

Proposta de algoritmo transgenético híbrido para o problema da Árvore Geradora Multiobjetivo baseada no operador OWA

Prof^a. Dr^a. Elizabeth Ferreira Gouvêa Goldberg
Islame Felipe da Costa Fernandes

Departamento de Informática e Matemática Aplicada - DIMAp
Universidade Federal do Rio Grande do Norte - UFRN

Natal, 14 novembro de 2017

Agenda

- 1 Introdução
- 2 Algoritmo transgenético
- 3 Experimentos computacionais

Introdução

- ① O objetivo desta pesquisa é:
 - ① propor uma abordagem memética *M-SA* e uma transgenética *T-SA* aplicada ao problema da *OWA-ST*;
 - ② comparar o desempenho do *M-SA* e do *T-SA* com o melhor exato da literatura (a saber, o modelo matemático de Fernández et al. (2017));
 - ③ comparar o desempenho das meta-heurísticas entre si, em termos de qualidade de solução e tempo de processamento;
 - ④ demonstrar que o desempenho da transgenética computacional, na grande maioria dos casos, é superior a do memético.
- ② Serão submetidas a experimentos computacionais 270 instâncias com 8 objetivos, possuindo de 30 a 1000 vértices, classificadas segundo critérios *k-trimmed*, *k-centrum* e *Hurwicz* e classes *correlated* e *anticorrelated*.

Algoritmo transgenético

- ① Cria a população P de $\#popSize$ endossimbiontes;
- ② Hospedeiro inicialmente vazio;
- ③ Copia o DNA dos endossimbiontes no hospedeiro;
- ④ Para cada geração ($\#mas_gen$):
 - ① Cria conjunto de plasmídeos;
 - ② Plasmídeos são sorteados para atacar cada endossimbionte, inclusive o ótimo corrente
 - ③ $primTrans$ tenta atacar com probabilidade $\#probTrans1$;
 - ④ SA_Trans tenta atacar com probabilidade $\#probTrans2$;
 - ⑤ Caso um vetor qualquer consiga atacar um endossimbionte $p \in P$, então o DNA original de p é inserido no hospedeiro

População inicial

- Idêntico ao *M-SA*
- Representação do endossimbionte: lista de arestas (RAIDL; JULSTROM, 2003)
- Tamanho: $\#popSize$
- Método: *rmcPrim* (KNOWLES, 2002)
- Vetores de escalarização ω para o *rmcPrim*:
 - ▶ SPEA/R (JIANG; YANG, 2017)
 - ▶ *k-centrum* : $k \in [1, M]$, $\omega_1 = \dots = \omega_k = 1/k$ e $\omega_{k+1} = \dots = \omega_M = 0$.
 - ▶ *k-trimmed* : $k \in [1, \frac{M}{2}]$, $\omega_1 = \dots = \omega_k = 0$,
 $\omega_{k+1} = \dots = \omega_{M-k} = \frac{1}{M-2k}$ e $\omega_{M-k+1} = \dots = \omega_M = 0$
 - ▶ *Hurwicz* : $\alpha \in \{0.2, 0.3, 0.4, 0.6, 0.7, 0.8\}$, $\omega_1 = \alpha$,
 $\omega_2 = \dots = \omega_{M-1} = 0$ e $\omega_M = 1 - \alpha$.

Hospedeiro

- Estratégia: *a posteriori*;
- Tamanho: $3 * \#popSize$;
- Sempre que o ataque de um vetor transgenético é bem sucedido, o hospedeiro recebe o DNA original do indivíduo atacado;
- O objetivo é, ao invés de se desfazer completamente da cadeia atacada, guardá-la e utilizar seu material genético posteriormente;
- Caso o hospedeiro esteja lotado no momento da inserção de uma nova cadeia, então a cadeia de maior *fitness* é excluída do hospedeiro, a fim de dar lugar à nova;
- O hospedeiro, neste caso, caracteriza-se como um repositório elitista de cadeias de DNA recicláveis.

Plasmídeo 1 - *plas1*

- Proposto por Monteiro, Goldbarg e Goldbarg (2010);
- A informação genética deste plasmídeo recombinado advém do *rmcPrim*;
- O *rmcPrim* é executado para obter p_{size} arestas, formando uma subárvore

Plasmídeo 2 - *plas2*

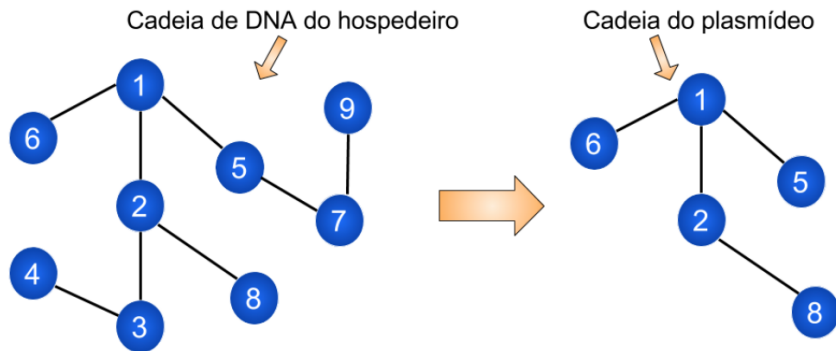


Figura 1: Esquematização do *plas2*

Plasmídeo 3 - *plas3*

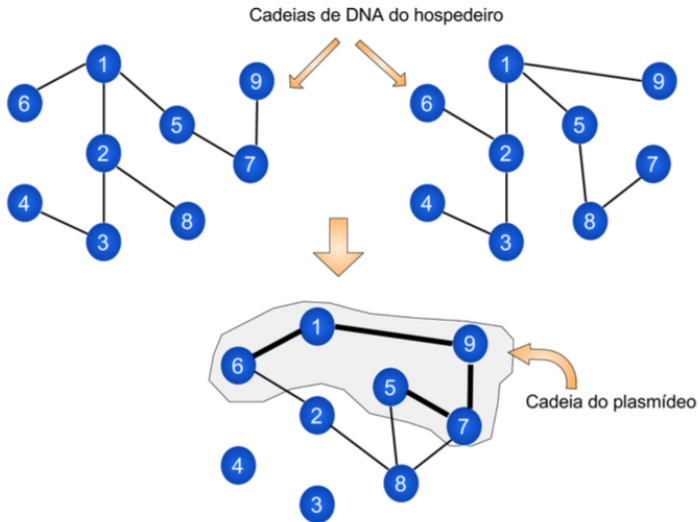


Figura 2: Esquematização do *plas3*

Transponson - *primTrans*

- Proposto por Monteiro, Goldberg e Goldberg (2010);
- A regra do *primTrans* consiste em, dado um endossimbionte $p \in P$, remover algumas arestas de p e completar a solução com o procedimento *rmcPrim* (KNOWLES, 2002);
- São removidas de 90% a 95% arestas;
- A escolha da aresta a ser removida depende do custo escalarizado obtido pela aplicação de um vetor de escalarização.

Operadores - transponson SA_trans

Algoritmo 2: *Simulated Annealing*

Entrada: sol: solução, T_0 : temperatura inicial, L_0 : quantidade inicial de iterações

```
1   $s = sol$ ;  
2   $T_i = T_0$ ;  
3   $L_i = L_0$ ;  
4  enquanto  $T_i > 0$  faça  
5      para  $i=1 \dots L_i$  faça  
6          escolha randomicamente dois pares de aresta  $(v_l, v_{l+1})$  e  $(v_j, v_{j+1})$  em  $s$ ;  
7          remova  $(v_l, v_{l+1})$  e  $(v_j, v_{j+1})$  de  $s$ ;  
8          religue os vértices  $v_l, v_{l+1}, v_j, v_{j+1}$  e produza  $s'$ ;  
9           $\Delta C = OWA(s') - OWA(s)$ ;  
10         se  $\Delta C \leq 0$  então  
11              $s = s'$ ;  
12             se  $OWA(s) < OWA(sol)$  então  
13                  $sol = s$ ;  
14             fim  
15         senão  
16             se  $random(0, 1) < \exp\{-\frac{\Delta C}{T_i}\}$  então  
17                  $s = s'$   
18             fim  
19         fim  
20     fim  
21      $T_i = \frac{T_i}{F_T}$ ;  
22      $L_i = L_i * F_L$ ;  
23 fim
```

Instâncias

- ① 270 instâncias
- ② 30 a 1000 vértices
- ③ 8 objetivos
- ④ 3 grupos de instâncias: $n.1$, $n.2$, $n.3$
- ⑤ Classes: *correlated* e *anti-correlated* (KNOWLES, 2002)
 - ① *correlated* Coeficientes dos respectivos grupos: 0.2, 0.5 e 0.85
 - ② *anti-correlated* Coeficientes dos respectivos grupos: -0.2, -0.5 e -0.85
- ⑥ Três critérios OWA: *k-trimmed*, *k-centrum*, *Hurwicz*.

Parâmetros - IRACE

Tabela 1: Parâmetros do *M-SA*

$\#popSize = 100$	$\#propCross = 96\%$
$\#max_gen = 48$	$\#propMutation = 18\%$

Tabela 2: Parâmetros do *T-SA*

$\#popSize = 100$	$\#numPlasTotal = 19$
$\#max_gen = 48$	$\#numPlas1 = 4$
$\#probTrans1 = 72\%$	$\#numPlas2 = 7$
$\#probTrans2 = 45\%$	$\#numPlas3 = 8$

Tabela 3: Parâmetros do recozimento simulado

$\#T_0 = 20$	$L_0 = 14$
$\#F_T = 1,78$	$F_L = 1,43$

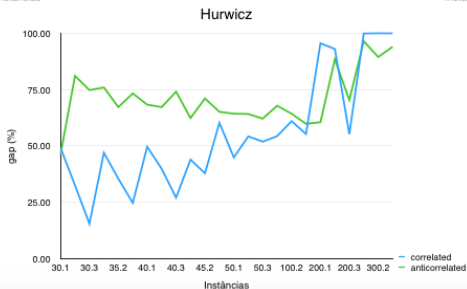
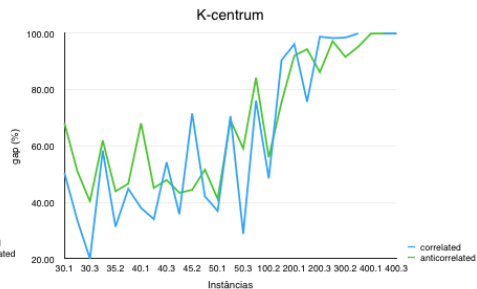
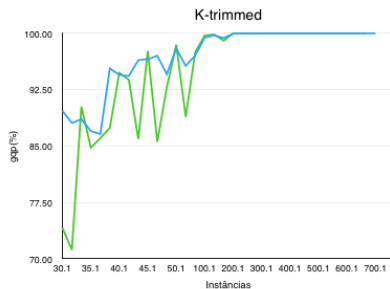
Metodologia dos experimentos

- Abordagens meta-heurísticas:
 - ▶ 30 execuções de cada instância;
 - ▶ qualidade final da solução retornada;
 - ▶ tempo para atingir a melhor solução;
 - ▶ tempo total;
 - ▶ contribuição de cada operador individualmente
- Abordagem exata:
 - ▶ Modelo de Fernández et al. (2017) melhorado;
 - ▶ Limite de tempo: 1h
- Comparar cada meta-heurística com o exato, calculando o desvio percentual da solução **mediana**:

$$d_p = \frac{OWA(s_h) - OWA(s_s)}{OWA(s_s)} * 100 \quad (1)$$

- *M-SA* vs *T-SA*:
 - ▶ Tempo total em segundos;
 - ▶ Tempo para atingir a melhor solução;
 - ▶ Qualidade de solução: teste *Mann-Whitney one-tailed*;

Comparação com o exato - *gap* do *solver*



Comparação com o exato

Abaixo, temos a porcentagem, por categoria, das instâncias em que o *T-SA* e *M-SA* conseguiram uma solução (**mediana**) melhor (desvio percentual negativo) que a do *solver* em 1h:

- *k-trimmed*

- ▶ *correlated T-SA* : 97,14% vs *M-SA* : 94,28%;
- ▶ *anticorrelated T-SA* : 93,94% vs *M-SA* : 84,85%;

- *k-centrum*

- ▶ *correlated T-SA* : 91,43% vs *M-SA* : 77,14%;
- ▶ *anticorrelated T-SA* : 87,88% vs *M-SA* : 84,85%;

- *Hurwicz*

- ▶ *correlated T-SA* : 80,00% vs *M-SA* : 57,14%;
- ▶ *anticorrelated T-SA* : 90,91% vs *M-SA* : 90,91%;

Comparação com o exato - Observações

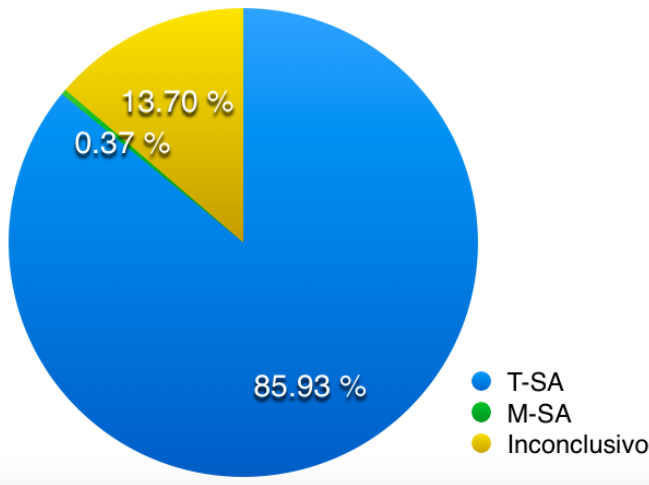
- As meta-heurísticas apresentam desvio percentual positivo para instâncias de até 100 vértices;
- Quanto maior a instância, mais negativo o desvio percentual;
- O desvio percentual máximo (positivo) respectivamente para *k-trimmed*, *k-centrum*, *Hurwicz*:
 - ▶ *M-SA*: 0,83%, 5,04% e 2,76%;
 - ▶ *T-SA*: 0,47%, 4,40% e 1,50%;

Qualidade de solução: *Mann-Whitney one-tailed*

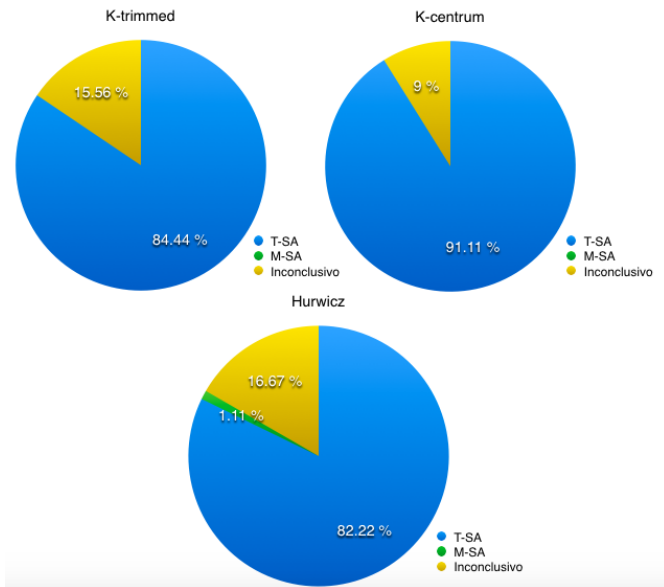
- Metodologia descrita por Knowles, Thiele e Zitzler (2006);
- Hipótese nula: "As amostras do *T-SA* e *M-SA* advêm da mesma distribuição";
- Hipótese alternativa: "A amostra do *T-SA* advém de uma distribuição melhor que *M-SA* (minimização)";
- Nível de significância: 0,05;
- *p-valores* menores que 0,05 indicam que a hipótese nula deve ser rejeitada em favor da hipótese alternativa;
- *p-valores* acima de 0.95 indicam que a hipótese nula deve ser aceita.

Qualidade de solução: *Mann-Whitney one-tailed*

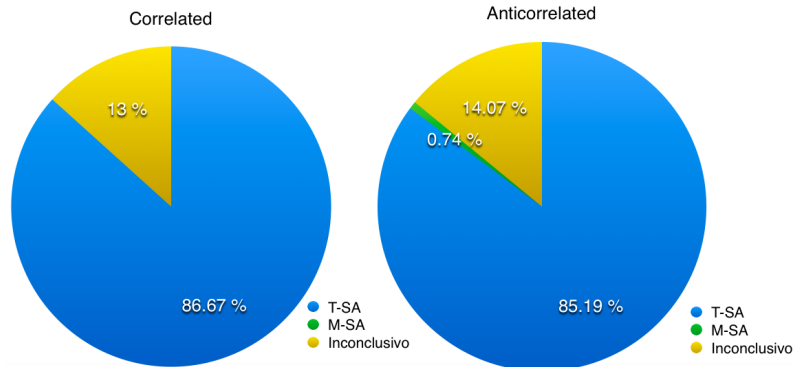
Comparação considerando todo o conjunto de instâncias



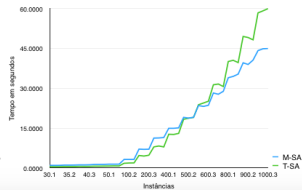
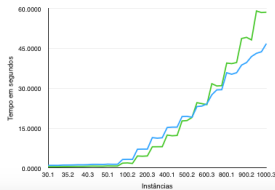
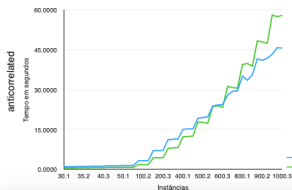
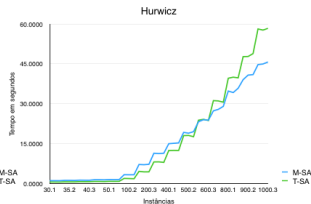
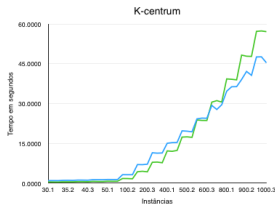
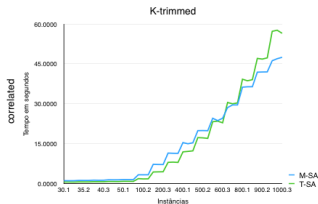
Qualidade de solução: *Mann-Whitney one-tailed*



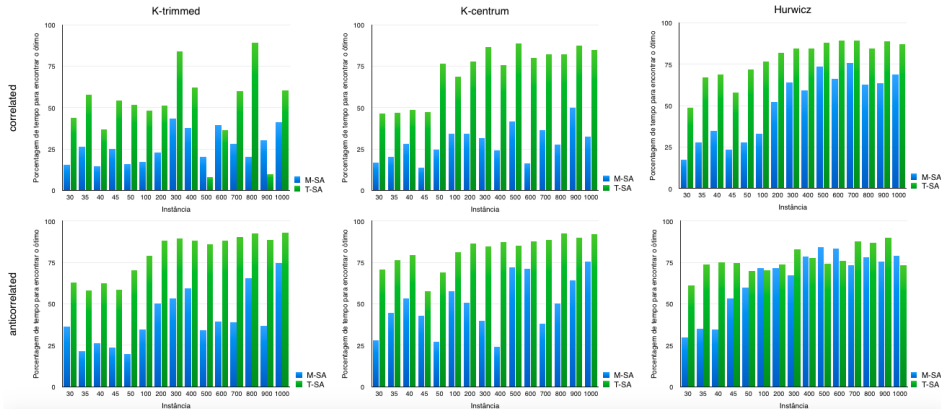
Qualidade de solução: *Mann-Whitney one-tailed*



M-SA vs T-SA: Tempo total



M-SA vs T-SA: Tempo para atingir o ótimo



Outras estatísticas

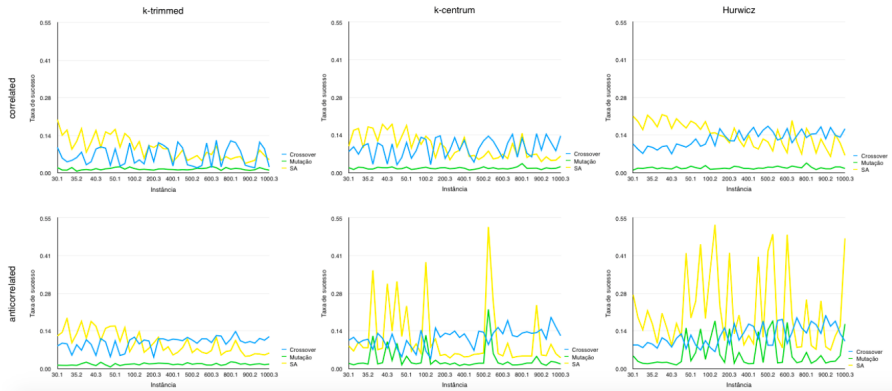


Figura 5: Taxas de sucesso dos operadores meméticos em função do tamanho das instâncias

Outras estatísticas

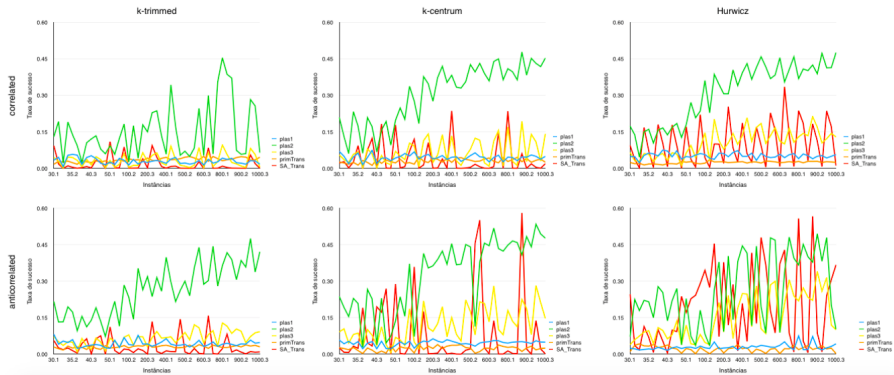


Figura 6: Taxas de sucesso dos operadores transgenéticos em função do tamanho das instâncias

Outras estatísticas

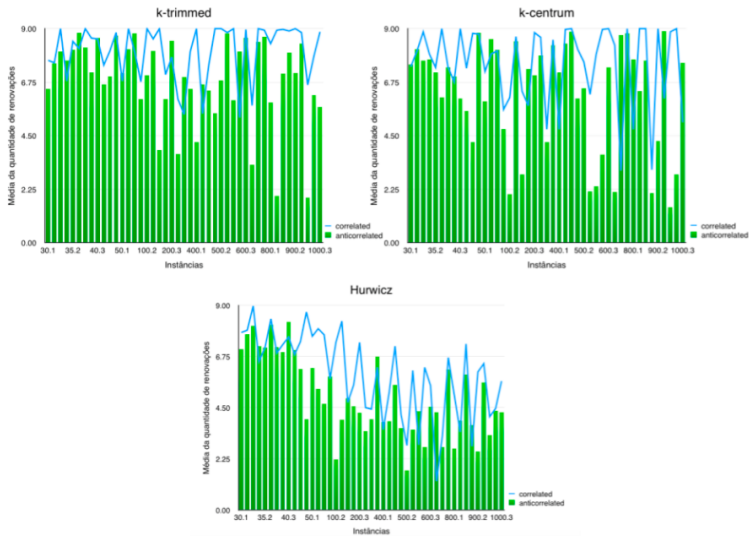




Figura 7: Taxas média de renovação da população, efetuada pelo *M-SA*, em função do tamanho das instâncias


Considerações Finais


- O transgenético consideravelmente foi melhor que o memético em:
 - ▶ Qualidade de solução;
 - ▶ Aproveitamento do tempo para melhorar o ótimo corrente;
- O transgenético também ganhou em tempo de processamento na maioria das instâncias;
- Trabalho futuro: testar para 10 (FEITO!!!!) e 5 objetivos (ainda não).

Referências Bibliográficas I



 FERNÁNDEZ, E. et al. Ordered weighted average optimization in multiobjective spanning tree problem. *European Journal of Operational Research*, Elsevier, v. 260, n. 3, p. 886–903, 2017.

 JIANG, S.; YANG, S. A strength pareto evolutionary algorithm based on reference direction for multiobjective and many-objective optimization. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 21, n. 3, p. 329–346, 2017.

 KNOWLES, J. *Local-search and hybrid evolutionary algorithms for Pareto optimization*. Tese (Doutorado) — Department of Computer Science, University of Reading, Reading, UK, 2002.

 KNOWLES, J.; THIELE, L.; ZITZLER, E. *A tutorial on the performance assessment of stochastic multiobjective optimizers*. [S.l.], 2006.

Referências Bibliográficas II

-  MONTEIRO, S. M. D.; GOLDBARG, E. F. G.; GOLDBARG, M. C. A new transgenetic approach for the biobjective spanning tree problem. In: *IEEE CEC 2010 Congress on Evolutionary Computation, 2010, Barcelona. Proceedings of IEEE CEC 2010 Congress on Evolutionary Computation*. [S.l.]: Piscataway, IEEE, 2010. v. 1, p. 519–526.
-  RAIDL, G. R.; JULSTROM, B. A. Edge sets: an effective evolutionary coding of spanning trees. *IEEE Transactions on Evolutionary Computation*, v. 7, n. 3, p. 225–239, 2003.