C-BIO
The Bio-inspired Computation Group
UFPR

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

A Meta-learning Algorithm Selection Approach for the Quadratic Assignment Problem

Augusto Lopez Dantas (aldantas@inf.ufpr.br)
Aurora Trinidad Ramirez Pozo (aurora@inf.ufpr.br)

IEEE Congress on Evolutionary Computation, Rio de Janeiro - Brazil

July 11, 2018

Motivation

- Automatic optimization problem solving
  - Reduce the requirement of expert knowledge
  - Raise de level of generality
- Meta-learning based approach
- Quadratic Assignment Problem [1]

[1] http://anjos.mgi.polymtl.ca/qaplib/inst.html

Summary

1. Quadratic Assignment Problem
2. Meta-features and Meta-labels
3. Experiments and Results
4. Meta-features Analysis
5. Conclusion

Quadratic Assignment Problem (QAP)

Proposed as a mathematical model in 1957.[2]

Definition

Given two sets of n facilities and n locations, the goal in QAP is to assign each facility into one unique location in order to minimize the total flow and distance between the associations. The problem can be formally defined as

$$\min_{\phi \in S_n} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij} d_{\phi(i)\phi(j)}$$

[2] Tjalling C Koopmans and Martin Beckmann. “Assignment problems and the location of economic activities”. In: *Econometrica: Journal of the Econometric Society* (1957).

QAP - Example

Permutation Encoding

3	4	5	1	2

F	1	2	3	4	5
1	0	4	2	1	3
2	4	0	1	3	0
3	2	1	0	4	2
4	1	3	4	0	1
5	3	0	2	1	0

Flow Matrix

D	1	2	3	4	5
1	0	50	60	94	50
2	50	0	22	50	36
3	60	22	0	44	14
4	94	50	44	0	50
5	50	36	14	50	0

Distance Matrix

QAP - Example

Permutation Encoding

3	4	5	1	2

F	1	2	3	4	5
1	0	4	2	1	3
2	4	0	1	3	0
3	2	1	0	4	2
4	1	3	4	0	1
5	3	0	2	1	0

Flow Matrix

D	1	2	3	4	5
1	0	50	60	94	50
2	50	0	22	50	36
3	60	22	0	44	14
4	94	50	44	0	50
5	50	36	14	50	0

Distance Matrix

$$f_{3,4}d_{\phi(3)\phi(4)} = f_{3,4}d_{1,2} = 4 * 50$$

Meta-features - Static

Abbreviation	Name	Description
n	Instance Size	Amount of associations
fd	Dominance	Variation coefficient ($100 * \frac{\mu}{\sigma}$)
dd		
fsp	Sparsity	Amount of zero values related to the total (n^2)
dsp		
fas	Asymmetry	Amount of values (i, j) different than (j, i) in relation to the total pairs $\left(\frac{n^2 - n}{2}\right)$
das		

Meta-features - Fitness Landscape Analysis (FLA)

Iterated Local Search (ILS)[3]

- 1000 times
- 500 iterations

Best Improvement Local Search (BI)

- Random restarts
- 5000 unique local optima
- 5 minutes timeout

[3] Thomas Stützle. "Iterated local search for the quadratic assignment problem". In: *European Journal of Operational Research* 174.3 (2006).

Meta-features - FLA

Abbreviation	Name	Description
n_opt	Amount of pseudo optima	Best unique solutions found by ILS
ils_dst	Average distance	Hamming distance to the closest pseudo optimum
bi_dst		
ils_fdc	Fitness Distance Correlation[4]	$FDC = \frac{\frac{1}{m}\sum_{i=1}^{m}(c_i - \mu_c)(d_i - \mu_d)}{\sigma_c . \sigma_d}$
bi_fdc		
acc	Autocorrelation coefficient[5]	Average fitness variation caused by a single step along the neighborhood
acl	Correlation length [5]	Number of required steps to make the correlation values statistically different

[4] Terry Jones and Stephanie Forrest. “Fitness Distance Correlation As a Measure of Problem Difficulty for Genetic Algorithms”. In: *Proceedings of the 6th International Conference on Genetic Algorithms*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995.

[5] Francisco Chicano, Gabriel Luque, and Enrique Alba. “Autocorrelation measures for the quadratic assignment problem”. In: *Applied Mathematics Letters* 25.4 (Apr. 2012).

Meta-labels

Algorithms

- Breakout Local Search (BLS)[6]

  $100 * n$ iterations
- Max-Min Ant System with BI Local Search (MMASBI)[7]

  $100 * n$ iterations
- Robust Tabu Search (RO-TS)[8]

  $2000 * n$ iterations

- 30 executions
- Average cost
- Average execution time

---

[6] Una Benlic and Jin-Kao Hao. “Breakout local search for the quadratic assignment problem”. In: *Applied Mathematics and Computation* (Jan. 2013).

[7] Thomas Stützle and Holger H Hoos. “MAX–MIN Ant System”. In: *Future Generation Computer Systems* 16.8 (2000).

[8] Éric Taillard. “Robust taboo search for the quadratic assignment problem”. In: *Parallel computing* 17.4-5 (1991).

Experiments

Full Dataset

Class	Instances
BLS	33
MMASBI	89
RO-TS	13

Classifier

Random Forest [9]

100 estimators and tree depth limit of 10

Validation

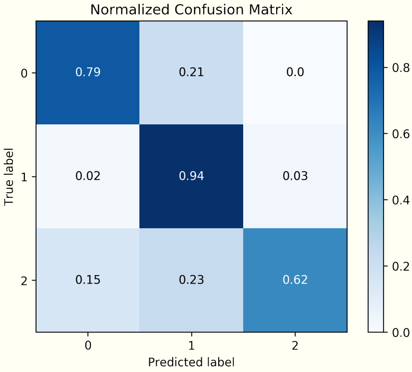
Stratified 10-Fold

[9] http://scikit-learn.org

Initial Results

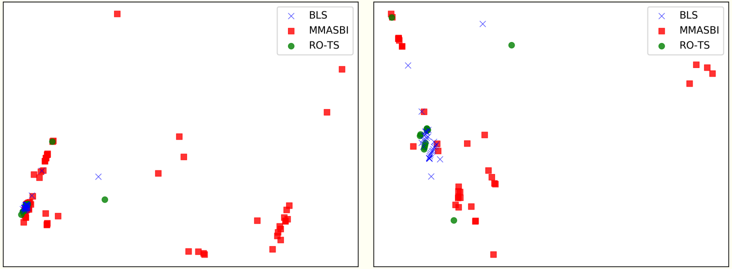
Performance

Accuracy	F-score			
	BLS	MMASBI	RO-TS	Average
0.8741	0.8254	0.9180	0.6667	0.8034



Data Cleansing

Principal Component Analysis



Full | Reduced

Reduced dataset

Class	Instances
BLS	33
MMASBI	35
RO-TS	13

Feature Selection

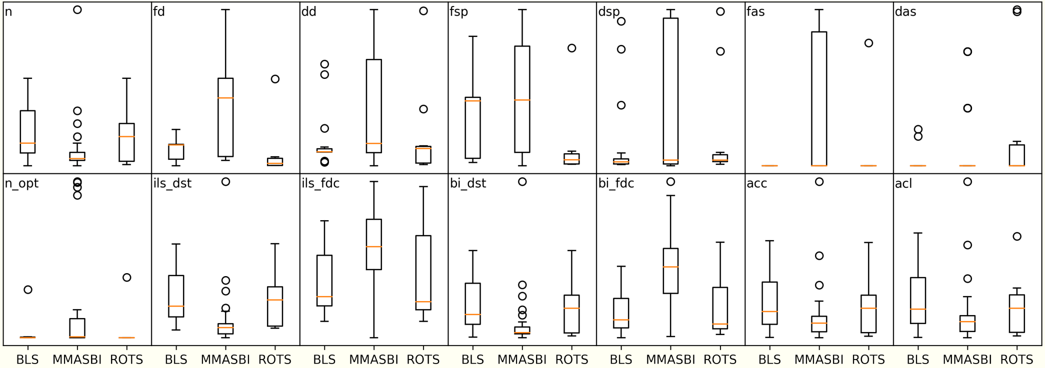
Best features for each dataset

Dataset	Best Features
Full	n, fd, dd, fsp, fas, ils_fdc, bi_fdc, acc
Reduced	fd, dd, dsp, fas, ils_dst, ils_fdc, bi_dst, bi_fdc

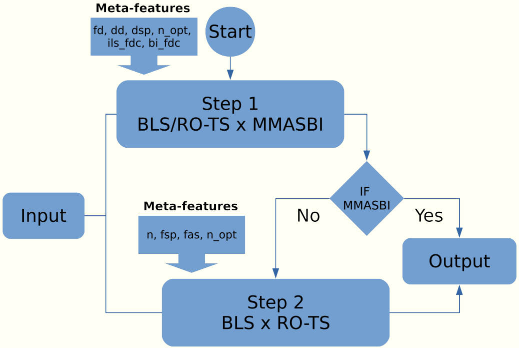
New results

Dataset	Accuracy	F-score			
		BLS	MMASBI	RO-TS	Average
Full	0.9037	0.8667	0.9355	0.75	0.8507
Reduced	0.8765	0.8923	0.8947	0.7619	0.8496

Meta-features Distributions



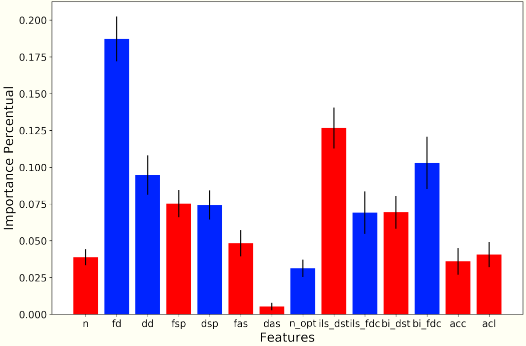
Cascade Classification Scheme



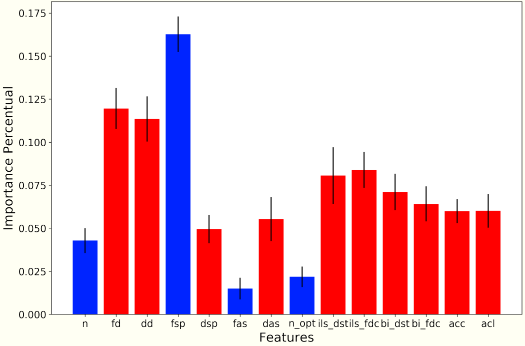
Performance

Accuracy	F-score			
	BLS	MMASBI	RO-TS	Average
0.8765	0.875	0.8912	0.8333	0.8665

Meta-features Analysis

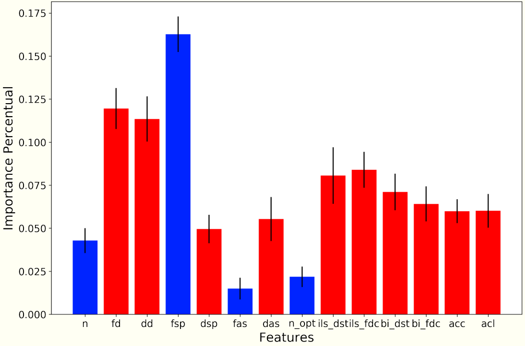


Step 1

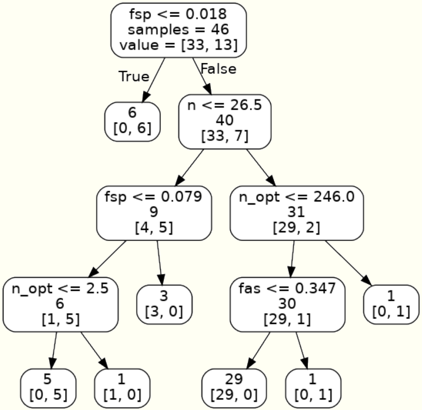


Step 2

Meta-features Analysis



Importances



Decision Tree Visualization

Conclusion

- Balanced class performances with data pre-processing and specialized trained models
- Difficulty on identifying the proper subset of meta-features
- Deprecation of the QAPLIB
- Future Works:
  - Include extra QAP instances
  - Extract more meta-features
  - Experiment with different sets of meta-heuristics
  - Multi-label classification
- Materials available online [10]

[10] github.com/aldantas/AS_QAP_CEC2018

C-BIO
The Bio-inspired Computation Group
UFPR

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

A Meta-learning Algorithm Selection Approach for the Quadratic Assignment Problem

Thank you!

Augusto Lopez Dantas (aldantas@inf.ufpr.br)

Aurora Trinidad Ramirez Pozo (aurora@inf.ufpr.br)

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico