Influence of the individual's size on the island model architecture

Autor: Mgr. Štěpán Balcar
Supervisor: Prof. RNDr. Peter Vojtáš, DrSc.

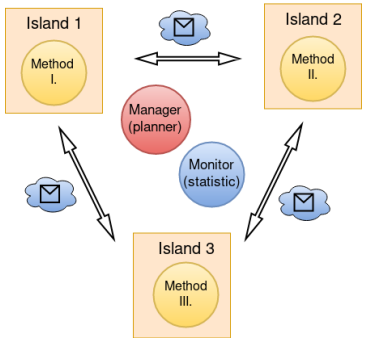
Charles University,
Faculty of Mathematics and Physics,
Prague, Czech Republic

Abstract:

Our motivation starts from recommendation systems on large data. We use hybrid evolutionary agent based system. Our system runs in parallel several optimization methods. Main challenge is the frequency of individuals migration. We describe technical restrictions which complicate parallelization techniques of helping each other computational methods. Size of data influences latency in communication and measurement. We experiment with architectural changes which enable to move the limits. At the same time we would like to give planners sufficient freedom to manage distributed computation.

Data a znalosti & WIKT 2018
Date: October 11-12, 2018,
Brno, Czech Republic

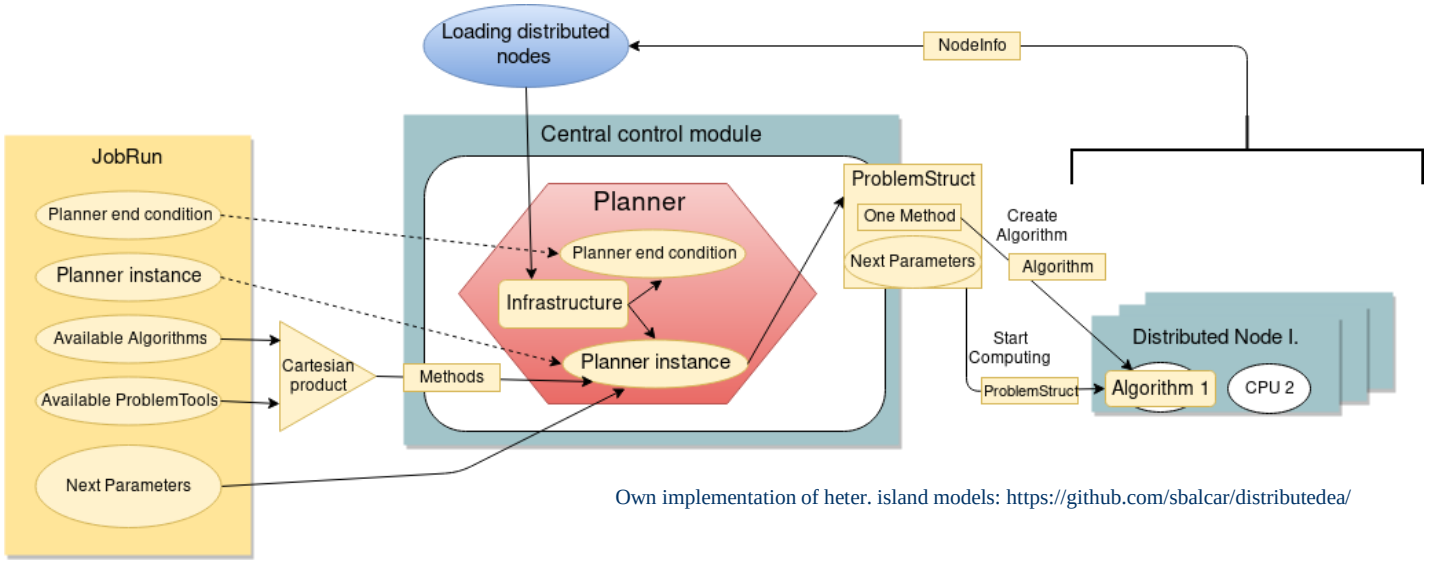
Multi-agent island model



- island models = migration of individuals

- problems = TSP, BinPacking problem, Vertex cover problem, Continuous optimization problem, hyperparameter tuning

- dataset = 1-20KB

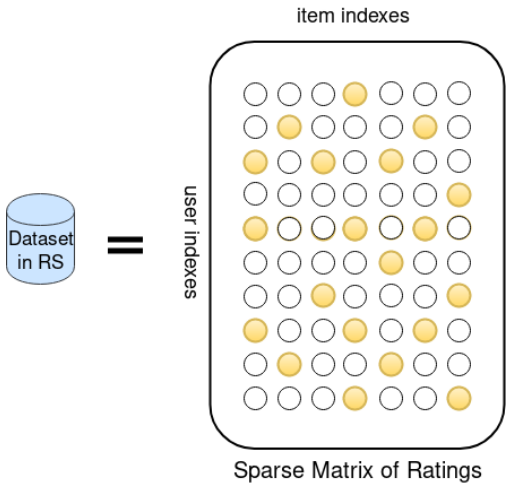


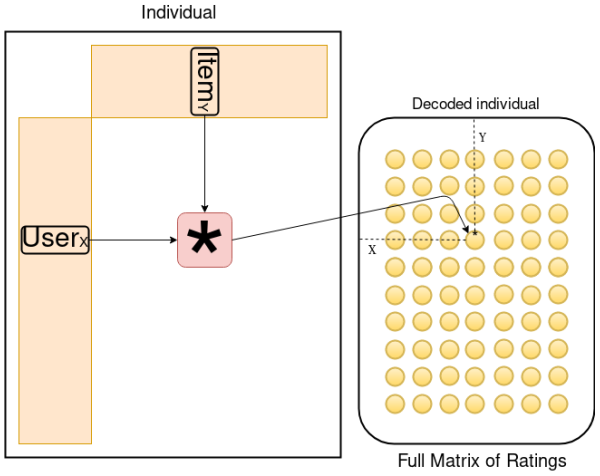
Own implementation of heter. island models: https://github.com/sbalcar/distributedea/

Balcar, S and Pilat, M, Heterogeneous Island Model with Re-planning of Methods, GECCO'18, short paper, 2018, Kyoto Japan
Balcar, S and Pilat, M, Online Parallel Portfolio Selection with Heterogeneous Island Model, ICTAI'2018, accepted as full paper, 2018, Volos, Greece

Supervisor's idea: use it for recommendation

Matrix factorization recomendation:

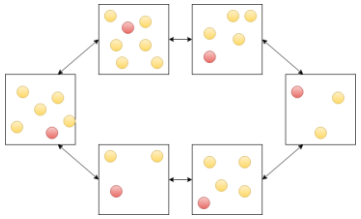




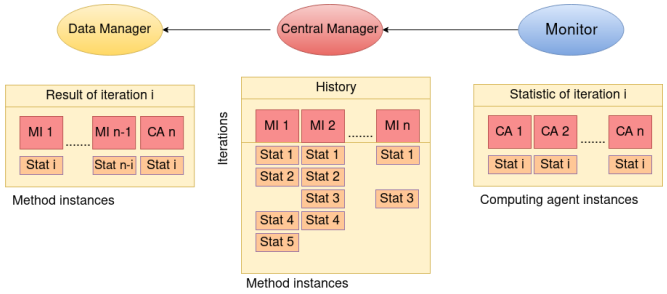
- migration = one type of individual
- individual = latent factors
- dataset = 20kB → 20MB

=> architectural problem

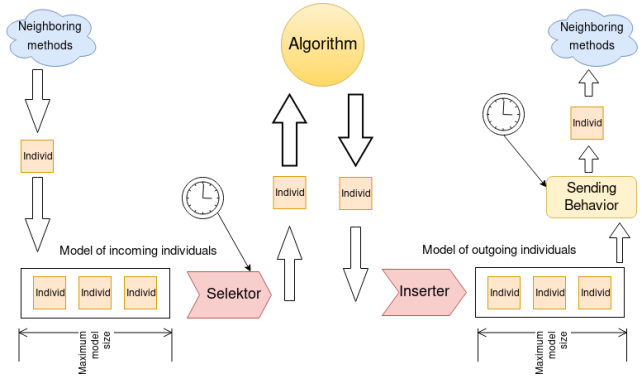
Architectural changes



- changing parameters:
  frequency of migration



- history based on hash objects and not on the whole models



- reduced queues of incoming and outgoing individuals

HPC - computational artificial intelligence

Jade framework - emphasis on adaptability http://jade.tilab.com/

Datasets MovieLens: https://movielens.org/

ML-100k 100000 ratings by 943 users on 1682 movies
ML-1M 1000209 ratings, by 6040 users on 3900 movies

Design of experiments:

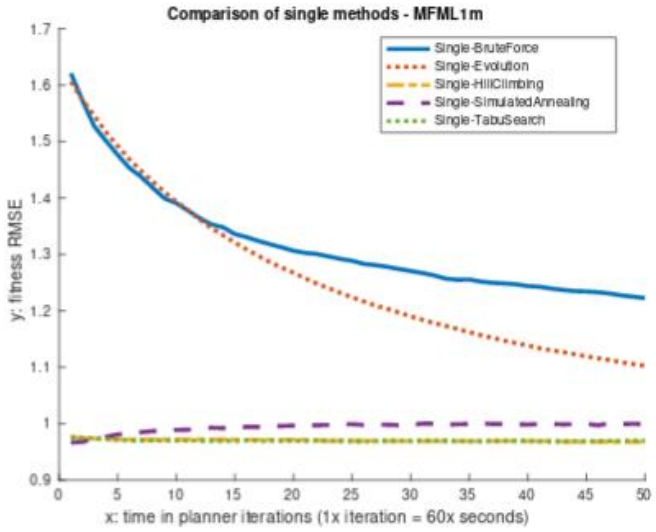
Parameter	value
Number of iterations	50, period = 60 seconds
Number of islands	4 (AMD Opteron 6376)
Neighbors of method	3 (distributed to everyone)
Migration frequency	5 seconds

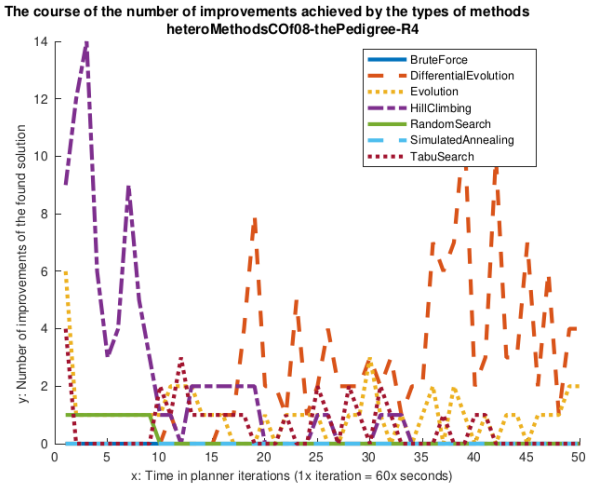
Table 2: Parameters of the island model

Experiments

Methods:

- HillClimbing
- RandomSearch
- Evolution
- BruteForce
- TabuSearch
- Sim.Annealing
- Diff.Evolution





Thank you

Questions?

Results - RMSE

Methods	Single-min	Islands-min	Single-average	Islands-average
BruteForce	0.98787115	0.94367838	0.99088728	0.94571058
DifferentialEvolution	1.50105714	1.49096267	1.51771324	1.49957429
Evolution	**0.88196787**	**0.87695296**	**0.88705747**	**0.87952888**
HillClimbing	0.98047412	0.97400051	0.98207866	0.97824895
RandomSearch	1.60448414	1.58333447	1.61703977	1.60802325
SimulatedAnnealing	1.06110779	1.03108626	1.07797515	1.03806128
TabuSearch	0.98048607	0.97681064	0.98217126	0.97963446

Table 4: Comparison of single methods and island models: Dataset - MFML100k, Runs - 9

Methods	Single-min	Islands-min	Single-average	Islands-average
BruteForce	1.22268511	1.18410586	1.23279212	1.2033589
DifferentialEvolution	1.55227103	1.55057106	1.55780106	1.55422412
Evolution	1.10254503	1.07069968	1.13526207	1.09140823
HillClimbing	0.96832015	**0.96727131**	0.97065502	**0.96840234**
RandomSearch	1.65483811	1.66265835	1.66738744	1.6660426
SimulatedAnnealing	**0.96655732**	0.97118289	**0.97048761**	0.9729405
TabuSearch	0.96873215	0.96735123	0.97106628	0.96854622

Table 5: Comparison of single methods and island models: Dataset - MFML1m, Runs - 9 (in red there are violations of island improvement)

- smaller models (individuals) → efficient parallelization
- larger models (individuals) → migration burdens too much the system