância

é utilizada é dada pela Equação 33 substituindo os valores da taxa de dispersão pelos valores da máxima variância.

4.5 Redução de Dimensionalidade *Online*

A ideia proposta neste trabalho é dar liberdade para que a tarefa de redução de dimensionalidade seja feita de acordo com a similaridade atual no espaço objetivos ao longo do processo de busca. A redução da dimensionalidade é realizada a cada geração de um Algoritmo Evolutivo Multiobjetivo Genético (MOEA). Usando as informações disponíveis, sem nenhuma avaliação de função extra, o objetivo é transformar o problema de muitos objetivos em um bi-objetivo.

Para problemas de muitos objetivos, foi demonstrado que o Algoritmo Evolutivo Multiobjetivo baseado em Decomposição (MOEA/D) Zhang e Li (2007) funciona bem quando comparado com algoritmos baseados em indicadores e Pareto dominância Ishibuchi et al. (2009), Ishibuchi, Akedo e Nojima (2015). Neste trabalho, a metodologia proposta é acoplada ao MOEA/D. No entanto, uma vez que o objetivo principal é demonstrar a eficácia da técnica de redução de dimensionalidade *online* proposta, a metodologia pode ser facilmente adaptada a qualquer outro MOEA.

A estrutura da abordagem proposta, nomeada como On-MOEA/D, é a mesma apresentado no MOEA/D clássico. A diferença do On-MOEA/D em relação à sua versão original é que o primeiro apresenta a aplicação de redução de dimensionalidade dentro do ciclo evolutivo do algoritmo. Para detalhes adicionais do MOEA/D e suas variantes, consultar o trabalho original Zhang e Li (2007), Ishibuchi, Doi e Nojima (2017).

O pseudocódigo para MOEA/D com redução de dimensionalidade *online* (On-MOEA/D) é apresentado no Algoritmo 3. O algoritmo apresentado foi construído com base no trabalho original Zhang e Li (2007) e alguns detalhes devem ser pontuados em relação ao algoritmo:

- Os vetores de peso são inicializados de acordo com uma distribuição uniforme;
- A população externa (em alguns trabalhos também é chamada de arquivo) é criada com base nas soluções não dominadas;
- O conjunto $B(i)$ contém os índices dos T vetores mais próximos do vetor de pesos λ^i , com $i = 1, \dots, N$;

O Algoritmo 3 apresenta a adição da redução de dimensionalidade como o primeiro passo do ciclo evolutivo. Pode ser visto que $C(t + 1)$ recebe os indivíduos da população $P(t + 1)$ com as funções objetivo reduzidas. A agregação dos objetivos é feita durante o passo de atualização de soluções vizinhas ao final do laço de repetição mais interno. Vale ressaltar que a redução de dimensionalidade implementada neste trabalho pode ser

Algoritmo 3: On-MOEA/D

Input: População P
Output: População externa EP
 Inicializa vetores de peso $\lambda^i = \lambda^1, \dots, \lambda^N$
 Inicializa população externa $EP(t = 0) = \emptyset$
 Calcula a distância Euclidiana entre os vetores de peso
 Obtém os T vetores de peso mais próximos $B(i) = \{i_1, \dots, i_T\}$ para λ^i
 Inicializa a população $P(t = 0) \leftarrow \{p^{(1)}, \dots, p^{(\mu)}\}$
 Inicializa o ponto ideal $z(t = 0)$ com os menores valores de cada função objetivo
while critério de parada não for satisfeito **do**
 $C(t + 1) \leftarrow$ Redução de dimensionalidade de $P(t + 1)$
 for $i=1, \dots, N$ **do**
 $S(i) \leftarrow$ Seleção ($C(t + 1)$)
 $O(i) \leftarrow$ Cruzamento ($S(i)$)
 $M(i) \leftarrow$ Mutação ($O(i)$)
 $z(i, t + 1) \leftarrow$ Atualiza ponto ideal ($z(i, t), M(i)$)
 $B(i) \leftarrow$ Atualiza soluções vizinhas ($B(i), M(i)$)
 end
 $EP(t + 1) \leftarrow$ Atualiza população externa ($EP(t), z(i, t), C(t)$)
 $t \leftarrow t + 1$
end

facilmente adaptada em qualquer algoritmo evolutivo multi-objetivo.

4.5.1 Detalhes do Algoritmo de Redução de Dimensionalidade *Online* Proposto

O algoritmo de redução de objetivos *online* proposto é baseado na análise de *cluster* hierárquica e uma correlação paramétrica e não-paramétrica para medir a similaridade dos objetivos a serem agregados. Este trabalho difere dos encontrados na literatura no sentido de que todos os objetivos do problema são utilizados sem eliminação aplicado-se uma medida paramétrica e não-paramétrica para determinar a similaridade entre as funções objetivo.

O algoritmo de redução de dimensionalidade proposto é aplicado de modo que cada *cluster* seja criado a partir dos elementos isolados. A cada iteração do algoritmo, funções com maior similaridade, medidas por Pearson ou τ de Kendal, são agregadas, de forma que a dimensão do problema seja reduzida. O algoritmo é aplicado até que dois *clusters* sejam obtidos, o que equivale a uma formulação bi-objetivo para o problema.

A segunda forma utilizada neste trabalho para a obtenção da formulação bi-objetivo é por meio do uso das técnicas de seleção de atributos não-supervisionadas descritas na Seção 4.4. Neste caso, os *clusters* são obtidos da mesma forma como descrito anteriormente e uma das funções objetivo é escolhida como representante do *cluster*, em que a

seleção de atributos é utilizada para a escolha da função.

Um outro detalhe sobre a redução de dimensionalidade que deve ser ressaltado é o fato das funções objetivo apresentarem unidades (por exemplo: número de veículos e tempo de deslocamento) e ordens de grandeza distintas. Uma vez que essas funções são agregadas ao se aplicar a clusterização hierárquica, é necessário criar uma forma de colocar as funções em uma mesma escala e dar o mesmo grau de importância para cada uma. Os detalhes de como a transformação das funções é feita serão apresentados na próxima seção.

4.5.2 Geração dos Pesos para os Objetivos

Na ausência de qualquer preferência de um tomador de decisão, os pesos para a formulação bi-objetivo precisam ser escolhidos de forma que a mesma importância seja dada a cada função objetivo. Em vez de definir uma escolha de peso *ad-hoc*, o seguinte procedimento é adotado. Seja $nruns$ o número de execuções de cada algoritmo em cada instância. $nruns$ populações iniciais contendo 100 indivíduos são geradas aleatoriamente e combinadas totalizando $100 \times nruns$ soluções. Usando essas soluções, para garantir a comparabilidade de todas as funções objetivo, os valores máximo e mínimo de cada objetivo (\max_{f_i} e \min_{f_i} , respectivamente) são obtidos e uma transformação linear é aplicada a cada expressão de função objetivo,

$$\bar{f}_i(\mathbf{x}) = \frac{f_i(\mathbf{x}) - \min_{f_i}}{\max_{f_i} - \min_{f_i}}, \quad \forall i = 1, \dots, 8. \quad (34)$$

Essa transformação apenas dimensiona os valores para caber no intervalo $[0, 1]$. As funções $\bar{f}_i, \forall i = 1, \dots, 8$ são usadas no processo de otimização. Observe que apenas os valores iniciais da população são ajustado para $[0, 1]$. Após o primeiro ciclo evolutivo, os valores podem estar fora dessa faixa.

4.6 Operadores Genéticos

Os operadores genéticos utilizados neste trabalho serão apresentados a seguir. Eles permitem a diversidade da população e definem quais características obtidas nas gerações anteriores podem ser mantidas durante o processo evolutivo. Foram escolhidos de acordo com o trabalho de [Mendes \(2016\)](#), em que foi feito um estudo dos melhores operadores evolutivos para o algoritmo NSGA-II.

4.6.1 Operador de Seleção

O operador de seleção torna os indivíduos mais adequados encontrados na população para terem maior chances de sobreviver durante o processo evolutivo. Indivíduos

com menor aptidão também têm chance de serem selecionados, evitando o processo de convergência prematura do algoritmo. Neste trabalho, foi utilizado o operador de seleção por roleta. Este operador consiste em selecionar um indivíduo de acordo com uma probabilidade de sua aptidão em relação a toda a população.

4.6.2 Operador de Cruzamento

Neste trabalho foi utilizado como operador de cruzamento o cruzamento em dois pontos. Neste operador, dois diferentes pontos $(x_1; x_2)$ são selecionados aleatoriamente entre 0 e $\min(|P_1|; |P_2|)$ e o conteúdo entre cada um desses pontos é trocado entre os indivíduos, gerando dois novos, chamados de filhos. Um exemplo do operador de cruzamento pode ser visto na Figura 17.

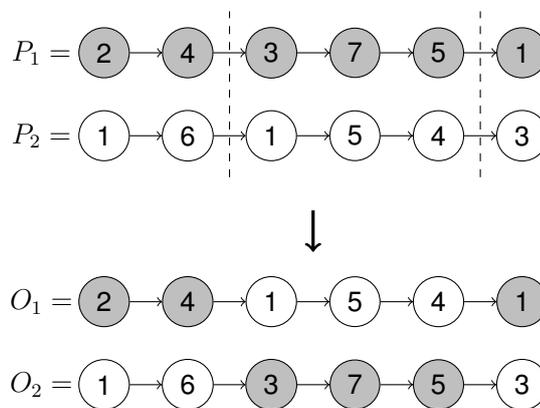
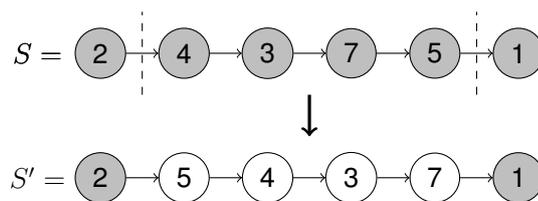


Figura 17 – Operador de cruzamento de dois pontos.

Os pontos $x_1 = 2$ e $x_2 = 5$ foram selecionados aleatoriamente. Os vértices $\{3, 7, 5\}$ de P_1 são trocados com os vértices $\{1, 5, 4\}$ de P_2 resultando em duas novas soluções, O_1 e O_2 .

4.6.3 Operador de Mutação

Como operador de mutação foi utilizado o operador *2-shuffle*, em que dois arcos não-adjacentes são selecionados e os vértices entre esses arcos são embaralhados. Dependendo de quais vértices são selecionados, o operador pode gerar soluções totalmente diferentes ou até mesmo muito próximas da solução original. A Figura 18 apresenta um exemplo do operador.

Figura 18 – Operador de mutação 2-*Shuffle*.

No exemplo do operador de mutação, os arcos (2, 4) e (5, 1) foram selecionados. A sequência de vértices 4, 3, 7, 5 é embaralhada, resultando em 5, 4, 3, 7 na nova solução.

5 Resultados Computacionais

A fim de avaliar a eficácia da abordagem proposta usando o redução de dimensionalidade *online* baseada em cluster com medida de correlação não-paramétrica, uma série de experimentos foram realizados. Os experimentos serão apresentados em duas partes. Na primeira parte, será apresentada a comparação entre o MOEA/D em relação às versões *online* e *offline*. Na segunda parte, os algoritmos serão comparados com as versões que fazem escolha das funções objetivo em cada *cluster* usando as técnicas de seleção de atributos descritas no capítulo anterior.

Inicialmente, será apresentado o conjunto de instâncias geradas e suas características. As implementações utilizadas foram feitas usando a linguagem Java com JDK na versão 8. Todos os experimentos descritos neste capítulo foram feitos usando estas instâncias. O acesso às instâncias poderá ser realizado com link ao final da seção. As abordagens a seguir serão testadas no PRVTRD com muitos objetivos reduzindo a dimensionalidade do problema para uma formulação bi-objetivo:

1. Aplicação do *cluster* hierárquico *offline* usando o coeficiente τ de Kendall com posterior execução do algoritmo MOEA/D (OffCL $_{\tau}$ -MOEA/D);
2. Aplicação do *cluster* hierárquico *offline* usando o coeficiente de correlação de Pearson com posterior execução do algoritmo MOEA/D (OffCL $_{\rho}$ -MOEA/D);
3. Utilização do *cluster* hierárquico *online* usando coeficiente τ de Kendall dentro do processo evolutivo do algoritmo MOEA/D (OnCL $_{\tau}$ -MOEA/D);
4. Utilização do *cluster* hierárquico *online* usando o coeficiente de correlação de Pearson dentro do processo evolutivo do algoritmo MOEA/D (OnCL $_{\rho}$ -MOEA/D);
5. Abordagem para *baseline*, a versão clássica do algoritmo MOEA/D.

5.1 Geração de Instâncias

As instâncias do problema foram geradas considerando a região central de Belo Horizonte contendo 50, 100, 150, 200 e 250 solicitações. Também foi utilizada uma frota homogênea (em relação à capacidade) de veículos com disponibilidade para 4 passageiros. Em cada instância, foram utilizados 12 pontos de parada, incluindo o depósito. O intervalo da janela de tempo foi definido para 10 minutos e o tempo máximo de espera de 3 minutos. Não foram realizados estudos prévios em relação aos parâmetros apresentados e para a definição dos pontos de parada.

A nomenclatura das instâncias utiliza o seguinte padrão: rAnBtwCkD, em que os valores de A, B, C e D são referentes a valores utilizados na instância. O valor de A

representa a quantidade de solicitações, B representa o número de pontos de parada, C o intervalo da janela de tempo e D é a capacidade dos veículos da frota. Por exemplo, para representarmos a instância com 50 solicitações, 12 pontos de parada, janela de tempo de 10 minutos e veículos com capacidade de 4 passageiros, usamos o seguinte nome: r050n12tw10k04.

Devido à dificuldade de acesso aos dados reais de deslocamentos de passageiros (por exemplo: origens e destinos), os pontos de parada foram definidos como os locais com uma alta concentração de pessoas. Para obter os valores da distância e o tempo de deslocamento entre os pontos de parada, foi usada a API do Google Maps.

Os valores de distância utilizados são aqueles que fazem o menor caminho entre os pontos de parada, levando em consideração o sentido das ruas. Vale ressaltar que o tempo de deslocamento coletado pela API é uma estimativa com base em uma série histórica de dados feita pelo próprio Google. Como a estimativa dos tempos de deslocamento pode variar com base na série histórica, foi necessário coletar os dados uma vez e armazená-los. Os dados foram coletados no dia 16/06/2017 e estão disponíveis no link <https://goo.gl/kKLxes>.

5.2 Configuração dos Parâmetros dos Algoritmos

Em todas as versões dos algoritmos usados neste trabalho, os parâmetros foram os mesmos e estão listados a seguir. A escolha dos valores foi feita baseada naqueles encontrados habitualmente na literatura, não sendo feito um ajuste fino.

- Probabilidade de cruzamento (p_c) = 0,7;
- Probabilidade de mutação (p_m) = 0,02;
- Número máximo de avaliações de função objetivo = 60.000;
- Tamanho da população = 100;
- Tamanho da vizinhança = 10;
- Número máximo de soluções substituídas = 5;
- Função de agregação no MOEA/D = *Boundary Intersection*;
- Tamanho do arquivo externo = 100.

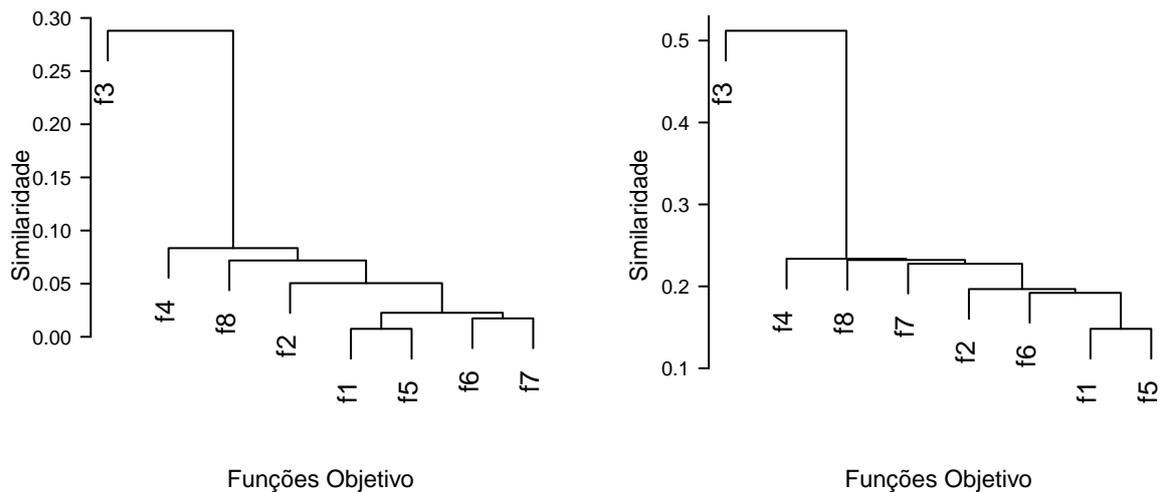
Os operadores genéticos (descritos na Seção 4.6) utilizados foram:

- Seleção por roleta;
- Cruzamento de dois pontos;
- Mutação 2-*Shuffle*.

A restrição de qualidade no atendimento dos passageiros, representando a porcentagem de solicitações não atendidas, é definida como 0%. Os algoritmos foram executados 30 vezes e foram geradas 30 populações iniciais diferentes, referentes a cada execução. Em todas as abordagens realizadas na tese foram utilizadas essas populações iniciais.

5.3 Redução de Dimensionalidade *Online* com Análise de *Cluster*

Para reduzir a dimensionalidade antes do processo de busca na abordagem *offline*, são utilizadas 2.500 soluções aleatórias para cada instância, a fim de obter uma amostra maior do espaço de soluções. Como foram utilizadas 5 instâncias, foram geradas um total de 12.500 soluções aleatórias. Uma perturbação, trocando-se dois pontos de parada aleatoriamente no vetor de rotas concatenadas, gerou cada solução na solução construída por um algoritmo guloso simples. As 12.500 soluções aleatórias são combinadas e avaliadas na formulação original com muitos objetivos (em \mathbb{R}^8) e usadas como uma entrada para a redução de dimensionalidade. A Figura 19 mostra os dendrogramas para o *cluster* hierárquico usando correlação de Pearson e τ de Kendall como medidas de similaridade entre as funções objetivo para as soluções aleatórias geradas anteriormente¹.



(a) Dendrograma para *cluster* hierárquico usando correlação de Pearson como medida de similaridade.

(b) Dendrograma para *cluster* hierárquico usando τ de Kendall como medida de similaridade.

Figura 19 – Dendrogramas para as 12.500 soluções aleatórias usando correlação de Pearson e τ de Kendall como medidas de similaridade entre as funções objetivo. As figuras foram feitas utilizando o pacote *hclust* da linguagem R.

Após aplicar a redução da dimensionalidade de *cluster* às soluções aleatórias usando tanto τ de Kendall quanto o coeficiente de correlação de Pearson, a formulação bi-objetivo é obtida. Embora a ordem em que as funções objetivo são agrupadas no mesmo *cluster* seja diferente considerando os coeficientes de correlação, a formulação bi-objetivo

¹ Os resultados apresentados podem ser encontrados no artigo [Mendes et al. \(2021\)](#)

final para o PRVTRD é a mesma para ambos os métodos e é dada por:

$$\begin{aligned} \text{Minimize } F_1 &= \lambda_1 f_1 + \lambda_2 f_2 + \lambda_4 f_4 + \lambda_5 f_5 + \lambda_6 f_6 + \lambda_7 f_7 + \lambda_8 f_8, \\ \text{Minimize } F_2 &= \lambda_3 f_3, \end{aligned} \quad (35)$$

em que λ_w , $w = 1, 2, \dots, 8$ representam os pesos associados a cada função objetivo. No caso dos experimentos realizados para a abordagem *offline*, os valores de λ_w foram todos definidos como 1 com o propósito de dar o mesmo peso para todas as funções objetivo. Como as formulações foram as mesmas em para as duas medidas de similaridade, a abordagem *offline* será denominada Off-MOEA/D sem o uso do índice como apresentado no início da seção.

Embora todos os algoritmos sejam aplicados para resolver a formulação bi-objetivo do PRVTRD, esta formulação pode ser diferente de geração para geração, e uma comparação de desempenho em \mathbb{R}^2 não é possível. No entanto, usando os resultados dos algoritmos em \mathbb{R}^2 , é possível obter suas contrapartes no espaço objetivo original, \mathbb{R}^8 . Desta forma, uma avaliação dos resultados no espaço objetivo original precisa ser realizada.

O objetivo é verificar se as abordagens de *cluster online*: (i) geram a mesma fronteira não dominada do problema original, ou (ii) alternativamente, preservam as relações de dominância do problema original. Uma medida de desempenho qualitativa deve ser usada. A avaliação de desempenho do algoritmo deve ser feita levando em consideração a propriedade de convergência (a minimização da distância do conjunto não dominado obtido para a verdadeira fronteira Pareto) e a propriedade de diversidade (proporcionando uma distribuição uniforme das soluções nesta fronteira e maximizando a extensão da fronteira não dominada obtida).

Muitas métricas famosas visam medir as propriedades de convergência e diversidade de diferentes algoritmos, mas a maioria delas requer o conhecimento da verdadeira fronteira Pareto ou tem baixa escalabilidade. Neste trabalho, o Indicador de Hipervolume (do inglês *Hypervolume Indicator* - HI) é aplicado para se avaliar a qualidade das soluções obtidas. O indicador foi proposto em nos trabalhos [Zitzler e Thiele \(1998\)](#), [Zitzler e Thiele \(1999b\)](#) e é um dos indicadores de qualidade para conjuntos mais usados para avaliar algoritmos multiobjetivos [Li e Yao \(2019\)](#). Ele avalia o resultado de um algoritmo levando em consideração a proximidade dos pontos à fronteira Pareto, diversidade e dispersão simultaneamente. Neste trabalho, o algoritmo *Walking Fish Group* (WFG) [Watanabe, Tatsukawa e Oyama \(2015\)](#) é aplicado para calcular o indicador de hipervolume de todos os conjuntos não dominados. A escolha do algoritmo foi feita pelo fato do algoritmo ter um custo computacional reduzido (em tempo) do cálculo do indicador

As figuras [20a](#), [20b](#), [20c](#), [20d](#) e [20e](#) mostram os *boxplots* dos valores do indicador de hipervolume obtidos nas 30 execuções para os algoritmos OffCL-MOEA/D, OnCL₇-MOEA/D,

OnCL $_{\rho}$ -MOEA/D e MOEA/D usando as instâncias com 50, 100, 150, 200 e 250 requisições, respectivamente. No entanto, uma vez que uma sobreposição das caixas pode ser vista para valores HI entre os algoritmos OnCL $_{\tau}$ -MOEA/D, OnCL $_{\rho}$ -MOEA/D e MOEA/D, um teste estatístico precisa ser realizado para identificar se há uma diferença entre os algoritmos.

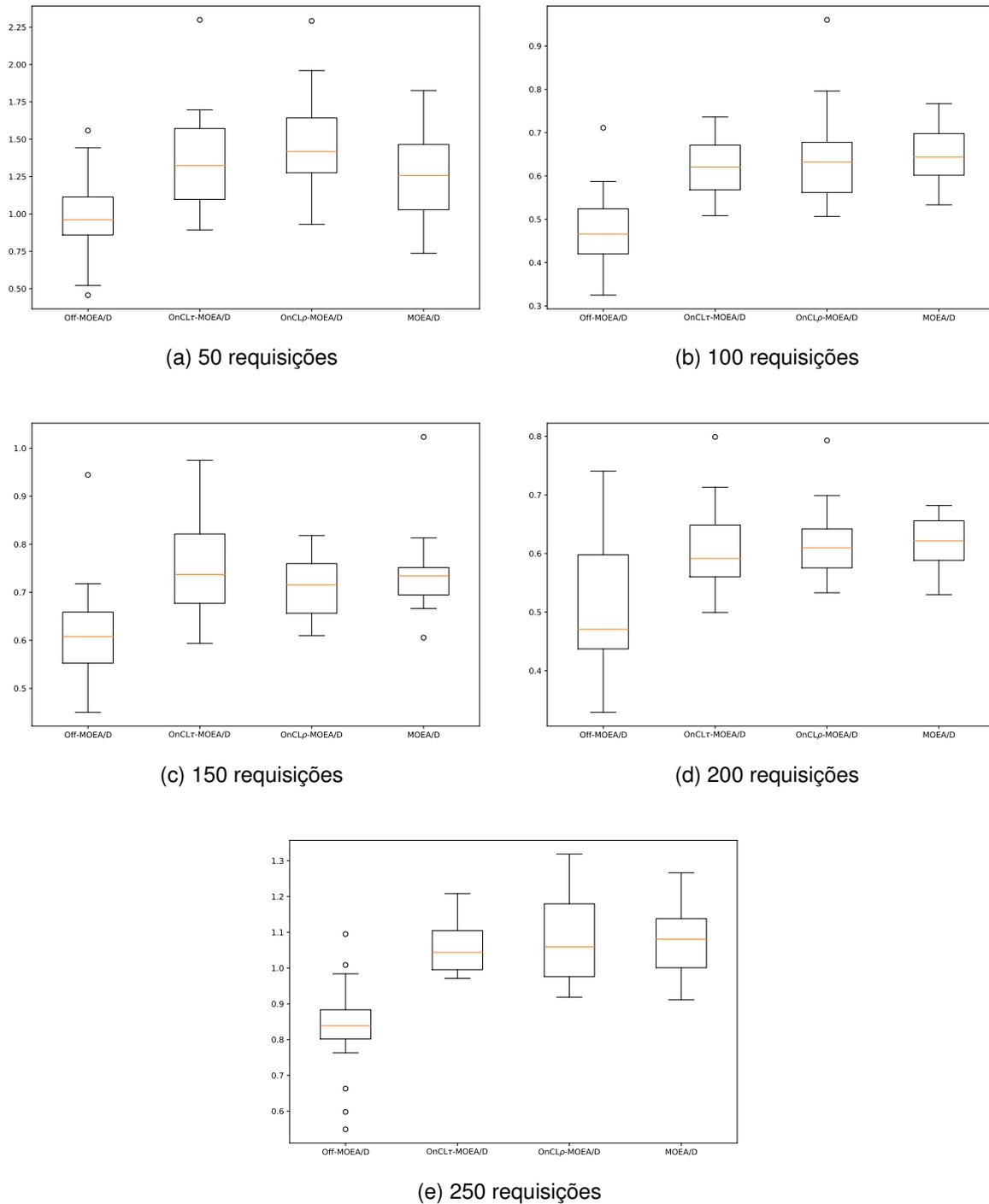


Figura 20 – *Boxplots* dos valores do indicador de hipervolume obtidos a partir das 30 execuções dos algoritmos OffCL-MOEA/D, OnCL $_{\tau}$ -MOEA/D, OnCL $_{\rho}$ -MOEA/D e MOEA/D usando 50, 100, 150, 200 e 250 requisições.

Um teste de aleatoriedade, encontrado em [Manly \(2006\)](#), é usado para determinar se os resultados obtidos são estatisticamente significativos. Os testes aleatórios são uma classe de testes não paramétricos amplamente aplicáveis e usam a mistura aleatória dos dados para obter a distribuição de uma estatística de teste sob uma hipótese nula. É um tipo de teste de significância estatística em que a distribuição da estatística de teste sob a hipótese nula é obtida calculando todos os valores possíveis da estatística de teste sob todos os rearranjos possíveis dos pontos de dados observados. A análise estatística é realizada usando $Média(A) - Média(B) = 0$ como uma hipótese nula e, desta forma, diferenças positivas favorecem A em relação a B .

Se o valor observado não for mais extremo do que uma fração dos resultados resultantes que corresponde a um nível de significância, então deve-se considerar a hipótese nula; de outra forma deve-se rejeitá-la. Pode ser observado que não há informação da distribuição verdadeira (amostral) da estatística de teste. No entanto, está sendo avaliado a distribuição numericamente usando a simulação de Monte Carlo. Uma aproximação do p -valor pode ser calculada pela proporção de valores de objetos simulados maiores do que o valor observado.

A Tabela 5.3 mostra os resultados do teste de aleatoriedade para todas as comparações de pares de algoritmos usando as instâncias com requisições 50, 100, 150, 200 e 250 usando os valores do indicador de hipervolume para 30 execuções de cada algoritmo. Embora os outros algoritmos tenham superado OffCL-MOEA/D em todas as instâncias, exceto para as 50 solicitações, a comparação entre pares contendo a versão *offline* é mantida na tabela para simplificar.

Usando um nível de significância de $\alpha = 0.05$, se a diferença observada estiver contida no meio de $(1 - \alpha)$ do intervalo de confiança, não é possível rejeitar a hipótese nula. Desta forma, não é possível afirmar que os resultados dos algoritmos são estatisticamente significativos. Caso contrário, a hipótese nula é rejeitada, e é possível dizer que os resultados dos algoritmos são diferentes no nível de significância α . Em cada comparação entre pares, caso a hipótese nula seja rejeitada, o algoritmo que apresenta desempenho superior é representado em negrito. Alternativamente, a ausência de fonte em negrito significa que não é possível afirmar que existe uma diferença estatística entre os resultados dos algoritmos. Além disso, os p -valores também são calculados.

Observando a Tabela 5.3, é possível ver que OffCL-MOEA/D é superado por todas as outras versões do algoritmo em todas as instâncias. Como esperado, as abordagens *online* alcançam melhores resultados em todas as instâncias testadas quando comparadas com a abordagem *offline*.

Tabela 3 – Testes aleatórios para todos os pares de algoritmos usando instâncias com 50, 100, 150, 200 e 250 requisições. O nível de significância é definido como 0.05. Se a hipótese nula for rejeitada, o algoritmo que apresenta desempenho superior é representado em **negrito**. Alternativamente, a ausência de **negrito** significa que não é possível afirmar que há uma diferença estatística entre os resultados dos algoritmos. Os p -valores para cada teste estatístico também são indicados.

Instância	Comparação Par-a-Par	p -valor
50 solicitações	OffCL-MOEA/D × OnCL_{τ}-MOEA/D	0.0156
	OffCL-MOEA/D × OnCL_{ρ}-MOEA/D	0.0002
	OffCL-MOEA/D × MOEA/D	0.0024
	OnCL _{τ} -MOEA/D × OnCL _{ρ} -MOEA/D	0.2214
	OnCL _{τ} -MOEA/D × MOEA/D	0.3430
	OnCL _{ρ} -MOEA/D × MOEA/D	0.0580
100 solicitações	OffCL-MOEA/D × OnCL_{τ}-MOEA/D	0.0002
	OffCL-MOEA/D × OnCL_{ρ}-MOEA/D	0.0002
	OffCL-MOEA/D × MOEA/D	0.0004
	OnCL _{τ} -MOEA/D × OnCL _{ρ} -MOEA/D	0.3384
	OnCL _{τ} -MOEA/D × MOEA/D	0.3482
	OnCL _{ρ} -MOEA/D × MOEA/D	0.4978
150 solicitações	OffCL-MOEA/D × OnCL_{τ}-MOEA/D	0.0002
	OffCL-MOEA/D × OnCL_{ρ}-MOEA/D	0.0002
	OffCL-MOEA/D × MOEA/D	0.0002
	OnCL _{τ} -MOEA/D × OnCL _{ρ} -MOEA/D	0.3008
	OnCL _{τ} -MOEA/D × MOEA/D	0.4188
	OnCL _{ρ} -MOEA/D × MOEA/D	0.3338
200 solicitações	OffCL-MOEA/D × OnCL_{τ}-MOEA/D	0.0368
	OffCL-MOEA/D × OnCL_{ρ}-MOEA/D	0.0092
	OffCL-MOEA/D × MOEA/D	0.0012
	OnCL _{τ} -MOEA/D × OnCL _{ρ} -MOEA/D	0.2498
	OnCL _{τ} -MOEA/D × MOEA/D	0.1596
	OnCL _{ρ} -MOEA/D × MOEA/D	0.3566
250 solicitações	OffCL-MOEA/D × OnCL_{τ}-MOEA/D	0.0010
	OffCL-MOEA/D × OnCL_{ρ}-MOEA/D	0.0002
	OffCL-MOEA/D × MOEA/D	0.0002
	OnCL _{τ} -MOEA/D × OnCL _{ρ} -MOEA/D	0.3884
	OnCL _{τ} -MOEA/D × MOEA/D	0.2412
	OnCL _{ρ} -MOEA/D × MOEA/D	0.3608

Comparando os algoritmos online, OnCL _{ρ} -MOEA/D e OnCL _{τ} -MOEA/D, em todos os casos, não é possível afirmar que haja diferença entre eles. É importante notar que, em

todos os casos, não é possível dizer que há diferença entre o MOEA/D e as duas versões *online*, OnCL _{ρ} -MOEA/D e OnCL _{τ} -MOEA/D.

Esta incomparabilidade entre o algoritmo baseline, MOEA/D, e as abordagens de cluster *online*, OnCL _{ρ} -MOEA/D e OnCL _{τ} -MOEA/D, corrobora a suposição de que as abordagens *online* propostas (i) geram o mesmo conjunto Pareto que o problema original, ou (ii) alternativamente, preservam as relações de dominância do problema original. Além disso, a abordagem *online* pode ser vista como uma abordagem eficaz de redução de dimensionalidade que transforma uma formulação de oito objetivos em uma formulação agregada bi-objetivo para o PRVTRD.

Uma forma eficaz de visualizar os conjuntos Pareto aproximados obtidos é essencial. Informações sobre a faixa de localização, formato da fronteira, avaliação de conflitos e o *trade-off* entre os objetivos, dentre outras, podem ser extraídas quando uma ferramenta de visualização é aplicada [Tusar e Filipic \(2015\)](#).

Uma ferramenta de visualização multifuncional [Meneghini, Koochaksaraei e Guimarães \(2014\)](#) é usada para analisar os resultados dos algoritmos no espaço objetivo original (\mathbb{R}^8). Esta ferramenta de visualização reúne três métodos de visualização, e um desses métodos, o diagrama de cordas [Koochaksaraei et al. \(2017\)](#), é usado aqui. Os dados de entrada são uma matriz de tamanho $N \times M$ contendo um conjunto de N pontos no espaço M -dimensional (aqui, $M = 8$). O diagrama de cordas é uma abordagem circular para exibir o vetor de um espaço M -dimensional em um gráfico bidimensional.

No diagrama de cordas, cada arco representa uma função objetivo do problema. O intervalo de cada objetivo é mostrado do limite inferior ao limite superior no sentido horário. Em nosso estudo, existem 8 arcos que mostram o alcance de todos os objetivos. Cada ponto no espaço objetivo de oito dimensões é exibido por curvas conectando os valores correspondentes em cada coordenada. O diagrama de cordas é implementado em R, e o R-script é disponibilizado pelos autores no site do MINDS Lab [Meneghini, Koochaksaraei e Guimarães \(2014\)](#).

Como o diagrama de cordas preserva algumas características do espaço de objetivos, ele pode ser usado de forma eficaz para estimar o alcance da área explorada e o desempenho dos algoritmos de uma forma qualitativa. O diagrama é aplicado para avaliar os resultados das versões com *cluster online* do algoritmo, OnCL _{τ} -MOEA/D e OnCL _{ρ} -MOEA/D. Essas versões foram consideradas equivalentes quando o indicador de hipervolume é utilizado para avaliar a qualidade das soluções. O objetivo neste ponto é avaliar se os algoritmos exploram diferentes regiões do espaço de objetivos. As Figuras [21a](#), e [21b](#), [21c](#) e [21d](#), [22a](#) e [22b](#), [22c](#) e [22d](#), e [22e](#) e [22f](#) mostram os conjuntos Pareto combinados em \mathbb{R}^8 usando o diagrama de cordas para os algoritmos OnCL _{ρ} -MOEA/D e OnCL _{τ} -MOEA/D nas instâncias com 50, 100, 150, 200 e 250 solicitações, respectivamente.

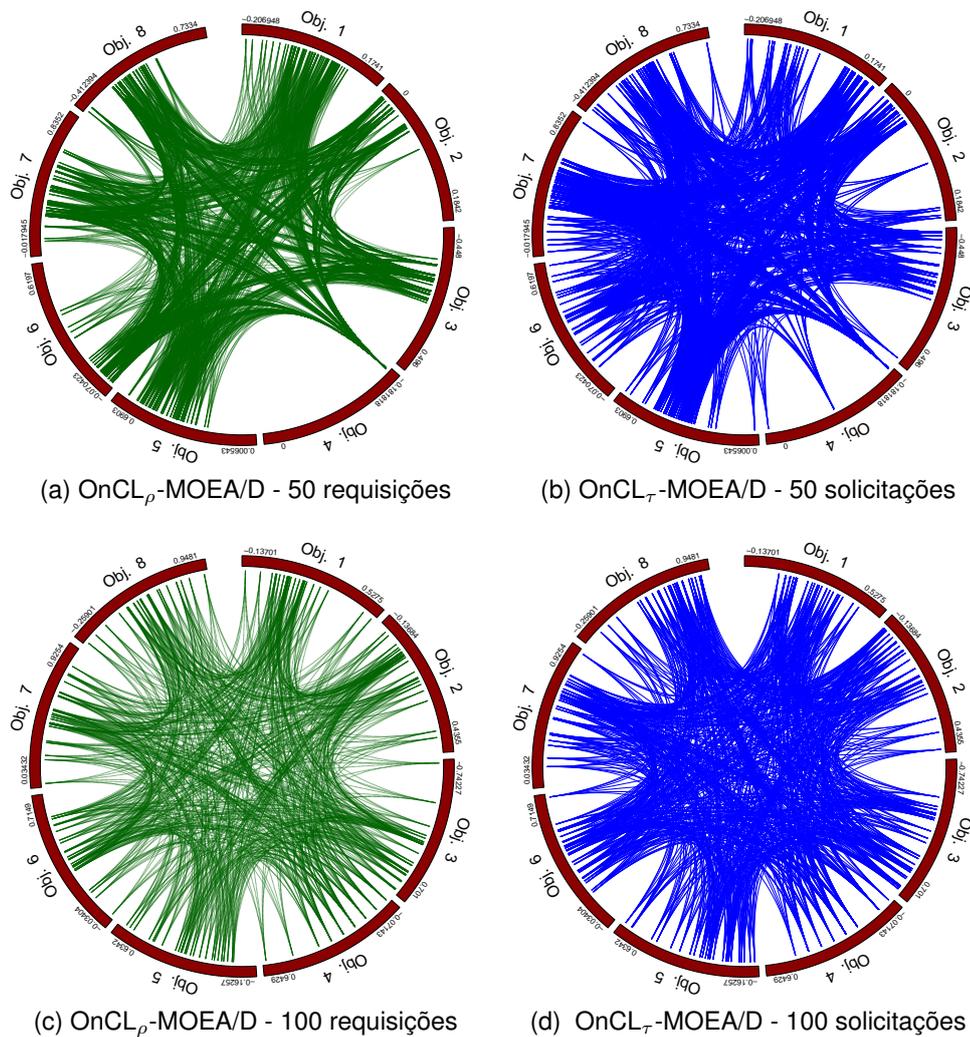


Figura 21 – Os diagramas de corda mostram a dispersão do conjunto Pareto combinado para as instâncias com 50 e 100 solicitações.

Quanto à distribuição, analisando a Figura 21a e 21b, pode-se notar que as soluções dos algoritmos OnCL $_{\rho}$ -MOEA/D e OnCL $_{\tau}$ -MOEA/D são diferentes em termos de diversidade quando aplicadas à instância com 50 solicitações. A Figura 21a apresenta mais soluções (62) quando comparada com a Figura 21b (52). As soluções obtidas pelo algoritmo OnCL $_{\rho}$ -MOEA/D são altamente fragmentárias em algumas regiões dos objetivos 2, 3 e 4 e há poucas soluções perto do limite inferior dos objetivos 3 e 5, e perto do limite superior dos objetivos 2, 4 e 8 quando comparadas com as soluções obtidas pelo algoritmo OnCL $_{\tau}$ -MOEA/D. As soluções obtidas pelo OnCL $_{\tau}$ -MOEA/D são mais distribuídas ao longo da faixa de todos os objetivos quando comparadas com as soluções obtidas pelo OnCL $_{\rho}$ -MOEA/D.

Uma análise semelhante pode ser feita observando as Figuras 21c e 21d, em que pode-se notar que as soluções de OnCL $_{\rho}$ -MOEA/D e OnCL $_{\tau}$ -MOEA/D são diferentes em termos de diversidade quando aplicadas à instância com 100 requisições. A Figura 21c apresenta aproximadamente o mesmo número de soluções (25) quando comparada com

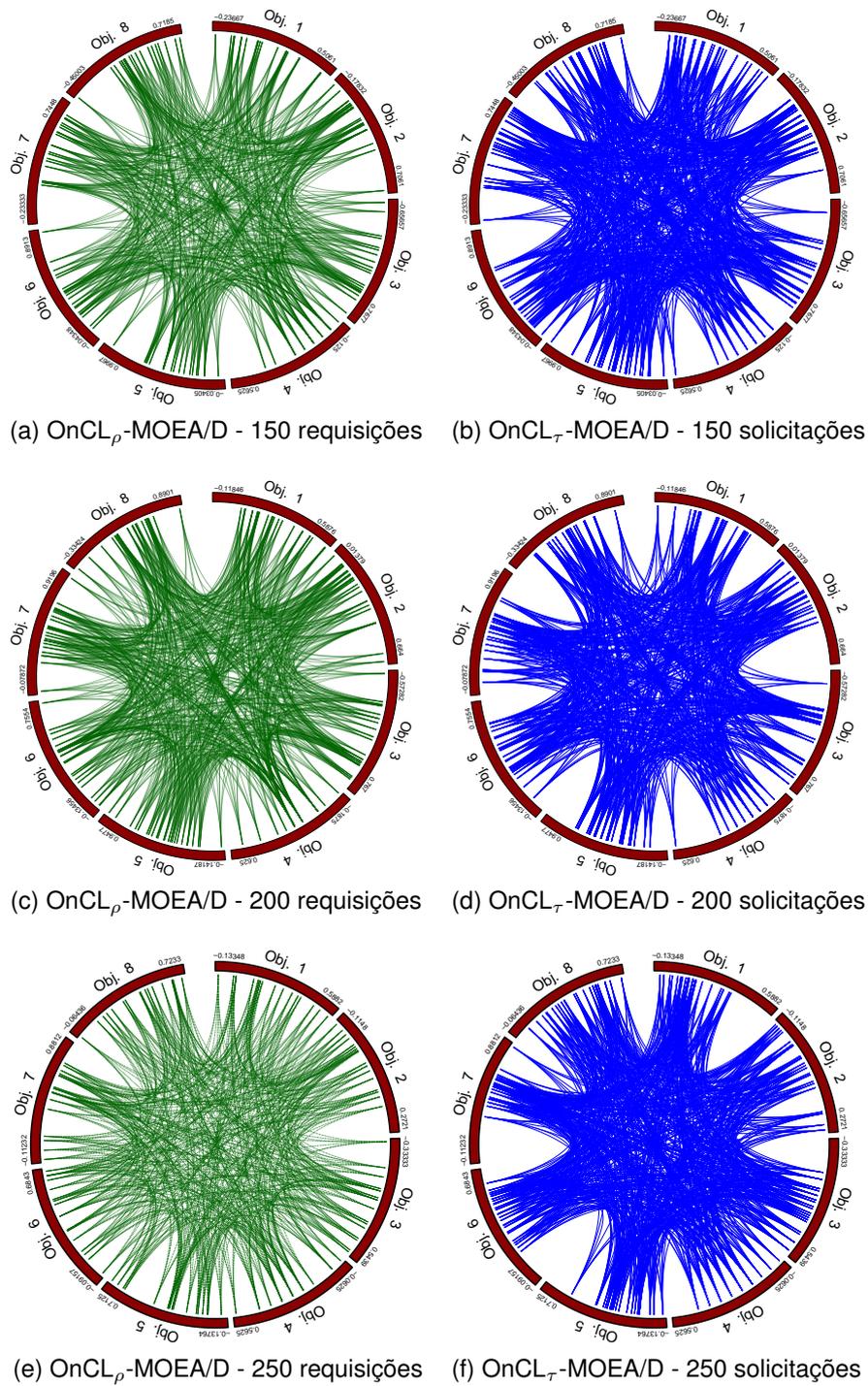


Figura 22 – Os diagramas de corda mostram a dispersão do conjunto Pareto combinado para as instâncias com 150, 200 e 250 solicitações.

a Figura 21d (27). As soluções obtidas pelo OnCL $_{\tau}$ -MOEA/D estão mais densamente concentradas em algumas regiões em todos os objetivos e não há soluções próximas ao limite inferior dos objetivos 3 e 6 quando comparadas com as soluções obtidas pelo OnCL $_{\rho}$ -MOEA/D.

Observando Figuras 22a e 22b, o número de soluções em ambas as figuras é

diferente (20 e 26, respectivamente). Pode-se notar que as soluções de $\text{OnCL}_\rho\text{-MOEA/D}$ e $\text{OnCL}_\tau\text{-MOEA/D}$ não são muito diferentes em termos de diversidade quando aplicadas à instância com 150 requisições. As soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$ estão pouco concentradas em algumas regiões dos objetivos 2, 3, 4 e 6 e não há soluções próximas ao limite superior do objetivo 6 quando comparadas com as soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$. Por outro lado, as soluções obtidas pelo $\text{OnCL}_\rho\text{-MOEA/D}$ são fragmentadas perto do limite inferior do objetivo 8.

A diversidade de soluções obtidas pelo $\text{OnCL}_\rho\text{-MOEA/D}$ e $\text{OnCL}_\tau\text{-MOEA/D}$ também são muito distintas quando aplicadas à instância com 200 requisições. A Figura 22c apresenta quase o mesmo número de soluções (24) quando comparada com a Figura 22d (23). As soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$ são fragmentadas em algumas regiões dos objetivos 2, 3, 4 e 5 e não existem soluções do limite inferior do objetivo 8 quando comparadas com as soluções obtidas pelo $\text{OnCL}_\rho\text{-MOEA/D}$. Além disso, observando as soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$, não há soluções próximas do limite superior dos objetivos 2 e 4 quando comparados com as soluções obtidas pelo $\text{OnCL}_\rho\text{-MOEA/D}$.

Observando Figuras 22e e 22f, embora o número de soluções em ambas as figuras seja diferente favorecendo $\text{OnCL}_\tau\text{-MOEA/D}$ (17 e 25, respectivamente), as soluções de ambos os algoritmos são muito diferentes em termos de diversidade quando aplicadas à instância com 250 requisições. As soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$ estão pouco concentradas em algumas regiões dos objetivos 1, 6 e 7, e não há soluções perto do limite superior dos objetivos 1 e 4, e do limite inferior do objetivo 3 quando em comparação com as soluções obtidas pelo $\text{OnCL}_\tau\text{-MOEA/D}$.

Os resultados mostraram que os algoritmos baseados em *cluster online* (usando correlação de Pearson e τ de Kendall) são estatisticamente melhores do que o algoritmo baseado em *cluster offline* em todas as cinco instâncias testadas. Outro ponto observado é que não é possível afirmar que existe diferença estatística entre os resultados dos algoritmos baseados em *cluster online* (algoritmos $\text{OnCL}_\rho\text{-MOEA/D}$, $\text{OnCL}_\tau\text{-MOEA/D}$) e o MOEA/D quando o indicador de hipervolume é usado para avaliar a diversidade das soluções obtidas em todas as instâncias testadas. Dessa forma, pode-se ver que a abordagem com *cluster online* é uma técnica eficaz de redução de dimensionalidade que transforma o problema de muitos objetivos em uma formulação agregada bi-objetivo para todas as instâncias testadas no PRVTRD. A redução da dimensionalidade tem um impacto favorável não só na eficiência da pesquisa e no custo computacional, mas também em permitir a aplicação de técnicas usuais de visualização.

Embora não haja diferença estatística entre os resultados das abordagens *online*, a ferramenta de visualização de diagramas de corda foi aplicada para estimar efetivamente a extensão da área explorada e o desempenho qualitativo dos algoritmos de *cluster online*. Os resultados indicaram que a diversidade das soluções obtidas pelo $\text{OnCL}_\rho\text{-MOEA/D}$ e

OnCL $_{\tau}$ -MOEA/D também são ligeiramente distintas quando aplicadas às instâncias. Um estudo de diversidade para avaliar a influência dos coeficientes de Pearson e τ de Kendall pode ser feito como trabalho futuro.

No que diz respeito à frequência de redução da dimensionalidade, o problema original com muitos objetivos é reduzido a uma formulação bi-objetivo em cada geração do algoritmo. O intuito é explorar as características do espaço de objetivos original ao longo do processo de busca.

Baseando nos resultados encontrados nesses experimentos, surgem duas questões a serem investigadas: Qual abordagem apresenta uma melhor fronteira Pareto final, a redução da dimensionalidade baseada na agregação apresentada anteriormente ou a eliminação de objetivos não conflitantes? O que é melhor, realizar a redução a cada geração do algoritmo ou após um determinado número predefinido de gerações? Na seção a seguir serão apresentados resultados dos experimentos realizados para investigar as questões levantadas.

5.4 Redução de Dimensionalidade *Online* Usando Seleção de Atributos

Nesta segunda parte dos experimentos, foram feitos testes relacionados às abordagens que eliminam os objetivos de cada *cluster* usando as técnicas de seleção de atributos, descritas na Seção 4.4. As abordagens são acopladas ao algoritmo MOEA/D. A correlação de Pearson foi utilizada como medida de similaridade pela sua facilidade de cálculo e pelo fato de não haver diferença entre resultados das abordagens online². Usando as mesmas cinco instâncias do PRVTRD descritas na seção anterior, os seguintes algoritmos são avaliados:

- Abordagem de *cluster online* usando o coeficiente de Pearson juntamente com MOEA/D: OnCL $_{\rho_g}$ -MOEA/D;
- Abordagem de *cluster online* usando o coeficiente de Pearson e a Taxa de Dispersão (DR) juntamente com MOEA/D: OnDR $_{\rho_g}$ -MOEA/D;
- Abordagem de *cluster online* usando coeficiente de Pearson e Máxima Variância (MV) juntamente com MOEA/D: OnMV $_{\rho_g}$ -MOEA/D;
- *Baseline*, a versão clássica do MOEA/D.

Todos os testes realizados nesta seção utilizaram os mesmos parâmetros e operadores descritos na Seção 5.2, com exceção do momento em que a redução de dimensionalidade foi realizada. O índice subscrito g indica a frequência da redução da dimensionalidade ao longo das gerações do MOEA/D. Independentemente da abordagem, a redução da

² Os resultados apresentados podem ser encontrados no artigo [Mendes et al. \(2021\)](#)

dimensionalidade é realizada (i) a cada geração ($g = 1$); e (ii) em 25 gerações ($g = 25$); (iii) em 50 gerações ($g = 50$); e (iv) na geração 100 ($g = 100$) do MOEA/D. Os resultados são mostrados para cada instância do PRVTRD.

Vale ressaltar dois pontos antes de apresentar os resultados nesta seção, primeiro, os *boxplots* foram feitos para todas as frequências (1, 25, 50 e 100), mas serão mostrados à seguir apenas para as frequências iguais a 1 e 100 com o objetivo de simplificar a visualização. As frequências foram escolhidas para de tal modo que pudesse ser observado o efeito da redução de dimensionalidade entre a maior e menor frequência. Os *boxplots* para todas as instâncias em todas as frequências testadas estão presentes na Seção A.1 do Apêndice A.

O segundo ponto é o nome do algoritmo MOEA/D foi removido do nome (eixo horizontal) em todos os *boxplots* da Figura 23 nos resultados dos algoritmos em que ocorreram a redução de dimensionalidade, tanto por agregação quanto por seleção de atributos. Foi deixado apenas a forma como a redução de dimensionalidade foi feita e a respectiva frequência. A remoção foi feita para facilitar a leitura dos gráficos nesta seção.

As Figuras 23a 23b, 23c, 23d e 23e mostram os *boxplots* dos valores do indicador de hipervolume obtidos nas 30 execuções para OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D, e MOEA/D em todas as instâncias com as frequências de redução de dimensionalidade iguais a 1 e 100. Para todas as instâncias testadas, existem dois grupos distintos compreendendo os seguintes algoritmos: grupo (a) MOEA/D, OnCL ρ_1 -MOEA/D, OnCL ρ_{100} -MOEA/D e grupo (b) OnDR ρ_1 -MOEA/D, OnDR ρ_{100} -MOEA/D, OnMV ρ_1 -MOEA/D e OnMV ρ_{100} -MOEA/D. O mesmo resultado pode ser observado com as reduções de dimensionalidade a cada 25 e 50 gerações.

Como os intervalos dos *boxplots* entre os dois grupos (a) e (b) não se sobrepõem nas figuras, pode-se dizer que esses dois grupos são estatisticamente diferentes, sendo o primeiro grupo estatisticamente superior ao segundo. Porém, dentro dos grupos, pode-se observar uma sobreposição dos *boxplots*. Portanto, um teste estatístico precisa ser realizado para identificar se há diferença entre os algoritmos.

O mesmo teste de aleatoriedade, descrito na Seção 5.3, é aplicado nos valores do indicador de hipervolume obtidos nas 30 execuções. A aplicação dos testes foi feita par-a-par dentro de cada grupo, gerando um total de 38 comparações. As Tabelas 4 e 5 do Apêndice A apresentam os p -valores para cada comparação realizada.

Quanto ao grupo (a), em todas as instâncias, não é possível afirmar que haja diferença estatística entre os resultados dos algoritmos (OnCL ρ_1 -MOEA/D, OnCL ρ_{25} -MOEA/D, OnCL ρ_{50} -MOEA/D, OnCL ρ_{100} -MOEA/D e MOEA/D) quando o nível de significância é definido como 5%. Levando em consideração o grupo (b), em todas as instâncias e quando o nível de significância é fixado em 0.05, não é possível afirmar que haja diferença estatís-

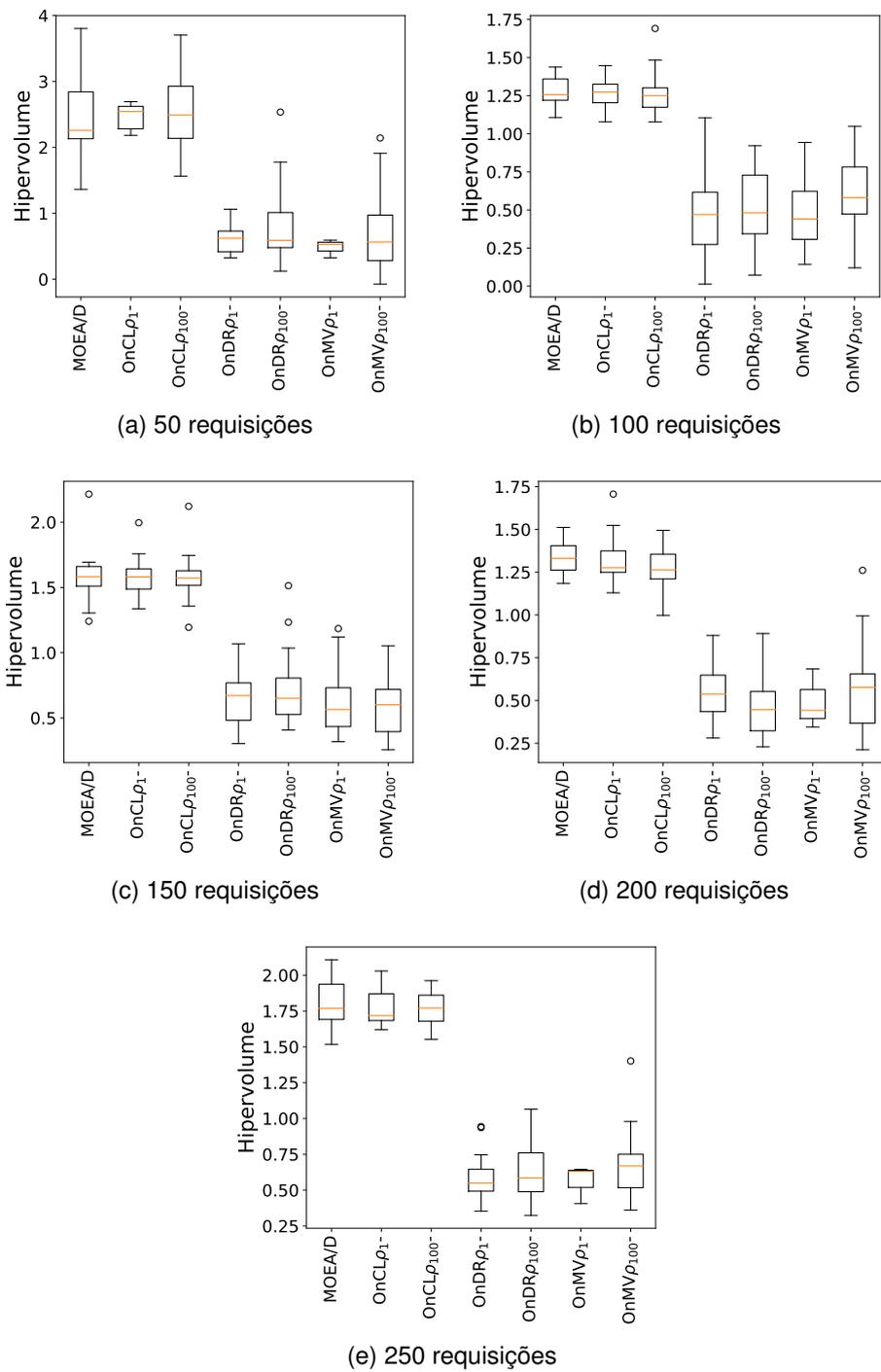


Figura 23 – *Boxplots* dos valores do indicador de hipervolume obtidos a partir das 30 execuções dos algoritmos MOEA/D, OnCL ρ_g -MOEA/D e OnMV ρ_g -MOEA/D com $g = 1, 100$ para as instâncias com 50, 100, 150, 200 e 250 requisições.

tica entre os resultados dos algoritmos OnDR ρ_1 -MOEA/D, OnDR ρ_{25} -MOEA/D, OnDR ρ_{50} -MOEA/D, OnDR ρ_{100} -MOEA/D, OnMV ρ_1 -MOEA/D, OnMV ρ_{25} -MOEA/D, OnMV ρ_{50} -MOEA/D e OnMV ρ_{100} -MOEA/D, independentemente da frequência da redução.

A incomparabilidade entre o MOEA/D e as abordagens de redução de dimensionalidade baseadas em agregação em conjunto com a diferença estatística favorecendo o

grupo (a) sobre o grupo (b) responde à pergunta levantada no final da seção anterior em relação a qual abordagem seria melhor para o problema: usar a agregação dos objetivos ou eliminar os objetivos não conflitantes. Para este conjunto de instâncias, é possível dizer que a redução de dimensionalidade baseada em agregação apresenta um desempenho superior que a baseada em seleção de atributos para eliminação de objetivos quando o hipervolume é usado como indicador de qualidade.

Além disso, os resultados também corroboram a suposição de que a relação entre os objetivos e todas as informações fornecidas pela formulação do problema original são preservadas na abordagem baseada em agregação. Um outro ponto observado é que, como afirmado em [Mendes et al. \(2021\)](#), a abordagem *online* de redução de dimensionalidade baseada em agregação (i) gera a mesma fronteira não dominada que o problema original, ou (ii) alternativamente, preserva as relações de dominância do problema original. Por outro lado, a eliminação das funções objetivo quando se aplica a redução da dimensionalidade baseada na seleção de atributos implica na perda de informações importantes sobre as características do problema.

O diagrama de cordas [Koochaksaraei et al. \(2017\)](#) também será utilizado para analisar os resultados dos Paretos combinados dos algoritmos em \mathbb{R}^8 . Neste caso, o diagrama será usado para avaliar os resultados dos algoritmos OnCL ρ_1 -MOEA/D, OnDR ρ_1 -MOEA/D, OnCL ρ_{100} -MOEA/D e OnDR ρ_{100} -MOEA/D. Os diagramas de cordas para todas as frequências em todas instâncias podem ser vistos no Apêndice B.

O algoritmo OnDR ρ_g -MOEA/D foi escolhido como o representante da abordagem baseada na seleção de atributos, mas resultados semelhantes são obtidos quando as versões com Máxima Variância são empregadas. Apesar do OnCL ρ_g -MOEA/D superar os outros dois algoritmos, o OnDR ρ_g -MOEA/D e OnMV ρ_g -MOEA/D foram considerados equivalentes na avaliação de desempenho. O objetivo aqui é analisar as diferentes regiões do espaço de busca exploradas por cada algoritmo. A Figura 24 apresenta os diagramas de corda para a instância de 50 requisições.

Pode ser visto que o algoritmo OnCL ρ_g -MOEA/D acaba se concentrando em determinadas regiões do espaço de objetivos quando comparado ao algoritmo OnDR ρ_g -MOEA/D para ambas as frequências de redução da dimensionalidade. Mesmo ocorrendo essa concentração, o algoritmo OnCL ρ_1 -MOEA/D (Figura 24a) consegue menores valores para os objetivos 1,3,4 e 8, e obtém um alcance similar no objetivo 5. Já o algoritmo OnCL ρ_{100} -MOEA/D (Figura 24c) obtém os menores valores para os objetivos 3, 4, 6, 7 e 8. Um ponto a ser observado é em relação à quantidade de soluções geradas pelos algoritmos. Percebe-se que os algoritmos com redução de dimensionalidade usando agregação dos objetivos apresenta um número maior de soluções ao final das execuções, podendo ser visto pela densidade das linhas geradas nas Figuras 24a e 24c.

Quanto ao espalhamento das soluções obtidas nas funções objetivo, o algoritmo

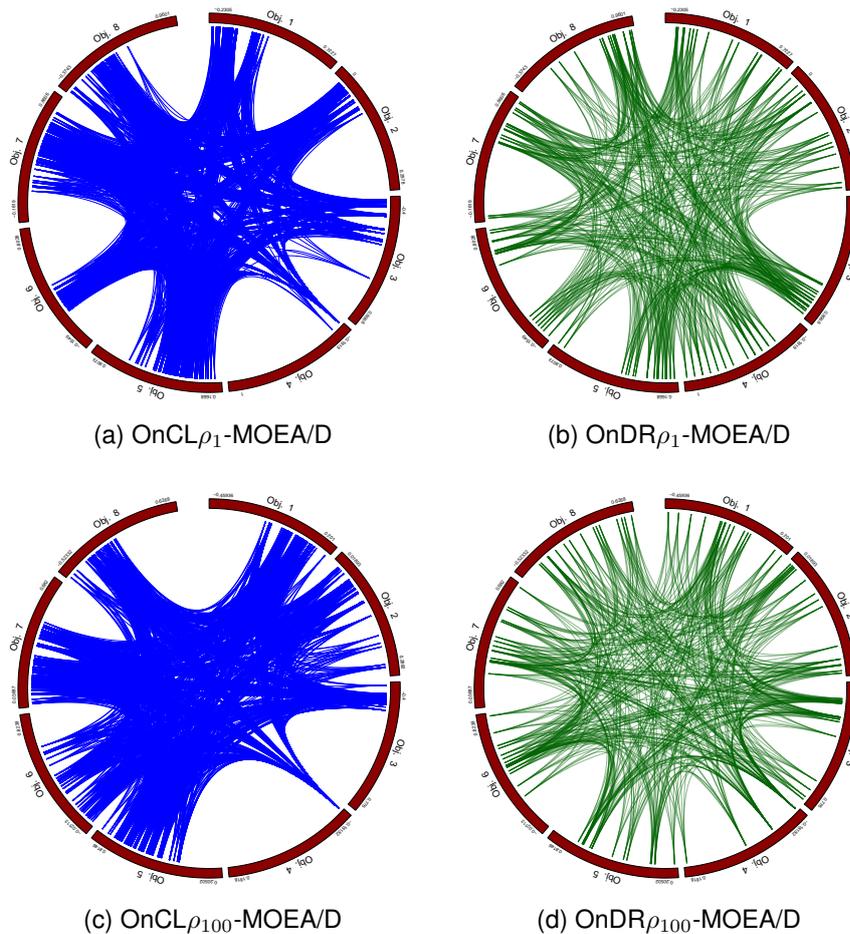


Figura 24 – Instância com 50 solicitações.

OnCL ρ_g -MOEA/D obtém uma melhor distribuição de valores nos objetivos 1,2 e 6, e possui uma distribuição similar nos objetivos 5, 7 e 8. Já para o algoritmo OnDR ρ_g -MOEA/D, ao se comparar as Figuras 24b e 24d, pode ser visto que há uma melhor distribuição nos objetivos 3, 5, 6, 7 e 8. Os diagramas de corda para a instância com 100 solicitações são apresentados na Figura 25.

Nas Figuras 25a e 25c, os diagramas mostram que há mais soluções obtidas pelo OnCL ρ_g -MOEA/D do que o algoritmo OnDR ρ_g -MOEA/D. Embora a diversidade das soluções nas Figuras 25a e 25c são semelhantes, o procedimento de eliminação e a frequência da redução afeta a diversidade das soluções obtidas ao final das execuções.

Um ponto que pode ser observado na instância de 100 solicitações ao se comparar os diagramas de corda de cada algoritmo: o aumento da frequência (intervalo entre as ocorrências da redução de dimensionalidade) proporcionou uma melhor distribuição das soluções. Essa melhoria nas distribuições pode ser observado ao se comparar a Figura 25a com a 25c e a Figura 25b com a 25d. Para o algoritmo OnCL ρ_g -MOEA/D, os objetivos 5, 6, 7 e 8, apresentam menores intervalos sem valores encontrados e o algoritmo OnDR ρ_g -MOEA/D consegue apresentar uma melhor distribuição em todos os objetivos.

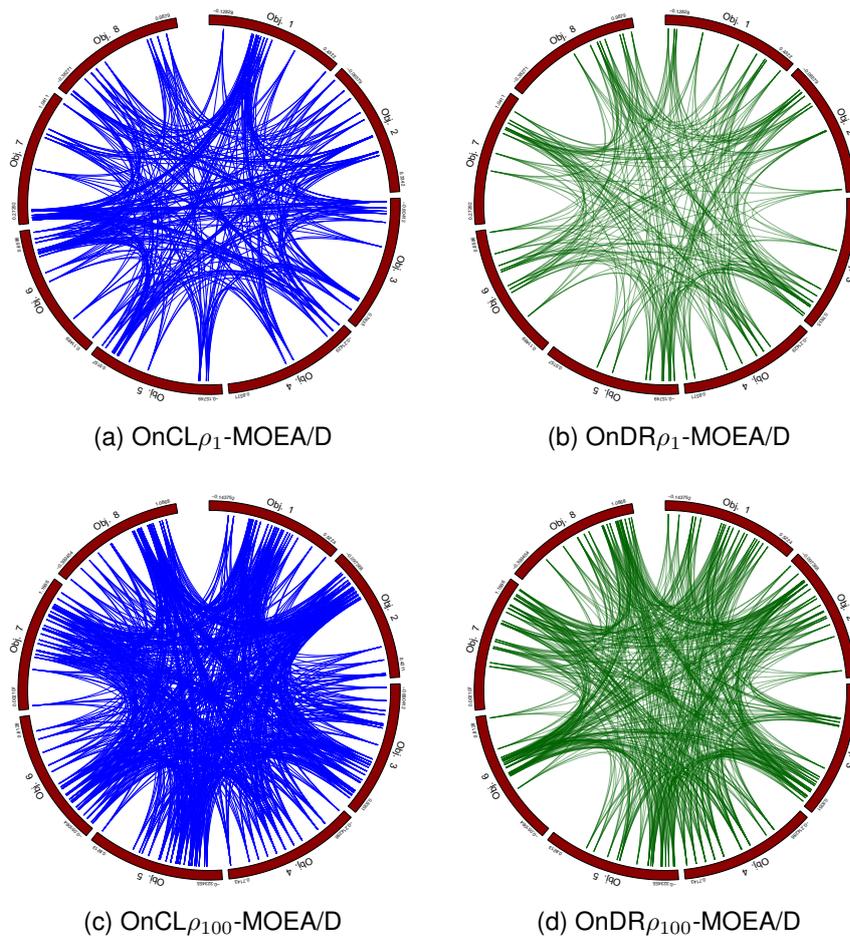


Figura 25 – Instância com 100 solicitações.

Para a instância com 150 solicitações, representada pelos diagramas de cordas dos conjuntos Pareto combinados na Figura 26, percebe-se que o algoritmo $\text{OnCL}_{\rho_g}\text{-MOEA/D}$ apresenta um número maior de soluções no conjunto Pareto combinado em ambas as frequências de redução de dimensionalidade. Pode ser visto que a densidade das linhas para o algoritmo $\text{OnDR}_{\rho_g}\text{-MOEA/D}$ é menor.

Em relação as regiões nas quais os algoritmos conseguem obter melhores valores, pode ser visto que quando a frequência da redução é igual a 1 o algoritmo $\text{OnDR}_{\rho_1}\text{-MOEA/D}$ consegue melhores resultados apenas nos objetivos 1 e 5, já o algoritmo $\text{OnCL}_{\rho_1}\text{-MOEA/D}$ consegue alcançar menores valores nos demais objetivos (2, 3, 4, 6 e 8). Agora, quando a frequência das reduções é igual a 100, o algoritmo $\text{OnCL}_{\rho_1}\text{-MOEA/D}$ consegue menores valores para os objetivos 3, 4, 6, 7 e 8.

Considerando os valores das funções objetivo alcançadas por cada algoritmo, o $\text{OnCL}_{\rho_g}\text{-MOEA/D}$, ao se comparar os diagramas de corda para as frequência iguais a 1 e 100 (Figuras 26a e 26a), não consegue uma melhor distribuição dos valores apenas para o objetivo 2, nos demais objetivos, pode ser visto uma maior distribuição dos valores dentro da amplitude obtida.

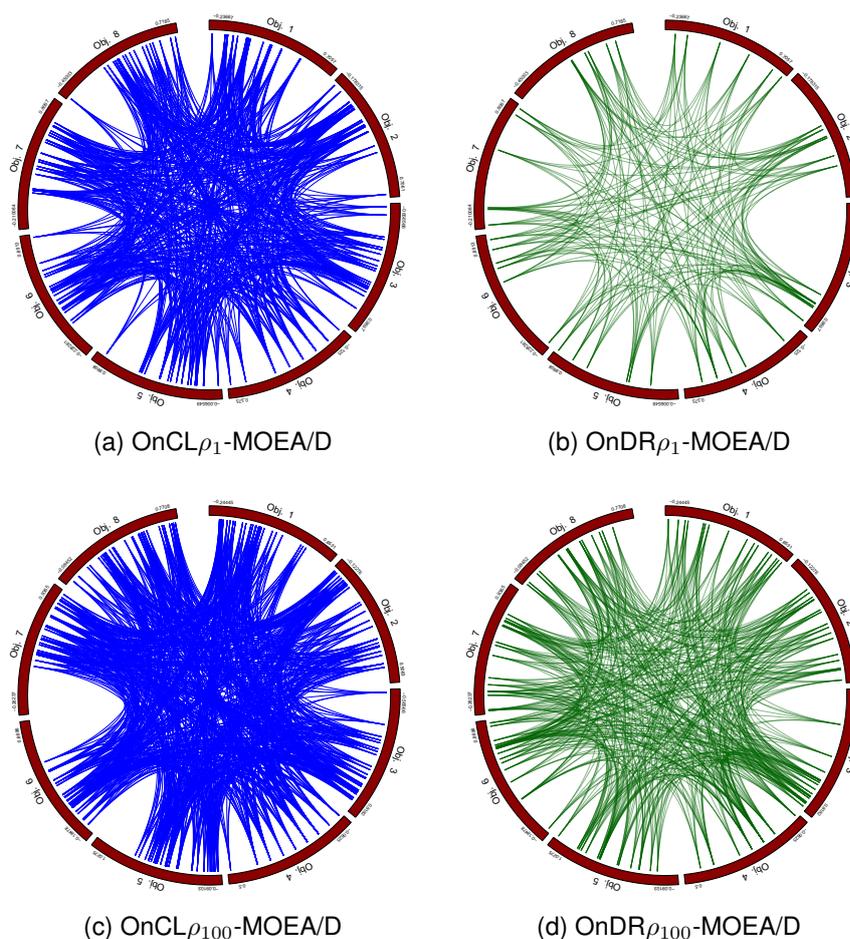


Figura 26 – Instância com 150 solicitações.

Os diagramas de corda para as soluções obtidas dos algoritmos OnCL ρ_g -MOEA/D e OnDR ρ_g -MOEA/D, para as frequências de redução iguais a 1 e 100 na instância com 200 requisições são apresentados na Figura 27. Em um primeiro momento, pode ser notado o aumento do número de soluções não dominadas em ambos os algoritmos quando comparamos as Figuras 27a com 27c e 27b com 27d.

Considerando os valores obtidos por cada algoritmo, para a frequência de redução igual a 1, o algoritmo OnCL ρ_1 -MOEA/D obtém menores valores para os objetivos 1, 2, 3, 4, 5 e 8. Quando a redução é feita a cada 100 gerações, o algoritmo OnCL ρ_{100} -MOEA/D consegue alcançar melhores valores para os objetivos com números iguais a 1, 3, 4, 5 e 8.

Ao se comparar as Figuras 27a com 27c e 27b com 27d pode ser observada uma melhor disposição dos valores obtidos dentro da amplitude alcançada por cada algoritmo quando a frequência de redução de dimensionalidade é igual a 100. Esse fato pode ser notado pela redução significativa dos intervalos sem apresentar nenhum valor em cada função objetivo.

Os diagramas de cordas para a instância de 250 solicitações é apresentado na

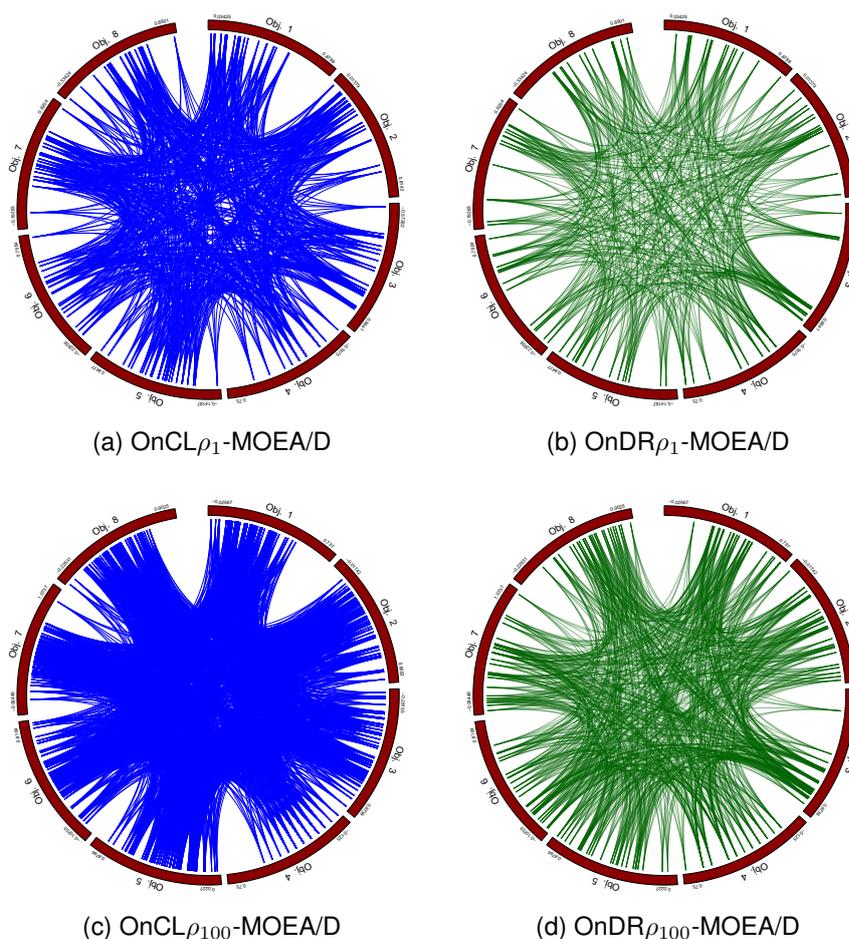


Figura 27 – Instância com 200 solicitações.

Figura 28. Nesta instância fica mais evidente apenas o aumento do número de soluções obtidas ao final das execuções apenas para o algoritmo OnDR ρ_g -MOEA/D quando os diagramas para as frequências iguais a 1 e 100 são comparados.

Corroborando os resultados obtidos para as instâncias anteriores, pode ser visto novamente que o algoritmo OnCL ρ_1 -MOEA/D consegue obter menores valores apenas para os objetivos 2 e 3 porém, pode ser visto que nos objetivos 4, 5, 6 e 7 houve uma concentração dos valores obtidos pelo algoritmo OnCL ρ_1 -MOEA/D em regiões não muito exploradas pelo OnDR ρ_1 -MOEA/D. Já o algoritmo OnCL ρ_{100} -MOEA/D alcança melhores valores nas funções objetivo de número 2, 3, 6, 7 e 8.

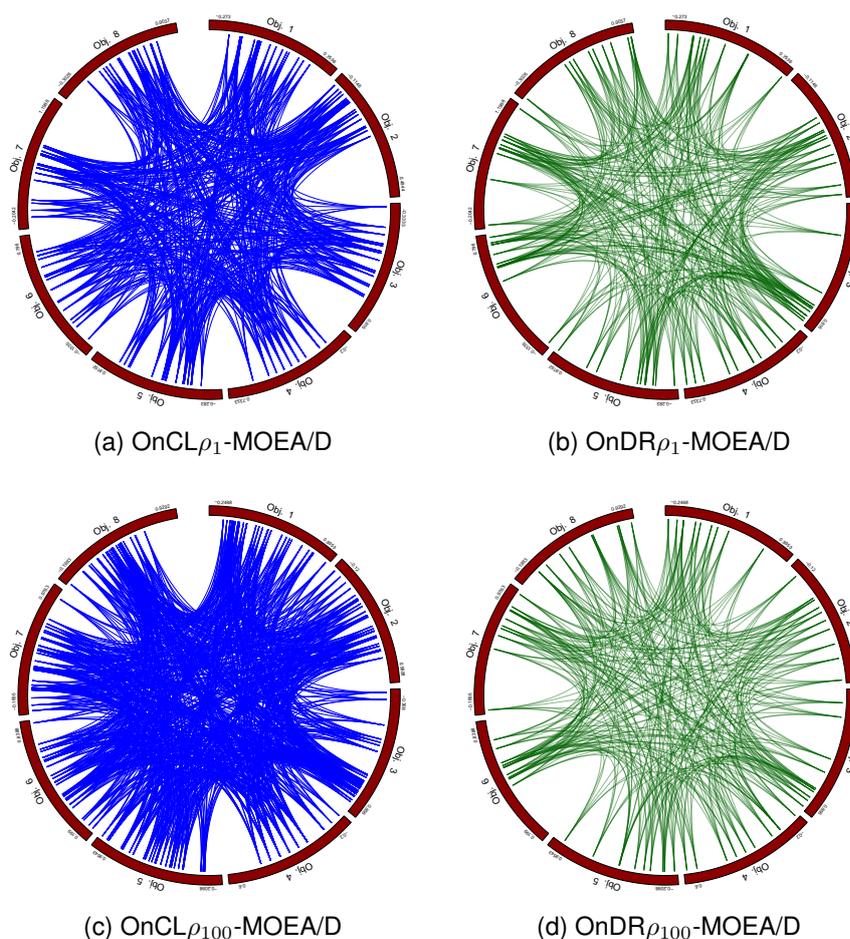


Figura 28 – Instância com 250 solicitações.

Os resultados obtidos nesta seção para as 30 execuções mostram que, independentemente da frequência da redução de dimensionalidade, a abordagem usando a agregação conseguiu melhores resultados em termos de convergência e qualidade das soluções obtidas quando comparados aos resultados obtidos pela abordagem usando seleção de atributos. Outro ponto importante que pôde ser observado é em relação à frequência da redução do número de objetivos que proporcionou tanto conjuntos Pareto combinados com melhores distribuições dos valores obtidos quanto um maior número de soluções não dominadas nas 5 instâncias testadas.

6 Conclusões e Trabalhos Futuros

Este trabalho apresentou uma formulação com muitos objetivos para o Problema de Roteamento de Veículos com Transporte Reativo a Demanda (PRVTRD). Em relação à formulação do problema, a função que avalia as requisições não atendidas foi removida e transformada em uma restrição de qualidade de serviço sendo um parâmetro de entrada no problema. Dessa forma, oito funções objetivo diferentes foram consideradas. A análise de *cluster online* baseada na Correlação de Pearson e no τ de Kendall para realizar a redução de dimensionalidade a cada geração do algoritmo MOEA/D, denominado OnCL $_{\tau}$ -MOEA/D, OnCL $_{\rho}$ -MOEA/D respectivamente, foi utilizada como forma de se reduzir a dimensionalidade do PRVTRD.

A primeira sequência de experimentos buscou avaliar o desempenho dos algoritmos propostos ao comparar as seguintes abordagens: (i) uma redução da dimensionalidade do *cluster offline* usando o coeficiente de correlação de Pearson; (ii) uma redução da dimensionalidade do *cluster offline* usando a correlação τ de Kendall; (iii) uma abordagem baseada em *cluster online* usando o coeficiente de correlação de Pearson; (iv) uma abordagem baseada em *cluster online* usando o coeficiente de correlação τ de Kendall; e (v) uma versão de *baseline*, um MOEA/D clássico.

Os algoritmos foram testados usando um conjunto de dados realistas contendo distâncias e tempos de viagem para a cidade de Belo Horizonte, usando 50, 100, 150, 200 e 250 solicitações. Para avaliar a dispersão das soluções obtidas pelos algoritmos, foi aplicado o indicador de hipervolume (HI), uma métrica de qualidade para conjuntos não dominados. Os algoritmos foram executados 30 vezes e os valores do indicador de hipervolume foram calculados em \mathbb{R}^8 .

Os resultados destes experimentos mostraram (i) que embora a diferença entre as versões *offline* dos algoritmos consistisse nos coeficientes de correlação usados na redução do *cluster*, a formulação bi-objetivo para o PRVTRD é exatamente a mesma para ambas as abordagens *offline*; (ii) os algoritmos baseados em *cluster online* (usando os coeficientes de correlação de Pearson e τ de Kendall) são estatisticamente melhores do que o algoritmo baseado em *cluster offline* em todas as instâncias testadas; (iii) não é possível afirmar que há diferença estatística entre os resultados dos algoritmos baseados em *cluster online* (algoritmos OnCL $_{\rho}$ -MOEA/D, OnCL $_{\tau}$ -MOEA/D) e o algoritmo MOEA/D quando o indicador de hipervolume é usado para avaliar a diversidade das soluções obtidas em todas as instâncias testadas.

Dessa forma, pode ser visto que o *cluster online* aplicado em conjunto com o algoritmo MOEA/D é uma técnica de redução de dimensionalidade eficaz que transforma

o problema de muitos objetivos em uma formulação agregada bi-objetivo para todas as instâncias testadas no PRVTRD. A redução da dimensionalidade tem um impacto favorável não só na eficiência no processo de busca dos algoritmos e no custo computacional, mas também em permitir a aplicação de técnicas usuais de visualização.

Embora não haja diferença estatística entre os resultados das abordagens *online*, a ferramenta de visualização diagrama de cordas foi aplicada para estimar efetivamente a extensão da área explorada e o desempenho qualitativo dos algoritmos de *cluster online*. Os resultados indicaram que a diversidade das soluções obtidas por OnCL $_{\rho}$ -MOEA/D e OnCL $_{\tau}$ -MOEA/D também são ligeiramente distintas quando aplicadas às instâncias.

A segunda sequência de experimentos realizados neste trabalho transforma o problema em uma formulação bi-objetivo para comparar duas abordagens de redução de dimensionalidade baseada em agregação e outra baseada em seleção de atributos. A primeira abordagem, igual a da primeira sequência de experimentos, agrega as funções objetivo em cada *cluster* obtido. Esta última elimina alguns objetivos usando técnicas não supervisionadas de seleção de atributos, denominadas Taxa de Dispersão (DR) e Máxima Variância (MV), para selecionar o representante do *cluster*. A frequência do impacto da redução da dimensionalidade no desempenho do algoritmo também foi analisada. As abordagens de redução de dimensionalidade foram acopladas ao algoritmo MOEA/D.

Três abordagens foram utilizados nos experimentos: (i) abordagem de *cluster online* usando o coeficiente de Pearson juntamente com MOEA/D: (OnCL $_{\rho_g}$ -MOEA/D); (ii) abordagem de *cluster online* usando o coeficiente de Pearson e a taxa de dispersão juntamente com MOEA/D (OnDR $_{\rho_g}$ -MOEA/D); (iii) abordagem de *cluster online* usando coeficiente de Pearson e Variância Máxima juntamente com MOEA/D: (OnMV $_{\rho_g}$ -MOEA/D). Os algoritmos foram testados nas mesmas instâncias dos experimentos iniciais e avaliados usando o indicador de hipervolume. Em seguida os resultados dos algoritmos foram comparados com o MOEA/D clássico no espaço objetivo original de oito dimensões.

Após as 30 execuções dos algoritmos, os resultados mostraram que, independentemente da frequência da abordagem de redução de dimensionalidade e das instâncias, os algoritmos MOEA/D e OnCL $_{\rho_g}$ -MOEA/D são estatisticamente superiores a OnDR $_{\rho_g}$ -MOEA/D e OnMV $_{\rho_g}$ -MOEA/D. Outro ponto que pôde ser observado é que não é possível afirmar que haja diferença estatística entre os resultados dos algoritmos baseados na seleção de atributos online em termos do indicador de hipervolume.

Os diagramas de cordas também foram utilizados para se avaliar, de forma qualitativa, o desempenho dos algoritmos *online* (usando agregação dos objetivos e seleção de atributos) como também a extensão da área explorada. Os resultados mostraram que há uma maior diversidade das soluções obtidas quando o algoritmo todas as funções objetivo são utilizadas na forma reduzida do problema e também quando os algoritmos permanecem por mais gerações em uma determinada formulação reduzida.

Durante o desenvolvimento desta tese, diversos questionamentos e oportunidades puderam ser encontrados. Com base nesses pontos levantados, algumas perspectivas e/ou iniciativas de pesquisa podem ser criadas uma vez que ainda há poucos trabalhos feitos na área bem como há oportunidades para diversas aplicações. Alguns dos pontos são listados abaixo como perspectivas de trabalhos futuros a serem explorados:

- Um estudo de diversidade das soluções obtidas e da velocidade de convergência para avaliar a influência dos coeficientes de Pearson e τ de Kendall no processo evolutivo;
- Fazer uma estratégia híbrida de frequências de redução como uma lista que armazene as últimas agregações realizadas em que a frequência em que foi feita a redução não possa se repetir em um determinado número de gerações do algoritmo;
- Usar o número de *clusters* variáveis, ou seja, deixar que a quantidade de *clusters* fosse definida de acordo com a região em que a população se encontra. Uma extensão dessa abordagem seria deixar que essa definição do número de *clusters* ocorresse dado um determinado número de gerações, como no caso das frequências testadas neste trabalho. Um exemplo de abordagem seria aplicar a técnica apresentada no trabalho de [Raykov et al. \(2016\)](#);
- Uso de técnicas de aprendizado profundo, como o caso de *autoencoders*, para descobrir uma representação latente do PRVTRD n -dimensional, com $n < 8$, com restrição dos valores dos pesos fazendo que a representação seja uma combinação linear dos objetivos do problema de modo que ocorra uma minimização dessa representação latente;
- Uso de técnicas de seleção de atributos supervisionadas. Uma vez que é necessário ter a associação entre o dado e uma classe, a ideia seria criar um critério para relacionar as funções objetivos avaliadas em uma solução com alguma classe, como por exemplo o caso de usar a dominância como a classe de cada solução;
- Aumentar o número de objetivos escolhidos pelas técnicas de seleção de atributos de modo que não ter apenas um objetivo por cluster, permitindo também ter números de objetivos diferentes em cada *cluster* e dessa forma reduzir a perda de informação;
- Explorar novas formas de medir a similaridade entre as funções objetivo, como no caso do cálculo da Divergência de Cauchy-Schwarz para medir a distância entre amostras, bem como o uso de outras técnicas de redução de similaridade (por exemplo: t-SNE);
- Analisar o uso de técnicas de decisão multicritério para identificar as soluções mais eficazes para serem aplicadas em uma situação real;
- Criação de uma técnica híbrida de agregação de objetivos usando o *cluster* hierárquico com as técnicas de seleção de atributos, de modo que em determinados momentos da busca ocorra a agregação e em outros ocorra a seleção, dando mais liberdade de escolha para o algoritmo.

Apêndices

APÊNDICE A – *Boxplots* e p -valores dos Testes de Aleatoriedade

Nesta seção, serão apresentados os *boxplots* em conjunto com os testes de aleatoriedade para os experimentos realizados e descritos na Seção 5.4. Inicialmente são apresentados os *boxplots* para cada uma das 5 instâncias testadas neste trabalho e em seguida os testes de aleatoriedade são apresentados em duas tabelas. Relembrando que há a divisão em dois grupos (a) e (b) sendo: grupo (a) correspondendo aos algoritmos MOEA/D, OnCL ρ_1 -MOEA/D, OnCL ρ_{100} -MOEA/D e o grupo (b) aos algoritmos OnDR ρ_1 -MOEA/D, OnDR ρ_{100} -MOEA/D, OnMV ρ_1 -MOEA/D e OnMV ρ_{100} -MOEA/D. Na Tabela 4 contém os p -valores para os testes par-a-par para cada algoritmo pertencente ao grupo (a) e a Tabela 5 os p -valores para as comparações do grupo (b).

A.1 *Boxplots*

Os *boxplots* apresentados foram feitos usando o indicador de hipervolume calculado usando o conjunto Pareto combinado após 30 execuções dos algoritmos. Em todas as figuras, pode ser visto dois grupos, um à esquerda contendo os algoritmos com agregação juntamente o MOEA/D e o outro à direita com os algoritmos que usam seleção de atributos.

A.1.1 *Boxplots* para Instância com 50 Solicitações

A Figura 29 apresenta os *boxplots* para as soluções não dominadas obtidas para os algoritmos usando agregação e seleção de atributos. Pode ser observado a sobreposição dos valores de hipervolume dentro do grupo (a) e dentro do grupo (b).

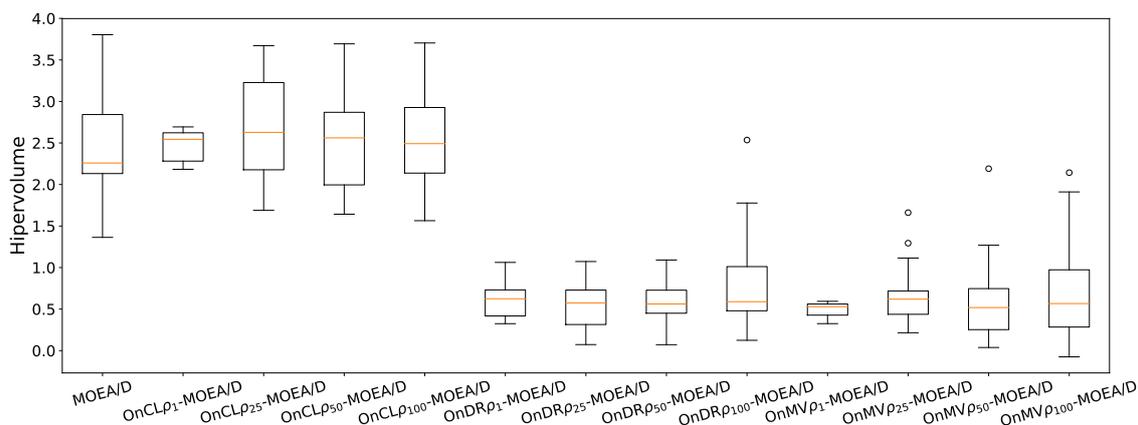


Figura 29 – *Boxplots* dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D e MOEA/D usando 50 solicitações com $g = 1, 25, 50,$ e 100 .

A.1.2 *Boxplots* para Instância com 100 Solicitações

Os *boxplots* para os valores do indicador de hipervolume obtidos após as 30 execuções dos algoritmos é apresentado na Figura 30. Percebe-se que não há sobreposição entre os grupos (a) e (b)

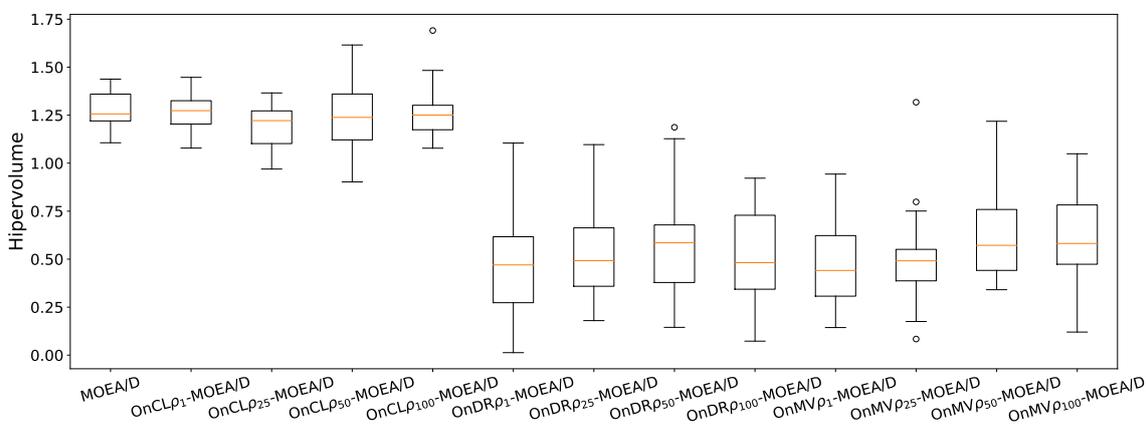


Figura 30 – *Boxplots* dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D e MOEA/D usando 100 solicitações com $g = 1, 25, 50,$ e 100 .

A.1.3 *Boxplots* para Instância com 150 Solicitações

A Figura 31 apresenta os valores obtidos para os indicadores de hipervolume dos conjuntos Pareto combinados após as execuções dos algoritmos. Novamente, pode ser visto a separação em dois grupos, um mais a esquerda e outro à direita, sendo o primeiro com desempenho superior.

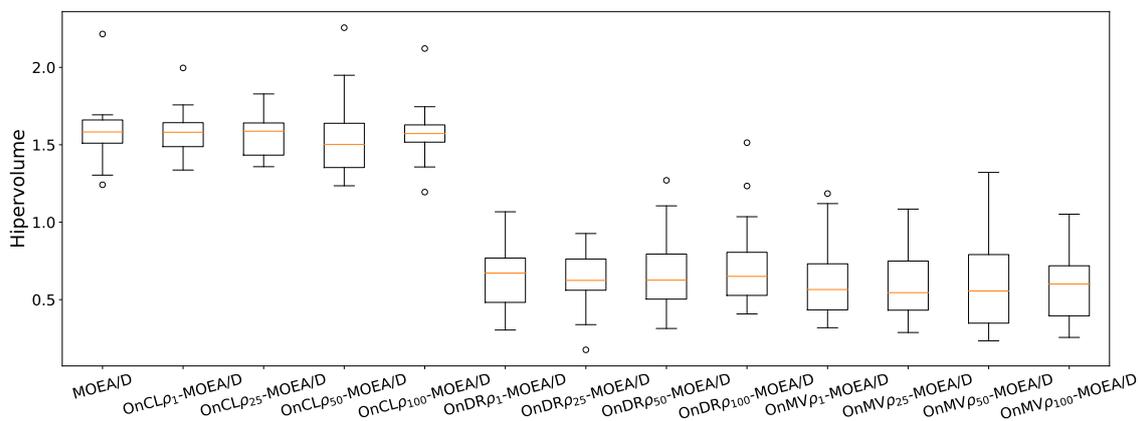


Figura 31 – *Boxplots* dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 150 solicitações com $g = 1, 25, 50, \text{ e } 100$.

A.1.4 *Boxplots* para Instância com 200 Solicitações

Os *boxplots* apresentados na Figura 32 foram feitos com base no hipervolume calculado para as soluções não dominadas obtidas pelos algoritmos para a instância de 200 solicitações. Com base na figura abaixo, pode ser visto que há intercessão dentro dos diagramas que pertencem ao grupo (a) e aos que pertencem ao grupo (b), porém não há intercessão entre os grupos.

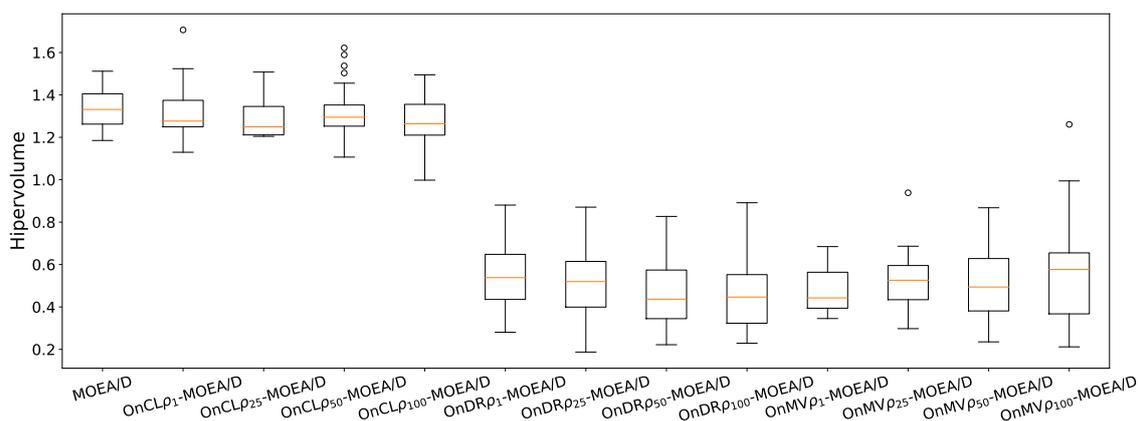


Figura 32 – *Boxplots* dos valores do indicador de hipervolume obtido amos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 200 solicitações com $g = 1, 25, 50, \text{ e } 100$.

A.1.5 *Boxplots* para Instância com 250 Solicitações

Na Figura 33 são apresentados os *boxplots* para a combinação dos conjuntos Pareto obtidos em cada execução dos algoritmos. Por não haver intercessão entre os diagramas

dos grupos (a) e (b), pode-se dizer que o primeiro grupo é estatsticamente superior em relação ao segundo.

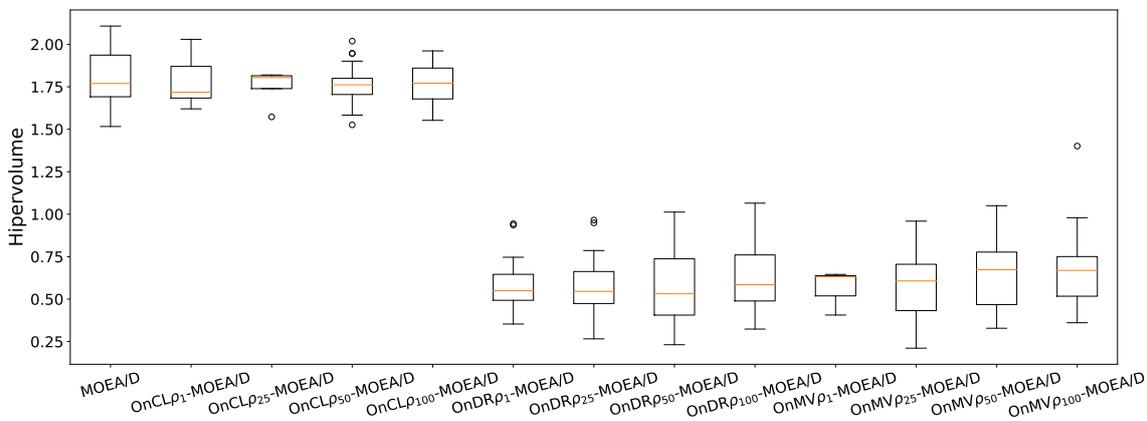


Figura 33 – *Boxplots* dos valores do indicador de hipervolume obtidos para 30 execuções para os algoritmos OnCL ρ_g -MOEA/D, OnDR ρ_g -MOEA/D, OnMV ρ_g -MOEA/D and MOEA/D usando 250 solicitações com $g = 1, 25, 50, e 100$.

A.2 Tabelas com Testes de Aleatoriedade

Uma vez que houve intercessão entre os *boxplots* apresentados na seção anterior, torna-se necessário aplicar algum teste estatístico para se detectar a existência ou não de diferença estatística entre os resultados obtidos. Conforme a Seção 5.4 o teste estatístico utilizado é teste de aleatoriedade e pode ser encontrado com maiores detalhes em Manly (2006). O nível de significância utilizado foi de 0.05.

Os testes foram divididos de acordo com os grupos (a) e (b) mencionados anteriormente. A Tabela 4 apresenta os p -valores obtidos ao se aplicar o teste de par-a-par. Com base nos valores obtidos, pode ser visto que, independentemente da frequência de redução, não há diferença estatística entre os algoritmos para todas as instâncias testadas.

Tabela 4 – p -valores para os testes de aleatoriedade para todos os pares de algoritmos do grupo (a) usando instâncias com 50, 100, 150, 200 e 250 reduções. O nível de significância é definido como 5%.

Comparação Par-a-Par	50	100	150	200	250
MOEA/D \times OnCL ρ_1 -MOEA/D	0.8994	0.5806	0.4386	0.3282	0.3186
MOEA/D \times OnCL ρ_{25} -MOEA/D	0.8440	0.4870	0.4438	0.6798	0.5934
MOEA/D \times OnCL ρ_{50} -MOEA/D	0.8804	0.5012	0.8564	0.2880	0.4090
MOEA/D \times OnCL ρ_{100} -MOEA/D	0.7622	0.5870	0.5848	0.2284	0.4280
OnCL ρ_1 -MOEA/D \times OnCL ρ_{25} -MOEA/D	0.4046	0.3140	0.5840	0.1526	0.8704
OnCL ρ_1 -MOEA/D \times OnCL ρ_{50} -MOEA/D	0.4482	0.4190	0.8114	0.5848	0.5818
OnCL ρ_1 -MOEA/D \times OnCL ρ_{100} -MOEA/D	0.3610	0.4158	0.5596	0.4366	0.5854
OnCL ρ_{25} -MOEA/D \times OnCL ρ_{50} -MOEA/D	0.5434	0.4612	0.2212	0.9704	0.6600
OnCL ρ_{25} -MOEA/D \times OnCL ρ_{100} -MOEA/D	0.4222	0.5190	0.4014	0.9092	0.2170
OnCL ρ_{50} -MOEA/D \times OnCL ρ_{100} -MOEA/D	0.4240	0.4710	0.7874	0.1794	0.9228

Os p -valores obtidos após a aplicação do teste de aleatoriedade para o grupo (b) são apresentados na Tabela 5. Os valores obtidos mostram que, para o nível de significância de 0.05 e dadas as frequências de redução de dimensionalidade, não há diferença estatística entre os algoritmos.

Tabela 5 – p -valores para os testes de aleatoriedade para todos os pares de algoritmos do grupo (b) usando instâncias com 50, 100, 150, 200 e 250 requisições. O nível de significância é definido como 5%.

Comparação Par-a-Par	50	100	150	200	250
OnMV ρ_1 -MOEA/D \times OnMV ρ_{25} -MOEA/D	0.5684	0.4686	0.7938	0.7964	0.3990
OnMV ρ_1 -MOEA/D \times OnMV ρ_{50} -MOEA/D	0.4258	0.8846	0.8020	0.5864	0.5860
OnMV ρ_1 -MOEA/D \times OnMV ρ_{100} -MOEA/D	0.5990	0.9010	0.8618	0.8734	0.6464
OnMV ρ_1 -MOEA/D \times OnDR ρ_1 -MOEA/D	0.5588	0.5820	0.9742	0.8068	0.9180
OnMV ρ_1 -MOEA/D \times OnDR ρ_{25} -MOEA/D	0.7702	0.6554	0.9332	0.9076	0.9024
OnMV ρ_1 -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.7558	0.8586	0.9282	0.5888	0.7494
OnMV ρ_1 -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.7580	0.7898	0.8800	0.5522	0.6416
OnMV ρ_{25} -MOEA/D \times OnMV ρ_{50} -MOEA/D	0.8284	0.3628	0.4114	0.3656	0.5924
OnMV ρ_{25} -MOEA/D \times OnMV ρ_{100} -MOEA/D	0.1946	0.9990	0.3678	0.6512	0.5900
OnMV ρ_{25} -MOEA/D \times OnDR ρ_1 -MOEA/D	0.5298	0.7136	0.5740	0.7396	0.7268
OnMV ρ_{25} -MOEA/D \times OnDR ρ_{25} -MOEA/D	0.9090	0.8090	0.4542	0.6614	0.6366
OnMV ρ_{25} -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.8944	0.9964	0.4566	0.4028	0.9228
OnMV ρ_{25} -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.3468	0.9468	0.5974	0.2432	0.9058
OnMV ρ_{50} -MOEA/D \times OnMV ρ_{100} -MOEA/D	0.5714	0.5060	0.4908	0.7302	0.4610
OnMV ρ_{50} -MOEA/D \times OnDR ρ_1 -MOEA/D	0.6260	0.2772	0.6872	0.7838	0.9932
OnMV ρ_{50} -MOEA/D \times OnDR ρ_{25} -MOEA/D	0.2412	0.3342	0.5422	0.7680	0.9686
OnMV ρ_{50} -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.2286	0.4638	0.5592	0.4420	0.5808
OnMV ρ_{50} -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.6360	0.3754	0.6704	0.4180	0.7252
OnMV ρ_{100} -MOEA/D \times OnDR ρ_1 -MOEA/D	0.6572	0.1244	0.7962	0.5640	0.9954
OnMV ρ_{100} -MOEA/D \times OnDR ρ_{25} -MOEA/D	0.1608	0.1730	0.6242	0.4960	0.9688
OnMV ρ_{100} -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.1534	0.4516	0.6358	0.2248	0.6588
OnMV ρ_{100} -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.4422	0.2474	0.5530	0.2674	0.7676
OnDR ρ_1 -MOEA/D \times OnDR ρ_{25} -MOEA/D	0.8432	0.5802	0.2512	0.4114	0.4342
OnDR ρ_1 -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.8122	0.8308	0.2524	0.3212	0.9014
OnDR ρ_1 -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.6756	0.7210	0.4096	0.1300	0.887
OnDR ρ_{25} -MOEA/D \times OnDR ρ_{50} -MOEA/D	0.4456	0.6870	0.1520	0.2382	0.9070
OnDR ρ_{25} -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.4924	0.5900	0.4432	0.2684	0.8840
OnDR ρ_{50} -MOEA/D \times OnDR ρ_{100} -MOEA/D	0.1122	0.2406	0.4946	0.5832	0.4226

APÊNDICE B – Diagramas de Corda

Os diagramas de corda [Koochaksaraei et al. \(2017\)](#) gerados para todos os conjuntos Pareto combinados obtidos após as execuções dos algoritmos serão apresentados nessa seção. Conforme mencionado no capítulo anterior, esta seção apresenta os diagramas para cada uma das frequências de redução de dimensionalidade testadas ($g = 1, 25, 50$ e 100). Os diagramas serão apresentados agrupados em 5 seções relativas às 5 instâncias usadas nos experimentos deste trabalho.

B.1 Diagramas de Corda para Resultados da Instância com 50 solicitações

A Figura 34 apresenta os diagramas de corda obtidos para os resultados na instância com 50 solicitações. Os diagramas foram feitos para todas as frequências de redução de dimensionalidade.

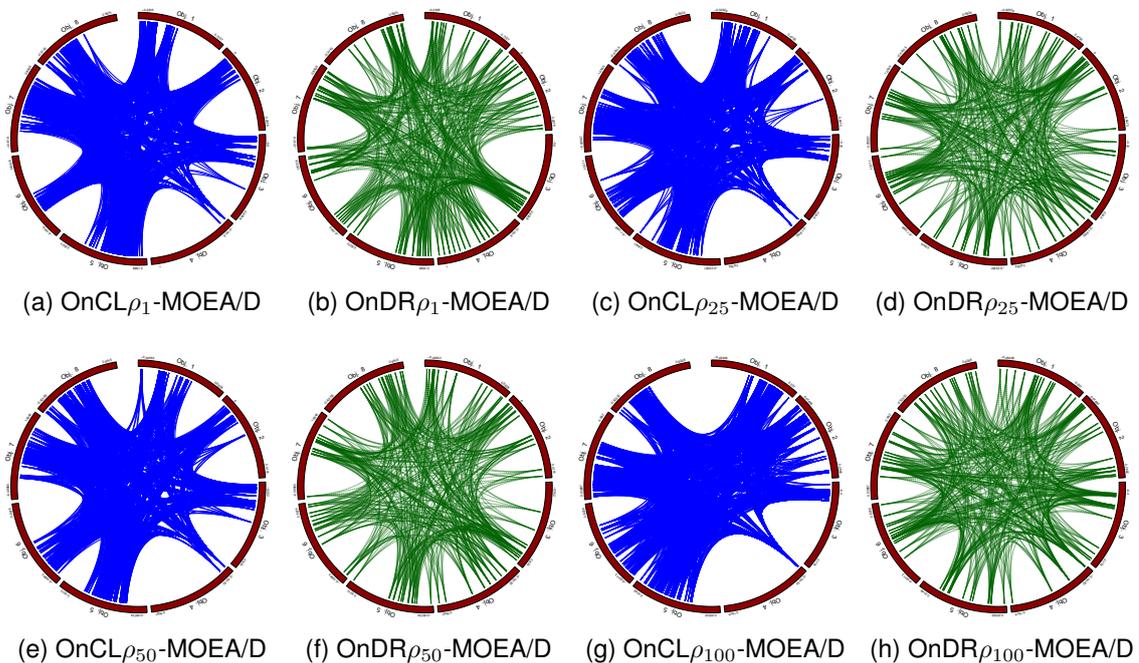


Figura 34 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 50 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).

B.2 Diagramas de Corda para Resultados da Instância com 100 solicitações

Os diagramas de corda para os conjuntos Pareto combinados após as execuções dos algoritmos na instância de 100 solicitações são apresentados na Figura 35. Os diagramas foram feitos para todas as frequências de redução de dimensionalidade.

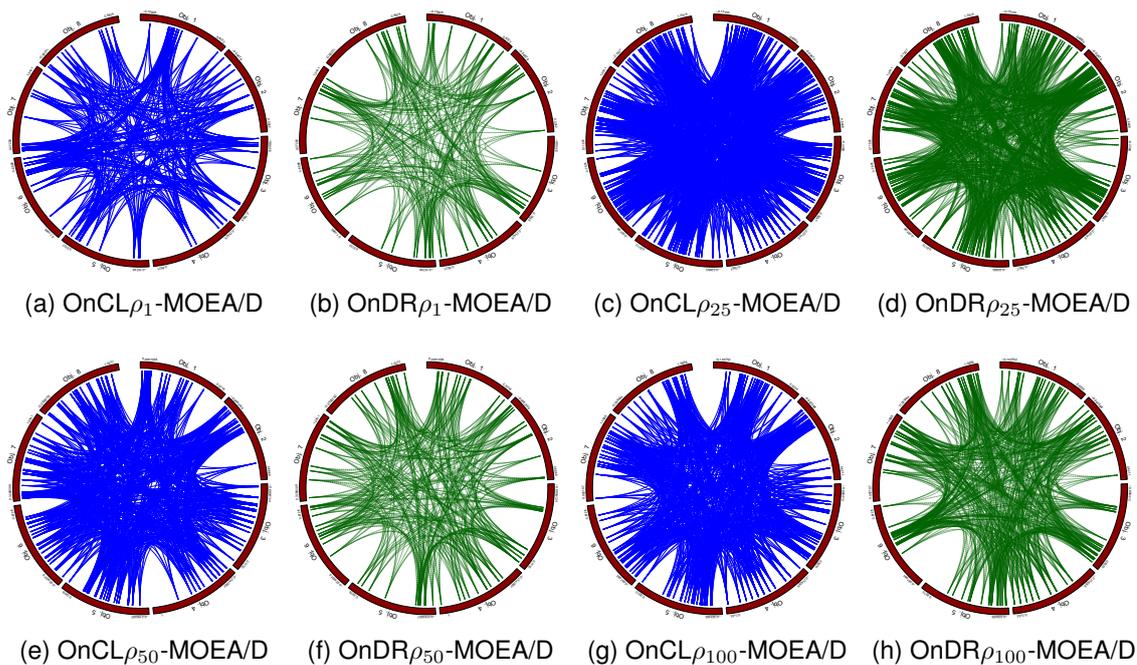


Figura 35 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 100 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).

B.3 Diagramas de Corda para Resultados da Instância com 150 solicitações

Na Figura 36 são apresentados os diagramas de corda para os resultados dos algoritmos na instância com 150 solicitações. Os diagramas foram feitos para todas as frequências de redução de dimensionalidade.

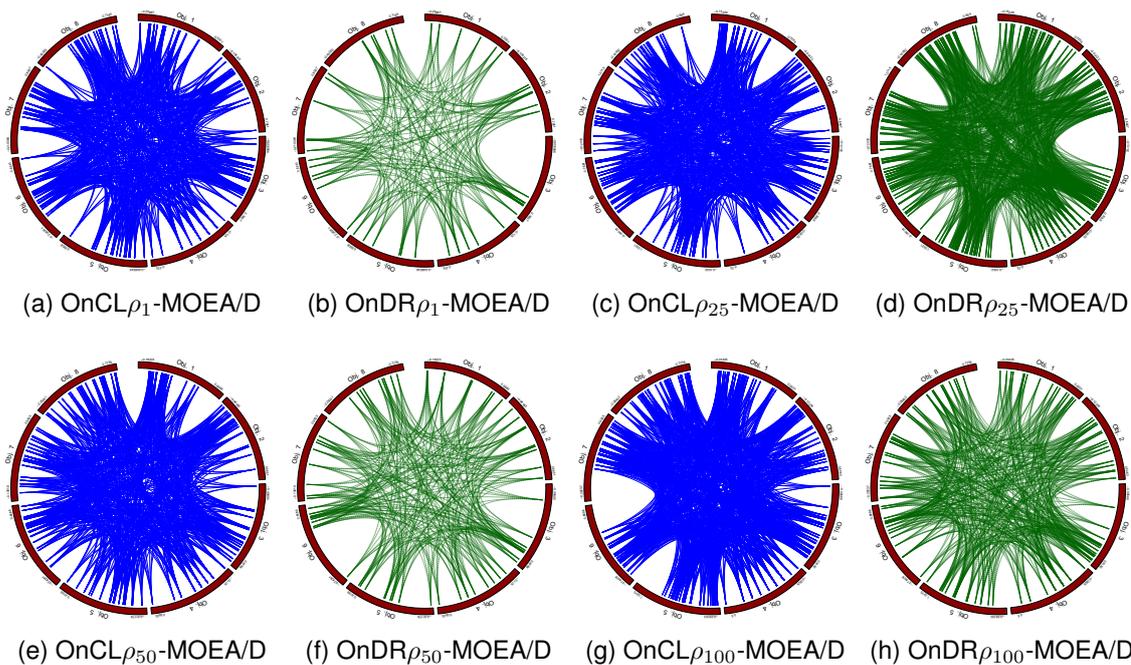


Figura 36 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 150 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).

B.4 Diagramas de Corda para Resultados da Instância com 200 solicitações

A Figura 37 apresenta os diagramas de corda obtidos para os resultados de cada algoritmo utilizado na instância com 200 solicitações. Os diagramas foram feitos para todas as frequências de redução de dimensionalidade.

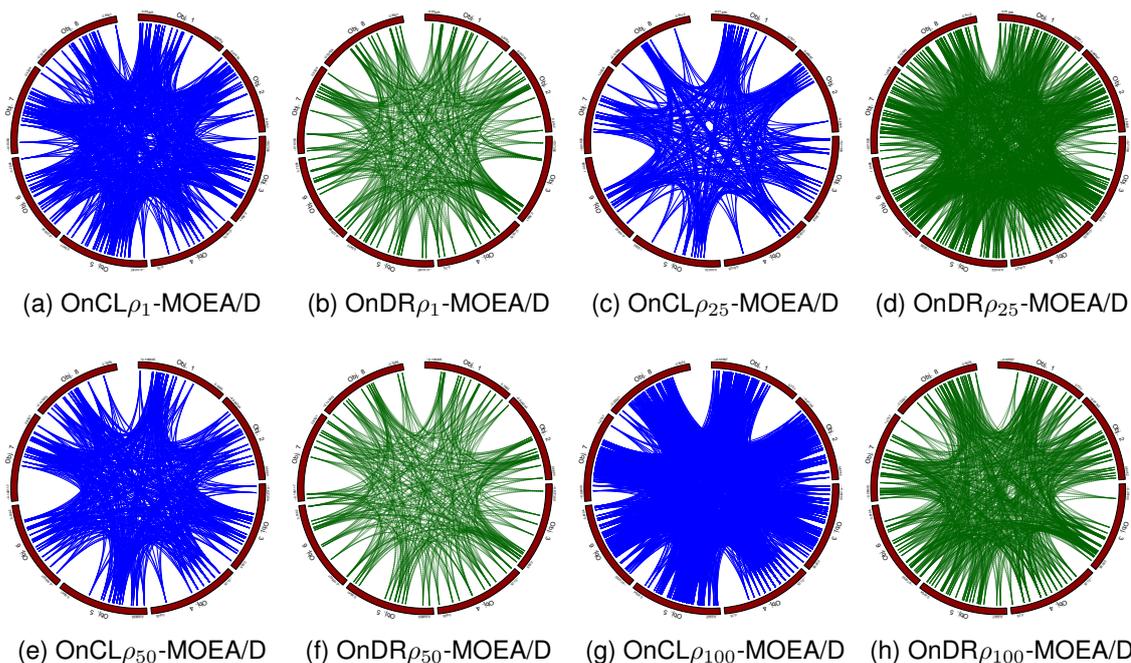


Figura 37 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 200 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).

B.5 Diagramas de Corda para Resultados da Instância com 250 solicitações

Os diagramas de corda para os resultados obtidos dos algoritmos na instância com 250 solicitações podem ser vistos na Figura 38. Os diagramas foram feitos para todas as frequências de redução de dimensionalidade.

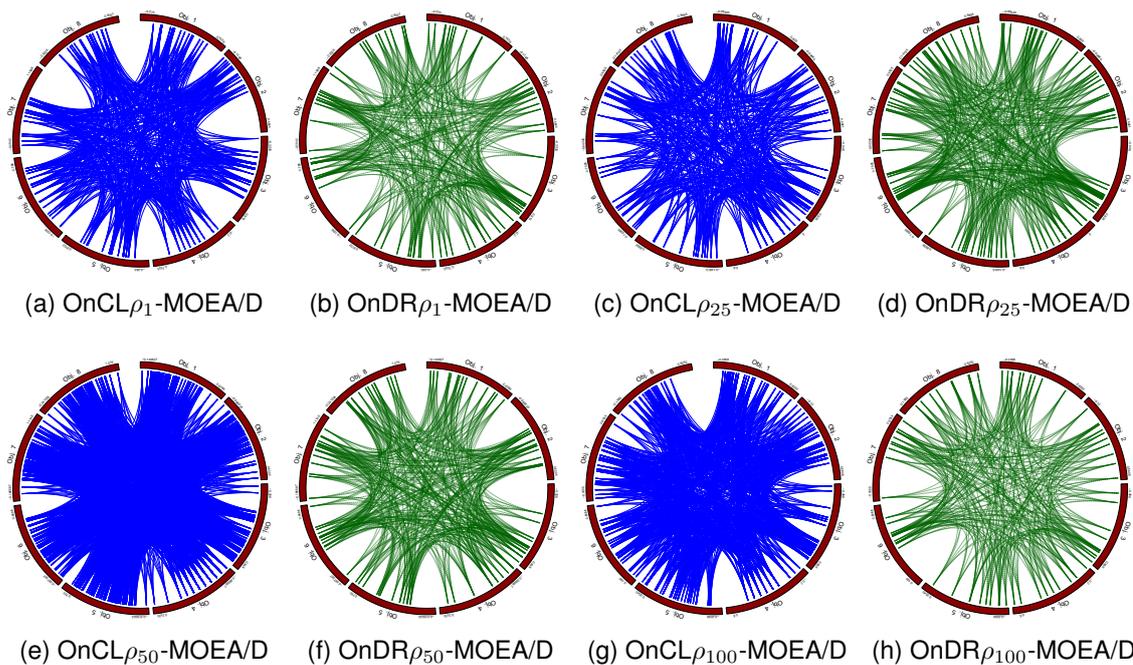


Figura 38 – Diagramas de corda para os resultados dos algoritmos baseados em agregação e seleção de atributos para a instância com 250 solicitações. São apresentados os diagramas para todas as frequências de redução de dimensionalidade ($g = 1, 25, 50$ e 100).

Referências

- ABDI, H. The Kendall Rank Correlation Coefficient. **Encyclopedia of Measurement and Statistics**, p. 508–510, 2007. ISSN 01621459. Citado na página 36.
- BADER, J.; ZITZLER, E. Hype: An algorithm for fast hypervolume-based many-objective optimization. **Evolutionary Computation**, v. 19, n. 1, p. 45–76, March 2011. ISSN 1063-6560. Citado 2 vezes nas páginas 29 e 31.
- BALCAN, M. F.; BLUM, A.; VEMPALA, S. A discriminative framework for clustering via similarity functions. **Proceedings of the 40th ACM Symposium on Theory of Computing**, ACM New York, NY, USA, 2008. Citado 2 vezes nas páginas 34 e 37.
- BREABAN, M. E.; IFTENE, A. Dynamic objective sampling in many-objective optimization. **Procedia Computer Science**, v. 60, p. 178 – 187, 2015. ISSN 1877-0509. Knowledge-Based and Intelligent Information and Engineering Systems 19th Annual Conference, KES-2015, Singapore, September 2015 Proceedings. Citado 2 vezes nas páginas 5 e 40.
- BRIDJ. 2014. Disponível em: <<http://www.bridj.com/>>. Citado na página 1.
- CAO, X. et al. Many-objective optimization of technology implementation in the industrial symbiosis system based on a modified nsga-iii. **Journal of Cleaner Production**, Elsevier Ltd, v. 245, 2 2020. ISSN 09596526. Citado na página 2.
- CARLSSON, C.; FULLER, R. Fuzzy multiple criteria decision making: Recent developments. **Fuzzy Sets and Systems**, v. 78, n. 2, p. 139 – 153, 1996. ISSN 0165-0114. Fuzzy Multiple Criteria Decision Making. Citado na página 32.
- CHE, J.; TANG, S. Research on integrated optimization design of hypersonic cruise vehicle. **Aerospace Science and Technology**, v. 12, n. 7, p. 567–572, 2008. ISSN 12709638. Citado na página 2.
- CHEUNG, Y. M.; GU, F.; LIU, H. L. Objective extraction for many-objective optimization problems: Algorithm and test problems. **IEEE Transactions on Evolutionary Computation**, v. 20, n. 5, p. 755–772, Oct 2016. ISSN 1089-778X. Citado na página 25.
- CHEVRIER, R. et al. Solving a dial-a-ride problem with a hybrid evolutionary multi-objective approach: Application to demand responsive transport. **Applied Soft Computing**, Sociedade Brasileira de Pesquisa Operacional, Bento Gonçalves, RS, Brasil, n. 12, p. 1247–1258, 2012. Citado 2 vezes nas páginas 16 e 17.
- CHIBA, K.; MAKINO, Y.; TAKATOYA, T. Evolutionary-based multidisciplinary design exploration for the silent supersonic technology demonstrator wing. **Journal of Aircraft**, v. 45, n. 5, p. 1481–1494, 2008. ISSN 15333868. Citado na página 2.
- CHIKUMBO, O.; GOODMAN, E.; DEB, K. Approximating a multi-dimensional pareto front for a land use management problem: A modified moea with an epigenetic silencing metaphor. In: **2012 IEEE Congress on Evolutionary Computation**. [S.l.: s.n.], 2012. p. 1–9. Citado na página 2.

COELLO, C. A. C. et al. Evolutionary multiobjective optimization: open research areas and some challenges lying ahead. **Complex & Intelligent Systems**, v. 6, p. 221–236, 2019. Citado 2 vezes nas páginas 5 e 22.

COELLO, C. A. C.; SIERRA, M. R. A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm. In: MONROY, R. et al. (Ed.). **MICAI 2004: Advances in Artificial Intelligence**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 688–697. ISBN 978-3-540-24694-7. Citado na página 30.

COELLO COELLO, C. A. et al. Evolutionary multiobjective optimization: open research areas and some challenges lying ahead. **Complex Intelligent Systems**, Springer International Publishing, v. 6, n. 2, p. 221–236, 2020. ISSN 2199-4536. Disponível em: <<https://doi.org/10.1007/s40747-019-0113-4>>. Citado na página 2.

CRUZ, A. R. et al. A multiobjective non-linear dynamic programming approach for optimal biological control in soy farming via nsga-ii. **IEEE Congress on Evolutionary Computation**, IEEE, 2007. Citado na página 20.

DEB, K. **Multi-Objective Optimization using Evolutionary Algorithms**. First. [S.l.]: WILEY, 2008. ISBN 978-0-470-74361-4. Citado 5 vezes nas páginas ix, 2, 21, 22 e 32.

DEB, K. et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. **IEEE Transactions on Evolutionary Computation**, v. 6, n. 2, p. 182–197, 2002. Citado 2 vezes nas páginas 3 e 4.

DEB, K.; JAIN, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with both constraints. **IEEE Computational Intelligence**, v. 18, n. 4, p. 577–601, 2014. Citado na página 31.

DEB, K.; SAXENA, D. K. **On finding Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems**. [S.l.], 2005. Citado 4 vezes nas páginas 23, 26, 27 e 39.

DEB, K. et al. Scalable test problems for evolutionary multiobjective optimization. In: _____. **Evolutionary Multiobjective Optimization: Theoretical Advances and Applications**. London: Springer London, 2005. p. 105–145. ISBN 978-1-84628-137-2. Disponível em: <https://doi.org/10.1007/1-84628-137-7_6>. Citado na página 27.

DENYSIUK, R.; COSTA, L.; SANTO, I. E. Clustering-based selection for evolutionary many-objective optimization. In: BARTZ-BEIELSTEIN, T. et al. (Ed.). **Parallel Problem Solving from Nature – PPSN XIII**. Cham: Springer International Publishing, 2014. p. 538–547. ISBN 978-3-319-10762-2. Citado 2 vezes nas páginas 39 e 40.

DURO, J. A. et al. Machine learning based decision support for many-objective optimization problems. **Neurocomputing**, v. 146, p. 30 – 47, 2014. ISSN 0925-2312. Bridging Machine learning and Evolutionary Computation (BMLEC) Computational Collective Intelligence. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231214008753>>. Citado 3 vezes nas páginas 2, 25 e 26.

DY, J. G.; BRODLEY, C. E. Feature selection for unsupervised learning. **Journal of Machine Learning Research**, JMLR.org, v. 5, p. 845–889, 2004. Citado 2 vezes nas páginas 19 e 42.

FANG, Y. et al. Evolutionary many-objective optimization for mixed-model disassembly line balancing with multi-robotic workstations. **European Journal of Operational Research**, Elsevier B.V., v. 276, p. 160–174, 7 2019. ISSN 03772217. Citado na página 2.

FARINA, M.; AMATO, P. On the optimal solution definition for many-criteria optimization problems. In: **2002 Annual Meeting of the North American Fuzzy Information Processing Society Proceedings. NAFIPS-FLINT 2002 (Cat. No. 02TH8622)**. [S.l.: s.n.], 2002. p. 233–238. Citado 2 vezes nas páginas ix e 24.

FEO, T. A.; RESENDE, M. Greedy randomized adaptive search procedure. **Journal of Global Optimization**, v. 2, n. 2, p. 109–133, 1995. Citado na página 16.

FREITAS, A. R. **Redução de Dimensionalidade em Problemas com Muitos Objetivos: Uma Aplicação em Composição Algorítmica**. Tese (Tese de doutorado) — Universidade Federal de Minas Gerais, 2013. Citado 2 vezes nas páginas 29 e 32.

FREITAS, A. R. R.; FLEMING, P. J.; GUIMARÃES, F. G. Aggregation trees for visualization and dimension reduction in many-objective optimization. **Information Sciences**, v. 298, n. 298, p. 288–314, 2015. Citado 2 vezes nas páginas 3 e 32.

FU, G. et al. Optimal Design of Water Distribution Systems Using Many-Objective Visual Analytics. **Journal of Water Resources Planning and Management**, v. 139, n. 6, p. 624–633, 2013. ISSN 0733-9496. Citado na página 2.

GOMES, J. R. **Dynamic Vehicle Routing For Demand Responsive Transportation Systems**. Tese (Tese de doutorado) — Universidade do Porto, Porto, Portugal, 2012. Citado 2 vezes nas páginas 8 e 17.

GOMES, R. J.; SOUZA, J. P.; DIA, T. G. A new heuristic approach for demand responsive transportation systems. In: **XLII Simpósio Brasileiro de Pesquisa Operacional (XLII SBPO)**. [S.l.: s.n.], 2010. p. 1839–1850. Citado na página 8.

GOMES, R. J.; SOUZA, J. P.; DIAS, T. G. A new heuristic approach for demand responsive transportation systems. **XLII Simpósio Brasileiro de Pesquisa Operacional (XLII SBPO)**, Sociedade Brasileira de Pesquisa Operacional, Bento Gonçalves, RS, Brasil, p. 1839–1850, 2010. Citado 5 vezes nas páginas 1, 8, 9, 13 e 16.

GOMES, R. J.; SOUZA, J. P.; DIAS, T. G. Design and operation of demand responsive transportation systems. **Procedia Social and Behavioral Sciences**, XV edition of the Euro Working Group on Transportation, Paris, França, p. 1–10, 2011. Citado 2 vezes nas páginas 8 e 17.

GOMES, R. J.; SOUZA, J. P.; DIAS, T. G. A grasp-based approach for demand responsive transportation. **International Journal of Transportation**, v. 2, n. 1, p. 21–32, 2014. Citado 2 vezes nas páginas 8 e 17.

GOULART, F.; CAMPELO, F. Preference-guided evolutionary algorithms for many-objective optimization. **Information Sciences**, Elsevier, v. 329, p. 236–255, 2016. Citado 2 vezes nas páginas 23 e 31.

GRABCOACH. 2017. Disponível em: <<https://www.grab.com/sg/coach/>>. Citado na página 1.

GUO, X. et al. A new objective reduction algorithm for many-objective problems: Employing mutual information and clustering algorithm. In: **2012 Eighth International Conference on Computational Intelligence and Security**. [S.l.: s.n.], 2012. p. 11–16. Citado 2 vezes nas páginas 39 e 40.

GUO, X.; WANG, Y.; WANG, X. Using objective clustering for solving many-objective optimization problems. **Mathematical Problems in Engineering**, v. 2013, 2013. Citado 2 vezes nas páginas 39 e 40.

HADKA, D. **MOEA Framework**. 2017. Disponível em: <<https://moeaframework.org/>>. Citado 3 vezes nas páginas ix, 25 e 27.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. New York, NY, USA: Springer New York Inc., 2001. (Springer Series in Statistics). Citado na página 41.

HE, Z.; YEN, G. G. Visualization and Performance Metric in Many-Objective Optimization. **IEEE Transactions on Evolutionary Computation**, IEEE, v. 20, n. 3, p. 386–402, 2016. ISSN 1089778X. Citado 2 vezes nas páginas 27 e 28.

HE, Z.; YEN, G. G.; DING, J. Knee-Based Decision Making and Visualization in Many-Objective Optimization. **IEEE Transactions on Evolutionary Computation**, v. 25, n. 2, p. 292–306, 2021. ISSN 19410026. Citado na página 28.

HONEST, N. A survey on feature selection techniques. **GIS Scienc Journal**, 09 2020. Citado na página 41.

HUANG, D. et al. A two-phase optimization model for the demand-responsive customized bus network design. **Transportation Research Part C: Emerging Technologies**, v. 111, p. 1–21, 2020. ISSN 0968-090X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0968090X1930395X>>. Citado na página 18.

IBRAHIM, A. et al. 3D-RadVis Antenna: Visualization and performance measure for many-objective optimization. **Swarm and Evolutionary Computation**, Elsevier, v. 39, n. July 2017, p. 157–176, 2018. ISSN 22106502. Disponível em: <<http://dx.doi.org/10.1016/j.swevo.2017.09.011>>. Citado na página 28.

IKEDA, K.; KITA, H.; KOBAYASHI, S. Failure of pareto-based moeas: does non-dominated really mean near to optimal? In: **Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546)**. [S.l.: s.n.], 2001. v. 2, p. 957–962 vol. 2. Citado 2 vezes nas páginas 24 e 25.

ISHIBUCHI, H.; AKEDO, A.; NOJIMA, Y. Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems. **IEEE Transactions on Evolutionary Computation**, v. 19, n. 2, p. 264–283, 2015. Citado na página 44.

ISHIBUCHI, H. et al. Behavior of EMO algorithms on many-objective optimization problems with correlated objectives. In: **2011 IEEE Congress of Evolutionary Computation (CEC)**. [S.l.: s.n.], 2011. p. 1465–1472. ISSN 1089-778X. Citado na página 30.

ISHIBUCHI, H. et al. Modified distance calculation in generational distance and inverted generational distance. In: GASPAR-CUNHA, A.; ANTUNES, C. H.; COELLO, C. C. (Ed.). **Evolutionary Multi-Criterion Optimization**. Cham: Springer International Publishing, 2015. p. 110–125. ISBN 978-3-319-15892-1. Citado na página 30.

ISHIBUCHI, H. et al. Evolutionary many-objective optimization by NSGA-II and MOEA/D with large populations. In: **Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics**. [S.l.]: IEEE Press, 2009. (SMC'09), p. 1758–1763. ISBN 9781424427932. Citado 2 vezes nas páginas 2 e 44.

ISHIBUCHI, I.; DOI, I.; NOJIMA, Y. On the effect of normalization in MOEA/D for multi-objective and many-objective optimization. **Complex Intelligent Systems**, v. 3, p. 279–294, 2017. Citado na página 44.

JAIMES, A. L. **Techniques to Deal with Many-objective Optimization Problems Using Evolutionary Algorithms**. Tese (Tese de doutorado) — Polytechnic Institute of Mexico, Cidade do México, México, 2011. Citado 6 vezes nas páginas ix, 20, 24, 25, 32 e 33.

JAIMES, A. L.; COELLO, C. A. C.; BARRIENTOS, J. E. U. Online objective reduction to deal with many-objective problems. In: _____. **Evolutionary Multi-Criterion Optimization: 5th International Conference, EMO 2009, Nantes, France, April 7-10, 2009. Proceedings**. [S.l.]: Springer International Publishing, 2009. p. 423–437. Citado 6 vezes nas páginas 5, 26, 27, 31, 39 e 40.

JAIMES, A. L.; COELLO, C. A. C.; CHAKRABORTY, D. Objective reduction using a feature selection technique. In: **Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation**. New York, NY, USA: Association for Computing Machinery, 2008. (GECCO '08), p. 673–680. ISBN 9781605581309. Disponível em: <<https://doi.org/10.1145/1389095.1389228>>. Citado na página 5.

JET, R. Multidisciplinary Design Optimization of Regional Jet Wing with Nacelle and Pylon. **Optimization**, n. January, 2003. Citado na página 2.

KENDALL, M. G. A new measure of rank correlation. **Biometrika**, [Oxford University Press, Biometrika Trust], v. 30, n. 1/2, p. 81–93, 1938. ISSN 00063444. Disponível em: <<http://www.jstor.org/stable/2332226>>. Citado na página 36.

KITA, H. et al. Multi-objective optimization by means of the thermodynamical genetic algorithm. In: VOIGT, H.-M. et al. (Ed.). **Parallel Problem Solving from Nature — PPSN IV**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996. p. 504–512. ISBN 978-3-540-70668-7. Citado na página 24.

KOOCHAKSARAEI, R. H. et al. A new visualization method in many-objective optimization with chord diagram and angular mapping. **Knowledge-Based Systems**, Elsevier, v. 138, p. 134–154, 2017. Citado 6 vezes nas páginas 4, 28, 29, 56, 63 e 79.

KOU, Y. N. et al. Many-objective optimization for coordinated operation of integrated electricity and gas network. **Journal of Modern Power Systems and Clean Energy**, Springer Heidelberg, v. 5, p. 350–363, 5 2017. ISSN 21965420. Citado na página 2.

LI, B. et al. Many-objective evolutionary algorithms: A survey. **ACM Comput. Surv.**, ACM, New York, NY, USA, v. 48, n. 1, p. 13:1–13:35, set. 2015. ISSN 0360-0300. Citado 5 vezes nas páginas 2, 5, 30, 31 e 39.

LI, L. et al. Preference-based evolutionary many-objective optimization for agile satellite mission planning. **IEEE Access**, Institute of Electrical and Electronics Engineers Inc., v. 6, p. 40963–40978, 7 2018. ISSN 21693536. Citado na página 2.

- LI, M.; YAO, X. Quality evaluation of solution sets in multiobjective optimisation: A survey. **ACM Comput. Surv.**, Association for Computing Machinery, New York, NY, USA, v. 52, n. 2, mar. 2019. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/3300148>>. Citado 2 vezes nas páginas 29 e 52.
- LIU, Q. et al. Many-objective best-order-sort genetic algorithm for mixed-model multi-robotic disassembly line balancing. In: . [S.l.]: Elsevier B.V., 2019. v. 83, p. 14–21. ISSN 22128271. Citado na página 2.
- LIU, R.; WANG, N. Data-driven bus route optimization algorithm under sudden interruption of public transport. **IEEE Access**, v. 10, p. 5250–5263, 2022. Citado na página 18.
- LIU, T.; CHEN, Y.; WU, M. An active feature selection strategy for dwt in artificial taste. **Journal of Sensors**, 2018. Citado 3 vezes nas páginas 19, 41 e 42.
- LOPEZ, A. J.; COELLO, C. A. C.; CHAKRABORTY, D. Objective reduction using a feature selection technique. In: **Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation**. New York, NY, USA: ACM, 2008. (GECCO '08), p. 673–680. ISBN 978-1-60558-130-9. Citado 3 vezes nas páginas 5, 39 e 40.
- LUCKEN, C. V.; BARÁN, B.; BRIZUELA, C. A survey on multi-objective evolutionary algorithms for many-objective problems. **Computational Optimization and Applications**, Springer, v. 58, p. 707–756, 2014. Citado na página 2.
- MAGEEAN, J.; NELSON, J. D. The evaluation of demand responsive transport services in europe. **Journal of Transport Geography**, Pergamon, v. 11, p. 255–270, 2003. Citado na página 2.
- MALTESE, J.; OMBUKI-BERMAN, B. M.; ENGELBRECHT, A. P. A Scalability Study of Many-Objective Optimization Algorithms. **IEEE Transactions on Evolutionary Computation**, IEEE, v. 22, n. 1, p. 79–96, 2018. ISSN 1089778X. Citado na página 29.
- MANLY, B. F. **Randomization, bootstrap and Monte Carlo methods in biology**. [S.l.]: CRC Press, 2006. v. 70. Citado 2 vezes nas páginas 54 e 77.
- MARQUES, J.; CUNHA, M.; SAVIĆ, D. Many-objective optimization model for the flexible design of water distribution networks. **Journal of Environmental Management**, Academic Press, v. 226, p. 308–319, 11 2018. ISSN 10958630. Citado na página 2.
- MELIS, L.; SORENSEN, K. The static on-demand bus routing problem: large neighborhood search for a dial-a-ride problem with bus station assignment. **International Transactions in Operational Research**, v. 29, n. 3, p. 1417–1453, 2022. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/itor.13058>>. Citado na página 18.
- MENDES, R. S. **Abordagem Multiobjetivo para o Problema de Roteamento de Veículos com Transporte Reativo a Demanda**. Dissertação (Master Thesis) — Centro Federal de Educação Tecnológica de Minas Gerais, 2016. Citado na página 46.
- MENDES, R. S. et al. Online clustering reduction based on parametric and non-parametric correlation for a many-objective vehicle routing problem with demand responsive transport. **Expert Systems with Applications**, v. 170, p. 114467, 2021. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417420311180>>. Citado 4 vezes nas páginas 15, 19, 51 e 63.

- MENDES, R. S. et al. Multiobjective approach to the vehicle routing problem with demand responsive transport. In: **Proceedings of IEEE Congress on Evolutionary Computation**. [S.l.: s.n.], 2016. Citado 3 vezes nas páginas 3, 18 e 40.
- MENDES, R. S. et al. Dimensionality reduction approach for many-objective vehicle routing problem with demand responsive transport. In: **Proceedings of International Conference on Evolutionary Multi-Criterion Optimization**. [S.l.]: Springer, 2017. p. 438–452. Citado 4 vezes nas páginas 3, 15, 16 e 19.
- MENDES, R. S. et al. Dimensionality reduction approach for many-objective vehicle routing problem with demand responsive transport. In: _____. **Evolutionary Multi-Criterion Optimization: 9th International Conference, EMO 2017, Münster, Germany, March 19-22, 2017, Proceedings**. [S.l.]: Springer International Publishing, 2017. p. 438–452. Citado na página 40.
- MENDES, R. S. et al. Aggregation or selection? Clustering many objectives for vehicle routing problem with demand responsive transport. In: **Proceedings of IEEE Congress on Evolutionary Computation**. [S.l.: s.n.], 2021. Citado 2 vezes nas páginas 19 e 60.
- MENDES, R. S. et al. Optimization of the vehicle routing problem with demand responsive transport using the NSGA-II algorithm. **Proceedings of IEEE Intelligent Transportation Systems Society Conference Management System**, IEEE, 2016. Citado 3 vezes nas páginas 3, 19 e 40.
- MENEGHINI, I. R.; KOOCHAKSARAEI, R. H.; GUIMARÃES, F. G. **Data Visualization in Many Objective Optimization Problems**. 2014. Accessed: Oct 30th, 2021. Citado na página 56.
- MILLER, G. A. The magical number seven, plus or minus two some limits on our capacity for processing information. *Psychological Review*, v. 101, p. 343–352, 03 1955. Citado na página 25.
- MIN, Y. et al. Unsupervised feature selection via multi-step markov transition probability. **arXiv**, 2020. Citado na página 41.
- MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada - Uma Abordagem Aplicada**. [S.l.]: Editora UFMG, 2005. Citado 3 vezes nas páginas 3, 33 e 36.
- MIRANDA, D. S. **Aplicação de metaheurísticas para o problema de roteamento de veículos dinâmico para o transporte reativo a demanda**. Dissertação (Dissertação de mestrado) — Universidade Federal de Viçosa, Viçosa, 2012. Citado 2 vezes nas páginas 12 e 16.
- MOIA. **New Mobility Modes**. 2017. Disponível em: <<https://www.volkswagenag.com/en/brands-and-models/moia.html>>. Citado na página 1.
- MOSTAGHIM, S.; SCHMECK, H. Distance based ranking in many-objective particle swarm optimization. In: RUDOLPH, G. et al. (Ed.). **Parallel Problem Solving from Nature – PPSN X**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 753–762. ISBN 978-3-540-87700-4. Citado na página 23.

- MULLEY, C.; NELSON, J. D. Flexible transport services: A new market opportunity for public transport. **Research in Transportation Economics**, Elsevier, v. 25, p. 39–45, 2009. Citado na página 1.
- NISBETT, R.; WILSON, T. Telling more than we can know: Verbal reports on mental processes. v. 84, p. 231–259, 03 1977. Citado na página 25.
- PAL, M.; SAHA, S.; BANDYOPADHYAY, S. DECOR: Differential evolution using clustering based objective reduction for many-objective optimization. **Information Sciences**, v. 423, p. 200 – 218, 2018. ISSN 0020-0255. Citado na página 40.
- PAL, S. K.; MITRA, P. **Pattern Recognition Algorithms for Data Mining: Scalability, Knowledge Discovery, and Soft Granular Computing**. GBR: Chapman and Hall, Ltd., 2004. ISBN 1584884576. Citado na página 41.
- RAJA, B. D.; JHALA, R. L.; PATEL, V. Many-objective optimization of cross-flow plate-fin heat exchanger. **International Journal of Thermal Sciences**, Elsevier Masson SAS, v. 118, p. 320–339, 8 2017. ISSN 12900729. Citado na página 2.
- RAYKOV, Y. et al. What to do when k-means clustering fails: A simple yet principled alternative algorithm. **PLOS ONE**, v. 11, p. e0162259, 09 2016. Citado na página 71.
- ROY, S. et al. Dispersion ratio based decision tree model for classification. **Expert Systems with Applications**, v. 116, p. 1–9, 2019. Citado 3 vezes nas páginas 19, 41 e 42.
- SAFI, H. H.; UCAN, O. N.; BAYAT, O. On the real world applications of many-objective evolutionary algorithms. In: . [S.I.]: Association for Computing Machinery, 2018. ISBN 9781450365369. Citado na página 2.
- SCHYNS, M. An ant colony system for responsive dynamic vehicle routing. **European Journal of Operations Research**, Elsevier, v. 245, p. 704–718, 2015. Citado na página 17.
- SEKIZAKI, S.; NISHIZAKI, I.; HAYASHIDA, T. Distribution network management under electricity deregulation using evolutionary many-objective optimization. **Journal of Multi-Criteria Decision Analysis**, John Wiley and Sons Ltd, v. 25, p. 17–30, 1 2018. ISSN 10991360. Citado na página 2.
- SEN, P.; YANG, J. **Multiple Criteria Decision Support in Engineering Design**. Springer London, 1998. ISBN 9783540199328. Disponível em: <<https://books.google.com.br/books?id=LlxRAAAAMAAJ>>. Citado na página 26.
- SOLER, J. et al. Data clustering and similarity. **Proceedings of the Twenty-Sixth International Florida Artificial Intelligence Research Society Conference**, Artificial Intelligence Research Society, p. 492–495, 2013. Citado na página 34.
- SOLOMON, B. **Toyota And Uber Join Forces In Ride-Sharing Sprint**. 2017. Citado na página 1.
- TAN, P. N.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining**. [S.I.]: Pearson, 2005. Citado 3 vezes nas páginas 33, 34 e 37.
- TUSAR, T.; FILIPIC, B. Visualization of pareto front approximations in evolutionary multi-objective optimization: A critical review and the projection method. **IEEE Transactions on Evolutionary Computation**, v. 19, n. 2, p. 225–245, April 2015. ISSN 1089-778X. Citado na página 56.

- UBER. 2009. Disponível em: <<https://www.uber.com/en-MX/>>. Citado na página 1.
- VELAGA, N. R. et al. Development of an integrated flexible transport systems platform for rural areas using argumentation theory. **Research in Transportation Business & Management**, Elsevier, v. 3, p. 62–70, 2012. Citado na página 1.
- VIANA, R. J. S.; SANTOS, A. G. Dynamic approach for a demand responsive transport. In: **Proceedings of XLI Latin American Computing Conference (CLEI)**. [S.l.]: IEEE, 2015. p. 1–12. Citado na página 8.
- VIANA, R. J. S.; SANTOS, A. G. Dynamic approach for a demand responsive transport. **XLI Latin American Computing Conference (CLEI)**, IEEE, p. 1–12, 2015. Citado na página 17.
- VIANA, R. J. S.; SANTOS, A. G. S.; ARROYO, J. E. C. Multi-objective evolutionary approach for optimizing a demand responsive transports. **XLVII Simpósio Brasileiro de Pesquisa Operacional (XLII SBPO)**, Sociedade Brasileira de Pesquisa Operacional, Porto de Galinhas, PE, Brasil, 2015. Citado na página 17.
- WATANABE, T.; TATSUKAWA, T.; OYAMA, A. On the fast hypervolume calculation method. **2015 IEEE Congress on Evolutionary Computation, CEC 2015 - Proceedings**, IEEE, p. 965–969, 2015. Citado 2 vezes nas páginas 30 e 52.
- WEGMAN, E. J. Hyperdimensional data analysis using parallel coordinates. **Journal of the American Statistical Association**, American Statistical Association, Taylor and Francis, Ltd., v. 85, n. 411, p. 664–675, 1990. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2290001>>. Citado na página 27.
- WHILE, L. et al. A faster algorithm for calculating hypervolume. **IEEE Transactions on Evolutionary Computation**, v. 10, n. 1, p. 29–38, 2006. Citado na página 29.
- XIONG, J. et al. Demand responsive service-based optimization on flexible routes and departure time of community shuttles. **Sustainability**, v. 12, n. 3, 2020. ISSN 2071-1050. Disponível em: <<https://www.mdpi.com/2071-1050/12/3/897>>. Citado na página 18.
- ZHANG, Q.; LI, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. **IEEE Transactions on Evolutionary Computation**, v. 11, n. 6, p. 712–731, Dec 2007. ISSN 1089-778X. Citado 3 vezes nas páginas 4, 19 e 44.
- ZHOU, A.; WANG, Y.; ZHANG, J. Objective extraction via fuzzy clustering in evolutionary many-objective optimization. **Information Sciences**, Elsevier Inc., v. 509, p. 343–355, 2020. ISSN 00200255. Disponível em: <<https://doi.org/10.1016/j.ins.2018.11.032>>. Citado na página 29.
- ZITZLER, E.; DEB, K.; THIELE, L. Comparison of multiobjective evolutionary algorithms: Empirical results. **Evolutionary Computing**, v. 8, n. 2, 2000. Citado na página 2.
- ZITZLER, E.; THIELE, L. Multiobjective optimization using evolutionary algorithms — a comparative case study. In: EIBEN, A. E. et al. (Ed.). **Parallel Problem Solving from Nature — PPSN V**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. p. 292–301. ISBN 978-3-540-49672-4. Citado na página 52.

ZITZLER, E.; THIELE, L. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. **Evolutionary Computation**, IEEE, Switzerland, p. 257–271, 1999. Citado na página 2.

ZITZLER, E.; THIELE, L. Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. **IEEE Transactions on Evolutionary Computation**, v. 3, n. 4, p. 257–271, 1999. Citado na página 52.

ZOGRAFOS, K. G.; ANDROUTSOPOULOS, K. N.; SIHVOLA, T. A methodological approach for developing and assessing business models for flexible transport systems. **Transportation**, v. 35, p. 777–795, 2008. Citado na página 1.