

Clustering Search - *CS



UFES

Marcos A. Spalenza

Doutorando em Ciência da Computação
Laboratório de Computação de Alto Desempenho - LCAD
Programa de Pós-Graduação em Informática - PPGI

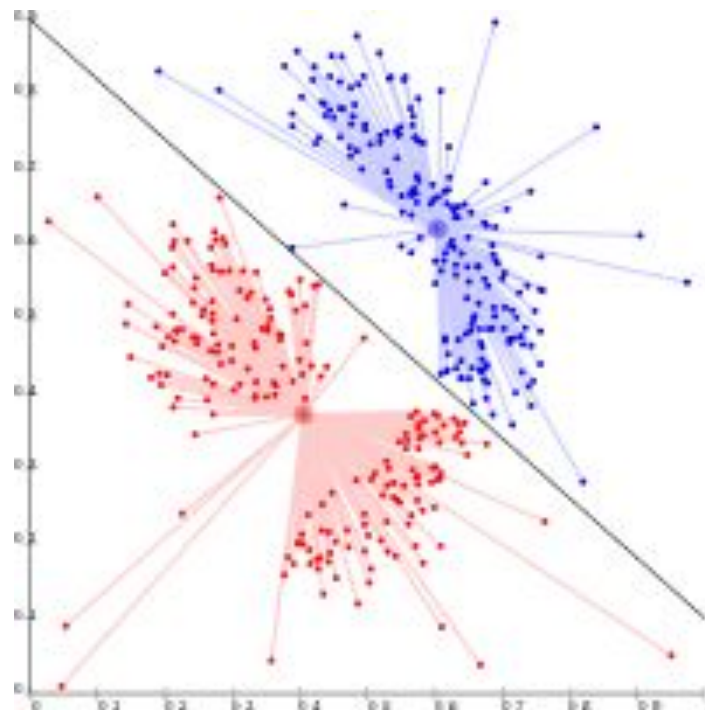
Apresentação

- Introdução
- Histórico
- Definição e Descrição
- Estado da Arte
- Exemplo de Aplicação

Introdução

Clustering Search - *CS é uma meta-heurística que visa selecionar os principais itens de uma população através da clusterização (Oliveira, 2007).

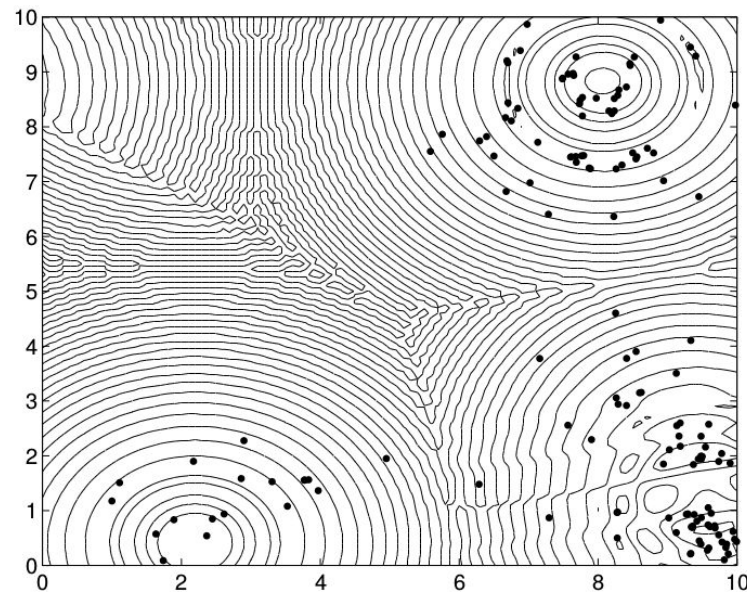
Essa técnica segue os conceitos de Estatística empregados em Machine Learning sobre a representatividade do centróide para as amostras agrupadas. Assim, aplica-se *Local Search - LS nos centróides* do particionamento.



Histórico

A metaheurística têm como início a aplicação de clusterização para o refinamento populacional entre as iterações do Algoritmo Genético, o *Evolutionary Clustering Search - ECS* (Oliveira et.al., 2004).

Posteriormente uma generalização do algoritmo definiu que o mesmo seria plausível com os conjuntos de respostas para outras meta-heurísticas (Oliveira et.al., 2007).



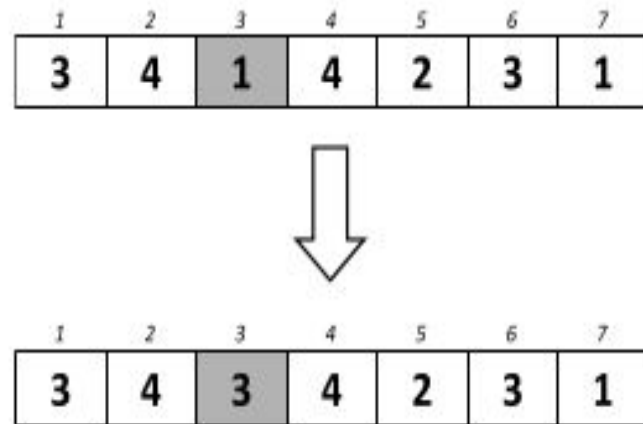
OLIVEIRA, Alexandre CM; LORENA, Luiz AN. Detecting promising areas by evolutionary clustering search. In: **Brazilian Symposium on Artificial Intelligence**. Springer, Berlin, Heidelberg, 2004. p. 385-394.

OLIVEIRA, Alexandre CM; LORENA, Luiz AN. Hybrid evolutionary algorithms and clustering search. In: **HYBRID evolutionary algorithms**. Springer, Berlin, Heidelberg, 2007. p. 77-99.

Definição

Clustering Search é uma meta-heurística de minimização por *Local Search* nas principais de soluções da vizinhança, resultantes da análise da densidade em clusters da população.

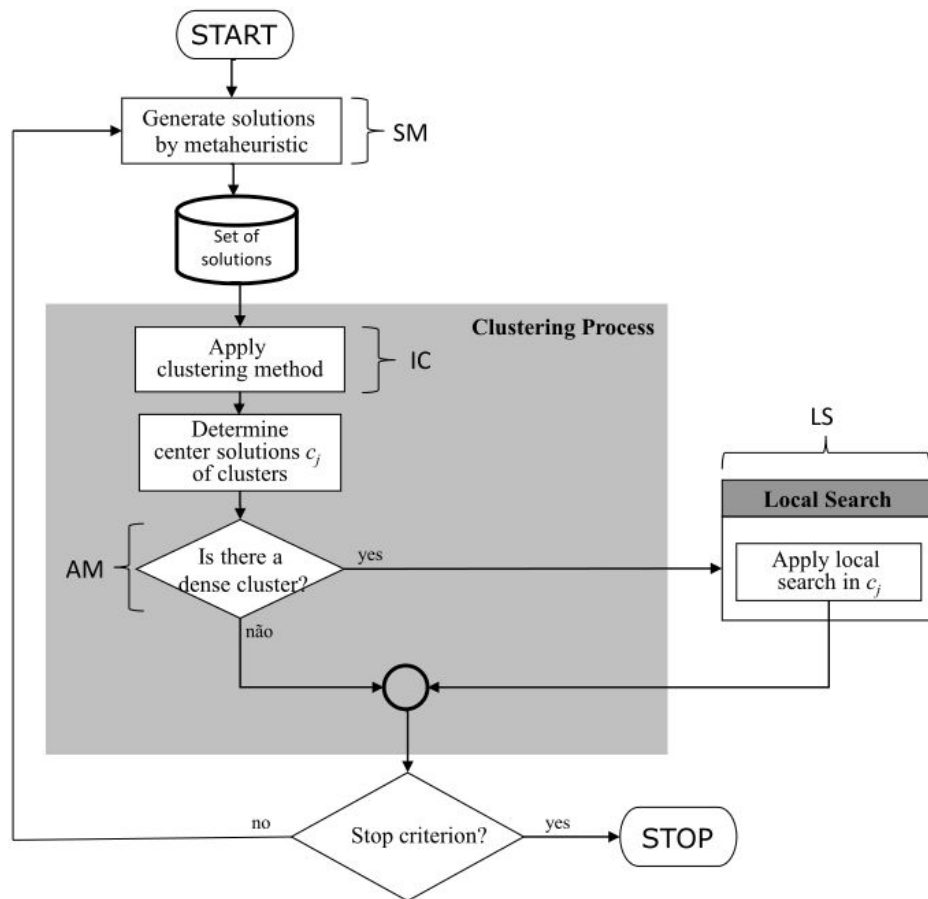
A minimização desse método atenta-se na tentativa agressiva de melhoria dos principais itens de vizinhança através da LS.



Descrição

É composta por 4 partes:

- Meta-heurística de Busca - SM
- Componente de Clusterização Iterativa - IC
- Módulo de Análise - AM
- Busca Local - LS



algorithm CS

create the initial clusters of CS

{ metaheuristic }

while termination condition not satisfied **do**

 generate a solution (s_k) by metaheuristic

{ clustering process }

 find the most similar cluster c_j to the solution s_k

 insert s_k into c_j ($\delta_j \leftarrow \delta_j + 1$)

$c_j \leftarrow \text{Assimilation}(c_j, s_k)$

if $\delta_j \geq \lambda$ **then**

$\delta_j \leftarrow 0$

{ local search }

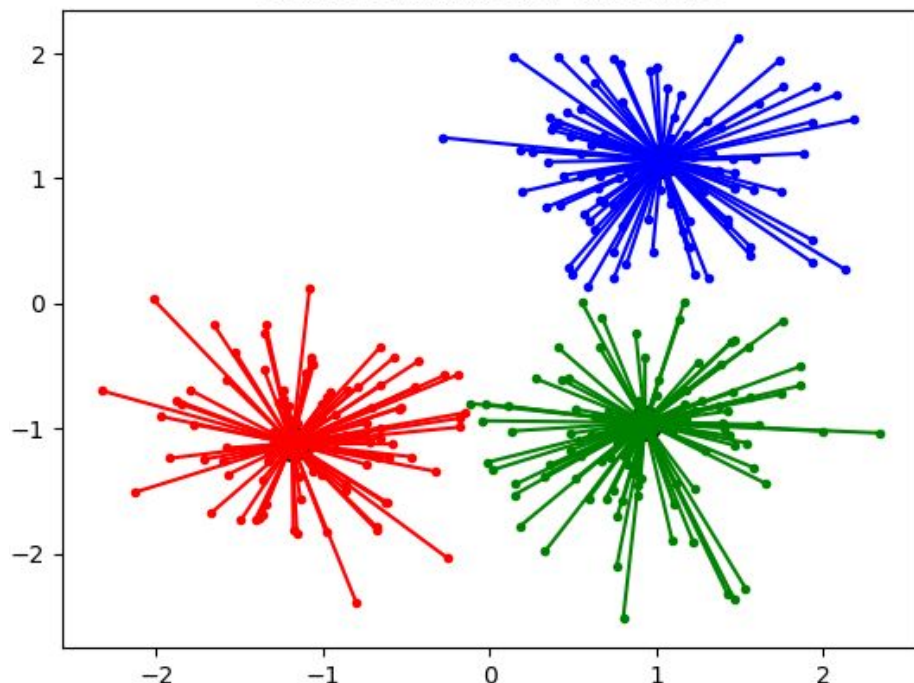
 find the best neighborhood of c_j

end if

 update the center c_j

end while

end algorithm



Estado da Arte

Destacam-se nesse algoritmo estudos da própria técnica de clusterização empregada (Araújo et.al., 2019). De origem em estruturas híbridas, o *Clustering Search* já foi estudado em conjunto com diversas outras meta-heurísticas:

- Algoritmos Genéticos
- Iterated Local Search
- Variable Neighborhood Search
- Simulated Annealing
- GRASP
- Busca Tabu

ARAÚJO, Eliseu J.; CHAVES, Antônio A.; LORENA, Luiz AN. Improving the Clustering Search heuristic: An application to cartographic labeling.

Applied Soft Computing, v. 77, p. 261-273, 2019.

OLIVEIRA, Alexandre César Muniz de; CHAVES, Antonio Augusto; LORENA, Luiz Antonio Nogueira. Clustering search. **Pesquisa operacional**, v. 33, n. 1, p. 105-121, 2013.

Estado da Arte

- **Otimização Contínua:** *Unconstrained Numerical Optimization.*
- **Problemas de Agendamento:** *Pattern Sequencing Problem; Permutation Flow Shop Scheduling Problem; M-Machine No-Wait Flow Shop Scheduling Problem; Assembly Line Worker Assignment and Balancing Problem.*
- **Problemas de Localização:** *Capacitated p -Median Problem; Probabilistic Maximal Covering; Location-Allocation Problem; Capacitated Centered Clustering Problem; Hub Location Problem; Capacitated Hub Location Problem.*
- **Problemas de Roteamento:** *Traveling Tournament Problem; Prize Collecting Traveling; Salesman Problem; Berth Allocation Problem; Dial-a-Ride Problem; Workover Rig Routing Problem.*

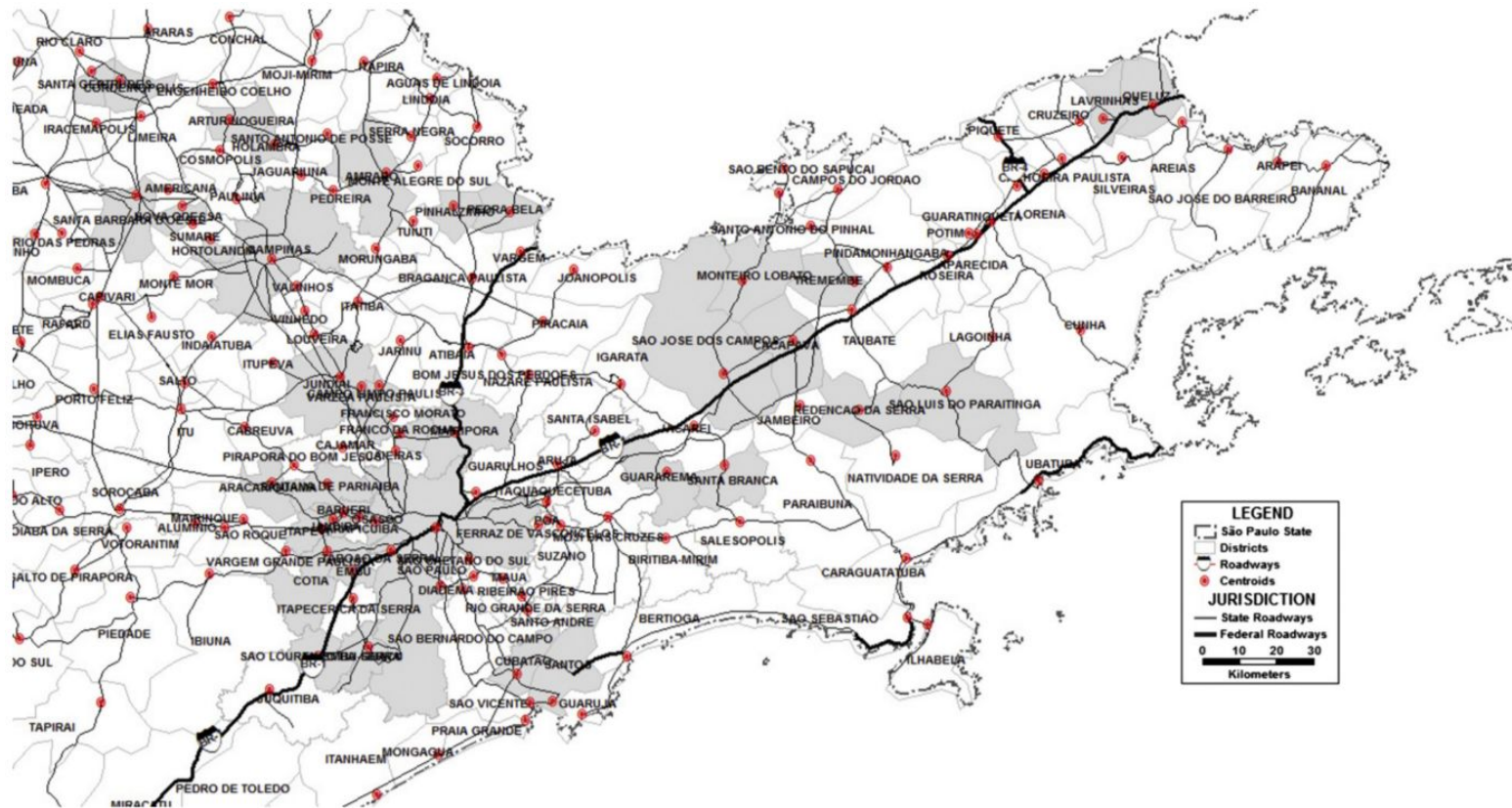
Exemplo de Aplicação

ARAÚJO, Eliseu J.; CHAVES, Antônio A.; LORENA, Luiz AN. Improving the Clustering Search heuristic: An application to cartographic labeling. **Applied Soft Computing**, v. 77, p. 261-273, 2019.

O problema de posicionamento de rótulos, é responsável por alocar descrições textuais em locais da imagem evitando sobreposição. Isso pode ser visto frequentemente em gráficos e mapas, onde o ideal é priorizar a legibilidade da imagem. Restringe-se, porém, ao posicionamento próximo do ponto de interesse.



Exemplo de Aplicação



Exemplo de Aplicação

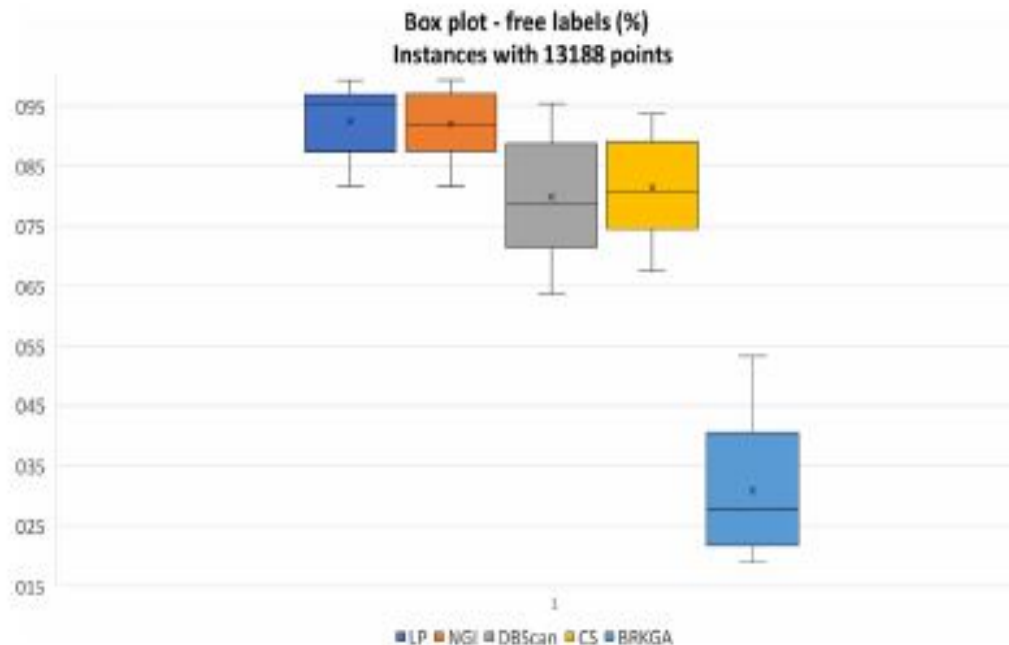
Label Propagation (LP)

Natural Group Identification (NGI)

DBSCAN (DCS)

Clustering Search (CS)

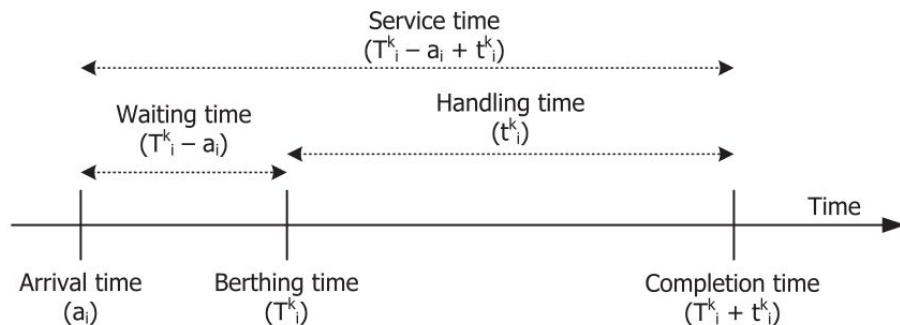
BRKGA



Exemplo de Aplicação

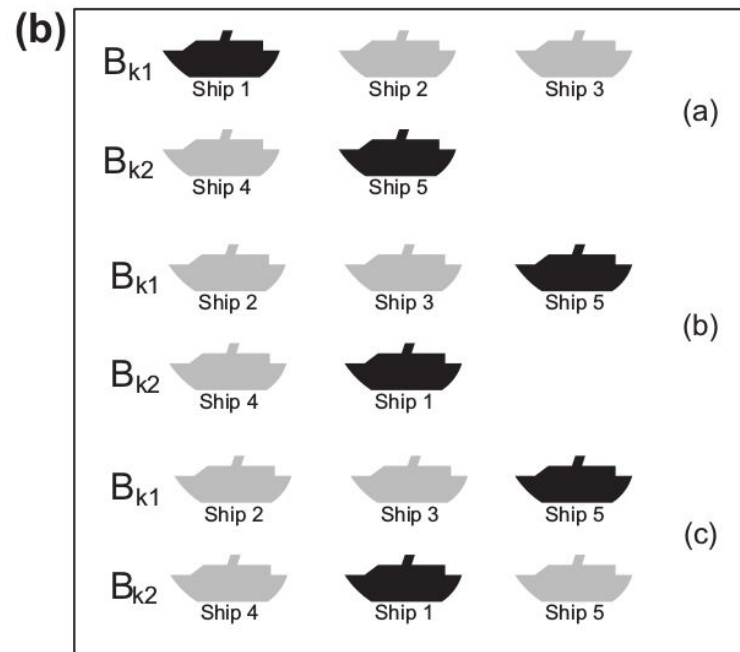
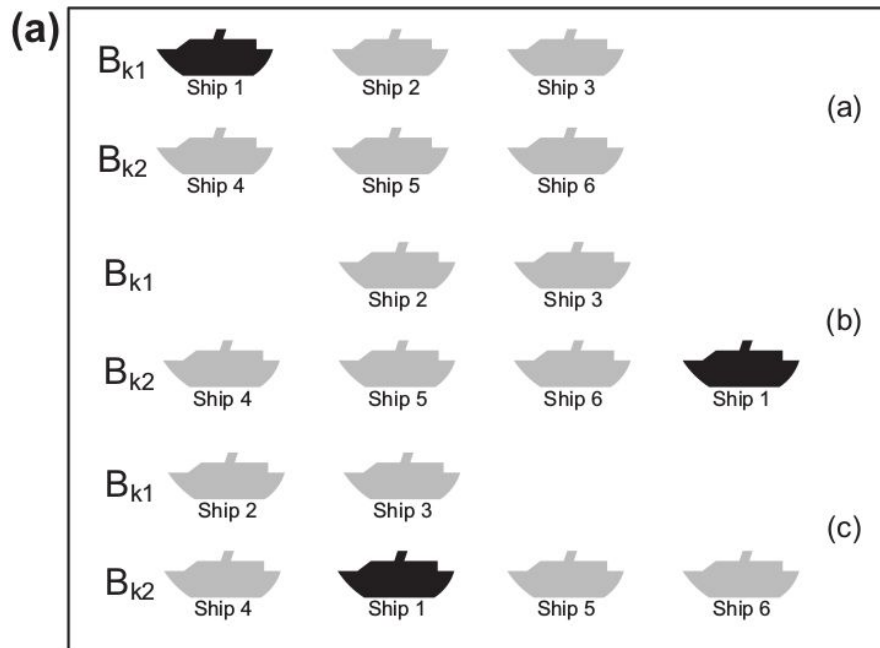
OLIVEIRA, Rudinei Martins; MAURI, Geraldo Regis; LORENA, Luiz Antonio Nogueira. Clustering search for the berth allocation problem. **Expert Systems with Applications**, v. 39, n. 5, p. 5499-5505, 2012.

Berth Allocation Problem (BAP) é considerado o maior problema de terminais marítimos. A proposta envolve minimizar o tempo de carga e descarga de *contêineres* dos navios. Para isso foi utilizado um híbrido de *Simulated Annealing* com o *Clustering Search*.



Exemplo de Aplicação

Foram utilizadas 30 instâncias diferentes, cada uma com 60 navios e 13 berths. Os dados foram coletados no porto de Gioia Tauro (Itália).



Exemplo de Aplicação

Os resultados foram comparados com cinco testes:

- *CPLEX*
- *Simulated Annealing with Re-Annealing (SA+RA)*
- *Tabu Search (TS)*
- *Population Training Algorithm with Linear Programming (PTA-LP)*
- *Generalized Set-Partitioning Problem (GSPP)*

As técnicas de otimização apresentam apenas pequenas diferenças em relação ao *CS quando destacadas ganho do CPLEX.

Apesar disso, o tempo 12.79 do *CS foi extremamente reduzido em relação ao CPLEX (3600.00), TS (120.00), SA+RA (60.26), PTA/LP (93.99) e GSPP (14.98).

IMPROVEMENT (%)				
CPLEX	TS	SA + RA	PTA/LP	GSPP
-	0.42	0.00	0.00	0.00
51.61	0.16	0.00	0.00	0.00
55.98	0.88	0.00	0.00	0.00
70.09	0.08	0.00	0.00	0.00
54.83	0.08	0.00	0.00	0.00
-	0.08	0.00	0.00	0.00
55.70	0.00	0.00	0.00	0.00
74.91	0.00	0.00	0.00	0.00
-	0.00	0.00	0.00	0.00
-	0.00	0.00	0.00	0.00
-	0.73	0.00	0.07	0.00
58.67	0.00	0.00	0.00	0.00
-	0.00	0.00	0.00	0.00
-	0.00	0.00	0.00	0.00
72.28	0.00	0.00	0.00	0.00
68.43	0.80	0.00	0.07	0.00
-	0.00	0.00	0.00	0.00
63.46	0.07	0.00	0.00	0.00
43.04	0.22	0.22	0.00	0.00
-	0.00	0.00	0.00	0.00
-	0.37	0.00	0.00	0.00
61.99	0.45	0.00	0.00	0.00
-	0.00	0.00	0.00	0.00
74.11	0.08	0.00	0.00	0.00
30.96	0.22	0.07	0.00	0.00
47.70	0.90	0.00	0.00	0.00
60.70	0.00	0.00	0.00	0.00
-	0.44	0.07	0.07	0.00
73.38	0.16	0.00	0.00	0.00
-	0.52	0.00	0.00	0.00

OTIMIZAÇÃO COMBINATÓRIA E META-HEURÍSTICAS



UFES

Marcos A. Spalenza

Doutorando em Ciência da Computação
Laboratório de Computação de Alto Desempenho - LCAD
Programa de Pós-Graduação em Informática - PPGI