evo* 2012

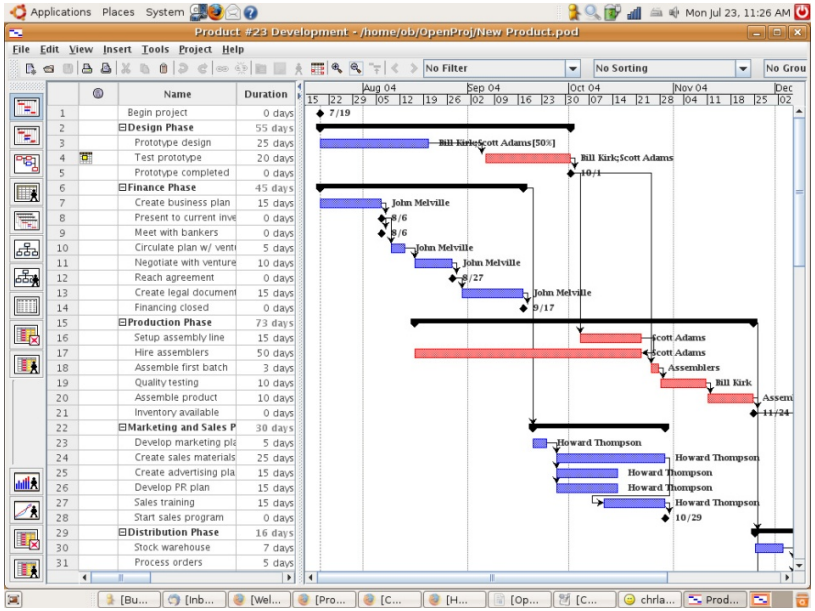
A Novel Multiobjective Formulation of the Robust Software Project Scheduling Problem

| M*: Multidisciplinar Multiobjective Metaheuristics
| TIN2008-06491-C04-01 Ministerio de Ciencia e Innovación (Spain)
| UMA - University of Málaga

Francisco Chicano, Alejandro Cervantes,
Francisco Luna, Gustavo Recio

Introduction

- **Software projects** usually involve **many people** and **many resources** that have to be **efficiently managed**
  - Poor planning → project failure
  - Goal: providing software engineers with an **automatic tool** for such a task
- **Optimization problem**: **assigning employees to tasks**
- **Particularities**
  - Several **people** assigned to tasks
  - Task cost estimation is usually **wrong**
- **Multiobjective** problem
  - **Objectives**: Cost and duration
  - **Constraints**: task completion
- **Robustness**
  - Task cost is inaccurate



Problem Formulation: Instance Information

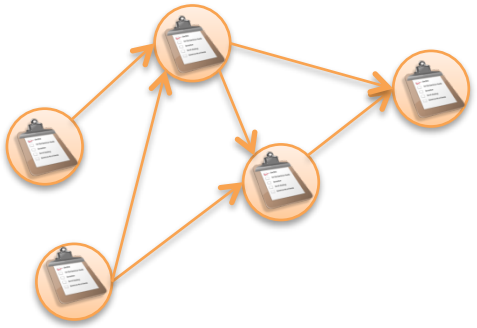
Employee

Salary

Task

Cost

TPG



Productivity matrix

	T1	T2	T3	T4	T5	T6
E1	0.3	0.2	0.5	0.7	1.0	0.0
E2	0.0	0.0	0.2	0.1	0.5	0.8
E3	0.2	0.0	0.0	0.6	1.0	1.0
E4	0.4	0.6	0.0	0.0	0.0	1.0

Problem Formulation: Solution

Priorities matrix

q	T1	T2	T3	T4	T5	T6
E1	3	1	5	0	0	0
E2	0	0	2	1	5	0
E3	2	0	0	0	1	1
E4	0	0	0	1	0	1

d
0.3
1.0
0.2
0.4

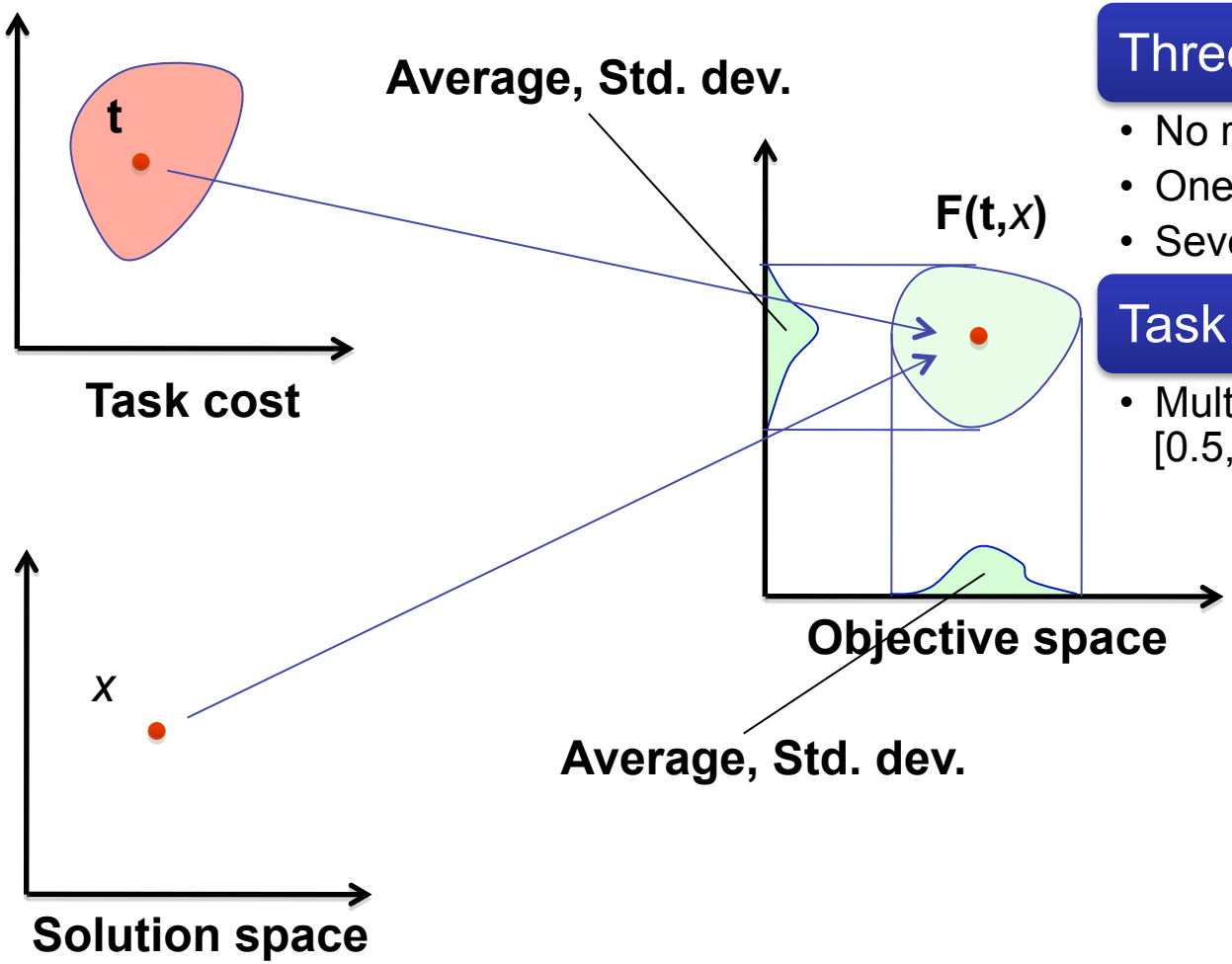
Dedication vector

Delays vector

r	3	2	5	7	1	0

- **The evaluation of a solution is based on a simulation of the project**
- **Objectives:**
  - **Makespan: the minimum time slot in which all tasks are done**
  - **Cost: salary multiplied by the dedication and worked hours**

Problem Formulation: Robustness



Three approaches

- No robustness (NR)
- One task changes (OTR)
- Several tasks change (STR)

Task change

- Multiply by a random value in [0.5,2]

Problem Formulation: Comparison with others

Shop Scheduling

- Tasks are the same, employees in SPS are like machines in SS
- Productivity is like length of tasks in machines
- In SS only one machine can perform a task, while in SPS a team works on a task
- In SS the concept of dedication of a machine does not exist

Resource-Constrained Project Scheduling

- Several kinds of resources in RCSP against one in SPS
- Each activity in RCSP requires different amounts of resources
- There is no minimum or maximum number of employees in SPS

Gutjahr and Chang models

- Complex models with lot of details and parameters
- Potential inaccuracy in the new parameters
- We try to keep the model simple

Algorithms in the Comparison

NSGA-II

- Generational GA
- Ranking & Crowding

SPEA2

- Generational GA + External Archive
- Strengh raw fitness & K-nearest neighbor

PAES

- (1+1) Evolution Strategy + External Archive
- Adaptive Grid

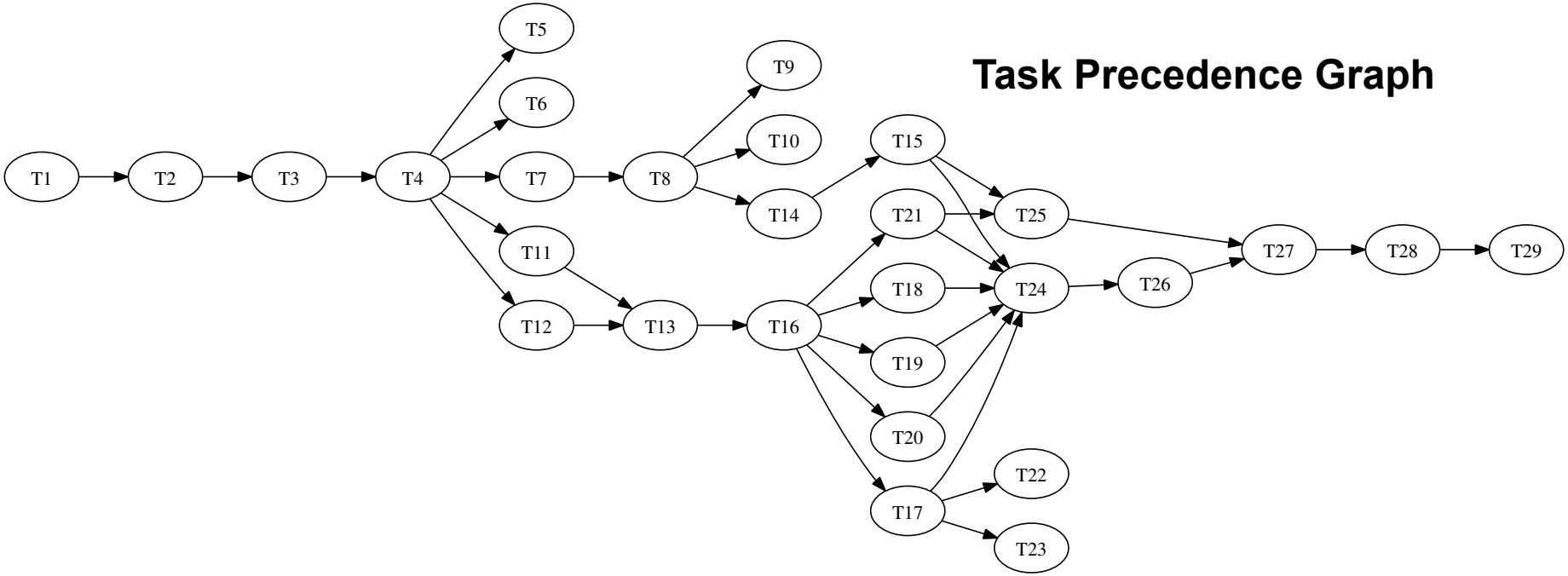
MOCell

- Cellular GA + External archive
- Ranking & Crowding from NSGA-II

Experiments: Instances

Problem instances

- 2 instances based on a MS Project repository real example: ms1 and ms2



Experiments: Instances

Problem instances

- 2 instances based on a MS Project repository real example: ms1 and ms2

Productivity Matrix

$Emp.$			Task (t_j)																											
e_i	e_i^s		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
e_1	50	ms1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		ms2	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e_2	40	ms1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1
		ms2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	.5	0	0	1	1
e_3	10	ms1	0	0	0	0	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0
		ms2	0	0	0	0	0	0	0	.3	.3	.3	0	0	.5	0	0	0	0	.5	0	.5	0	0	0	.5	0	.5	0	0	0
e_4	15	ms1	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	1	0	1	0	0	0	0	1	1	1	0
		ms2	0	0	0	0	0	0	0	1	1	1	.5	.5	.5	0	0	0	0	0	.8	0	.8	0	0	.8	.8	.8	.8	.8	0
e_5	20	ms1	0	1	1	1	1	1	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1	0
		ms2	0	.5	.5	.5	.5	.5	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	1	1	1	1	1	1	1	0
e_6	30	ms1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0
		ms2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	.8	0	0	.8	0
e_7	30	ms1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0
		ms2	0	.7	.7	.7	.7	.7	.7	.7	.7	.7	.7	.7	.7	.7	.7	0	0	0	0	0	0	0	0	1	1	1	1	1	0
t_j^c			6	680	408	8	10	10	378	10	10	10	162	48.6	8.8	720	6	198	180	6	108	6	30	36	36	18	540	120	180	450	3

Experiments: Algorithm-Specific Parameters

NSGAII	SPEA2	PAES	MOCell
Population: 100	Population: 100	Population: 1	Population: 100
Binary tournament	Binary tournament		Binary tournament
DPX (p_c=0.9)	DPX (p_c=0.9)		DPX (p_c=0.9)
Uniform mutation (p_m=1/L)	Uniform mutation (p_m=1/L)	Uniform mutation (p_m=1/L)	Uniform mutation (p_m=1/L)

Experiments: Global Parameters

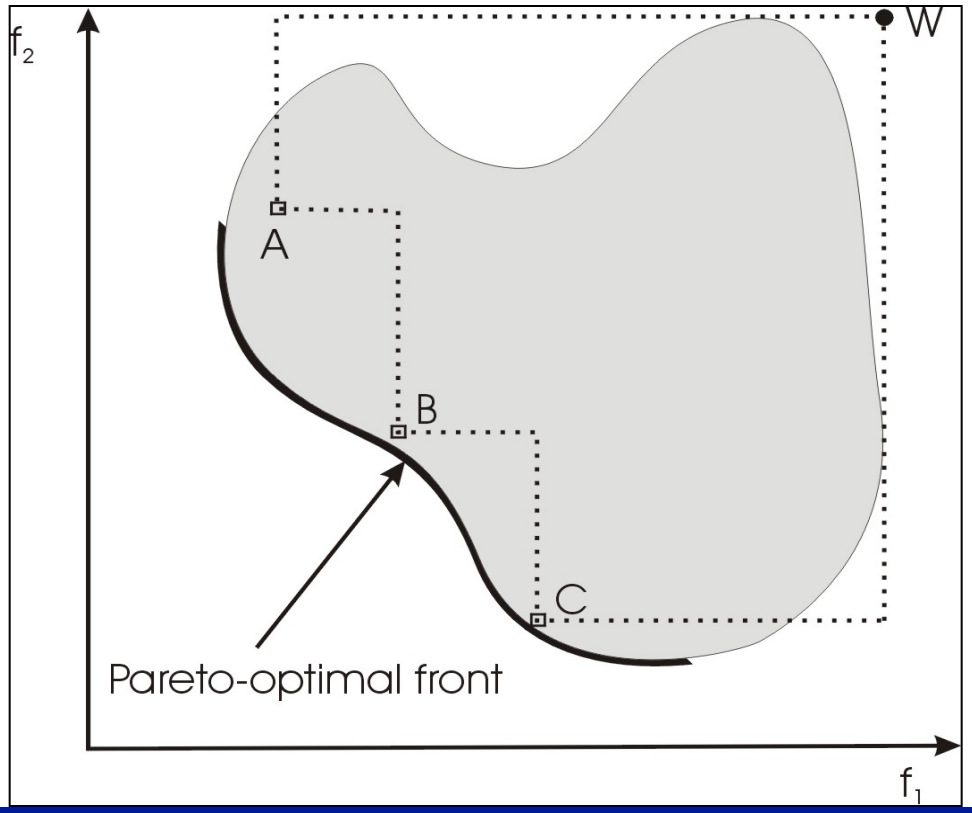
Global Parameters

- Stopping condition: 1 000 000 function evaluations
- Approximated Pareto front size: 100 solutions
- Sampling H=100
- 100 independent runs for each algorithm-instance
- Statistical tests for significance differences (95%)
- Representation: integer matrix + real vector + integer vector

Experiments: Quality Indicators

- Hypervolume (HV)
  - Volume covered by members of the non-dominated set of solutions
  - Measures both convergence and diversity in the Pareto front
  - Larger values are better

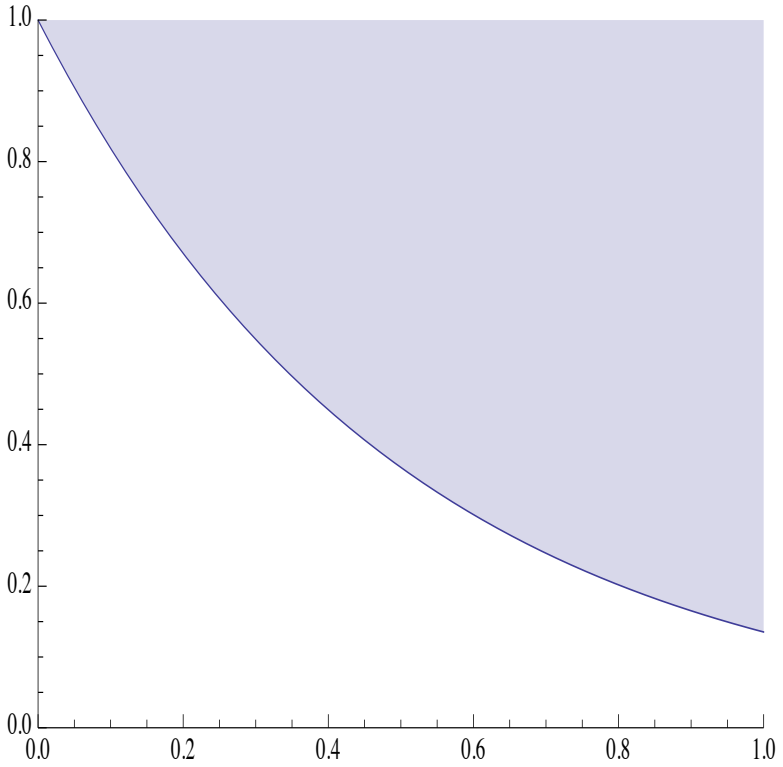
- Attainment surfaces
  - Localization statistics for fronts
  - The same as the median and the interquartile range in the mono-objective case



Experiments: Quality Indicators

- Hypervolume (HV)
  - Volume covered by members of the non-dominated set of solutions
  - Measures both convergence and diversity in the Pareto front
  - Larger values are better

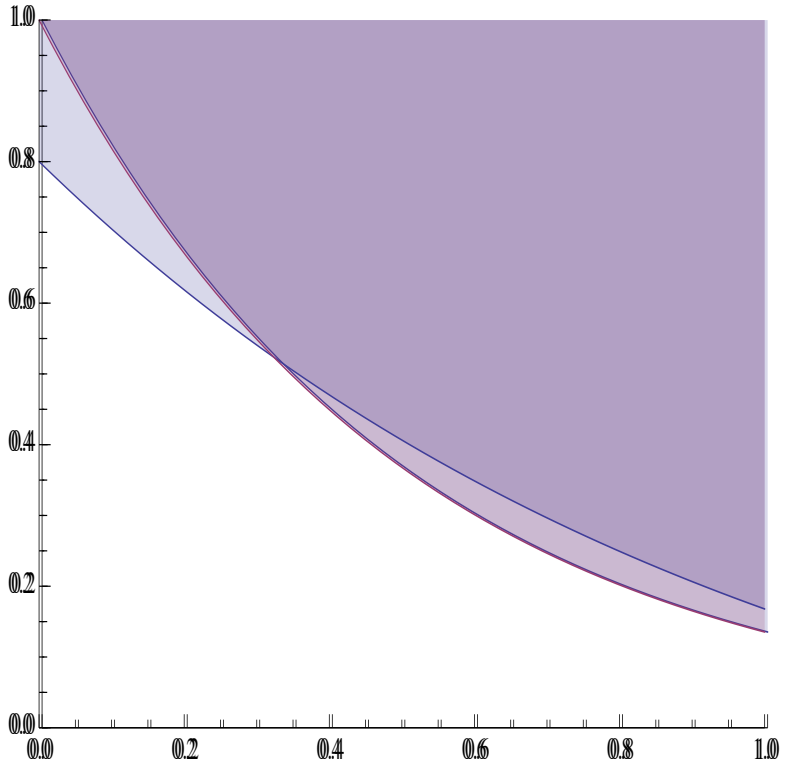
- Attainment surfaces
  - Localization statistics for fronts
  - The same as the median and the interquartile range in the mono-objective case



Experiments: Quality Indicators

- Hypervolume (HV)
  - Volume covered by members of the non-dominated set of solutions
  - Measures both convergence and diversity in the Pareto front
  - Larger values are better

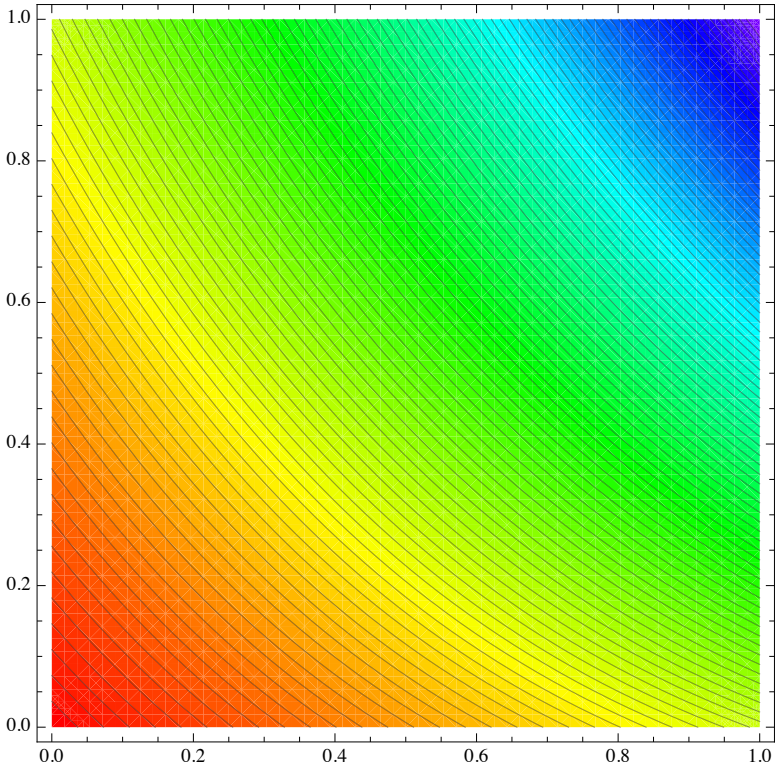
- Attainment surfaces
  - Localization statistics for fronts
  - The same as the median and the interquartile range in the mono-objective case



Experiments: Quality Indicators

- Hypervolume (HV)
  - Volume covered by members of the non-dominated set of solutions
  - Measures both convergence and diversity in the Pareto front
  - Larger values are better

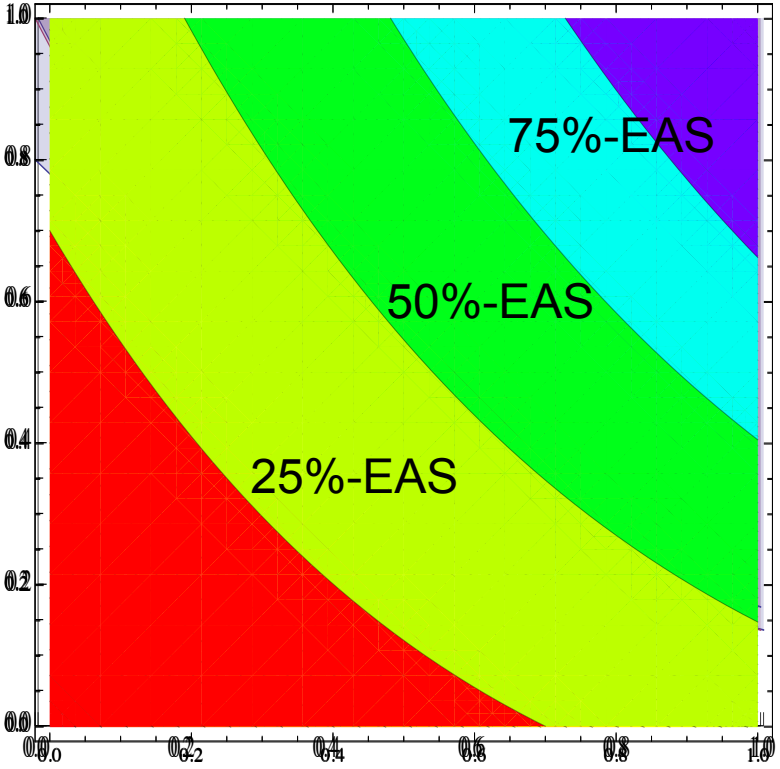
- Attainment surfaces
  - Localization statistics for fronts
  - The same as the median and the interquartile range in the mono-objective case



Experiments: Quality Indicators

- Hypervolume (HV)
  - Volume covered by members of the non-dominated set of solutions
  - Measures both convergence and diversity in the Pareto front
  - Larger values are better

- Attainment surfaces
  - Localization statistics for fronts
  - The same as the median and the interquartile range in the mono-objective case



Results: Hypervolume Comparison

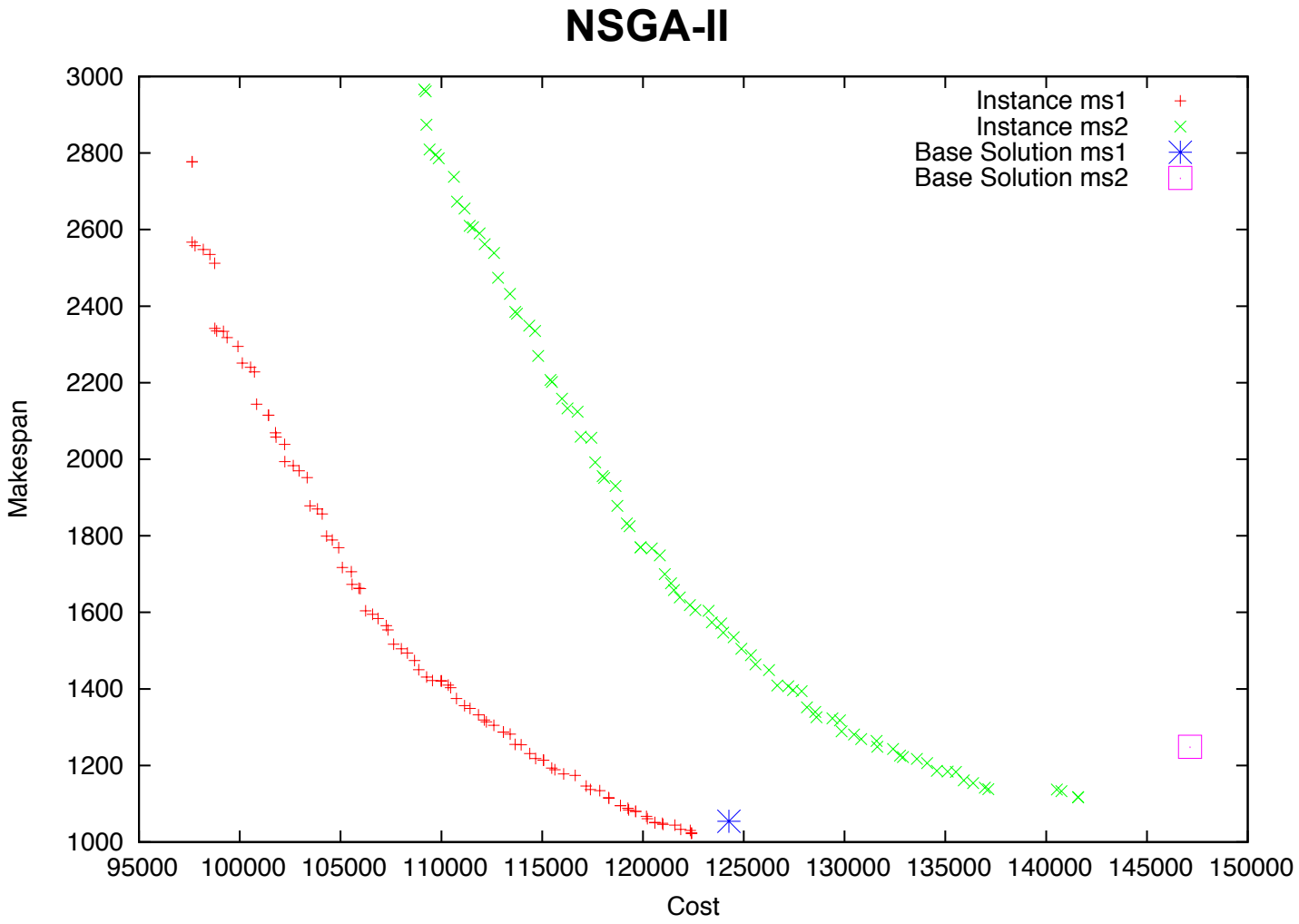
Hypervolume (HV)

- NSGA-II and MOCell are the best algorithms
- NSGA-II is specially good in robust versions of the problem
- MOCell is good in the non-robust version
- PAES is the worst algorithm in the comparison
- Running time between 2.5 and 5 minutes in NR and around 5 hours in OTR and STR

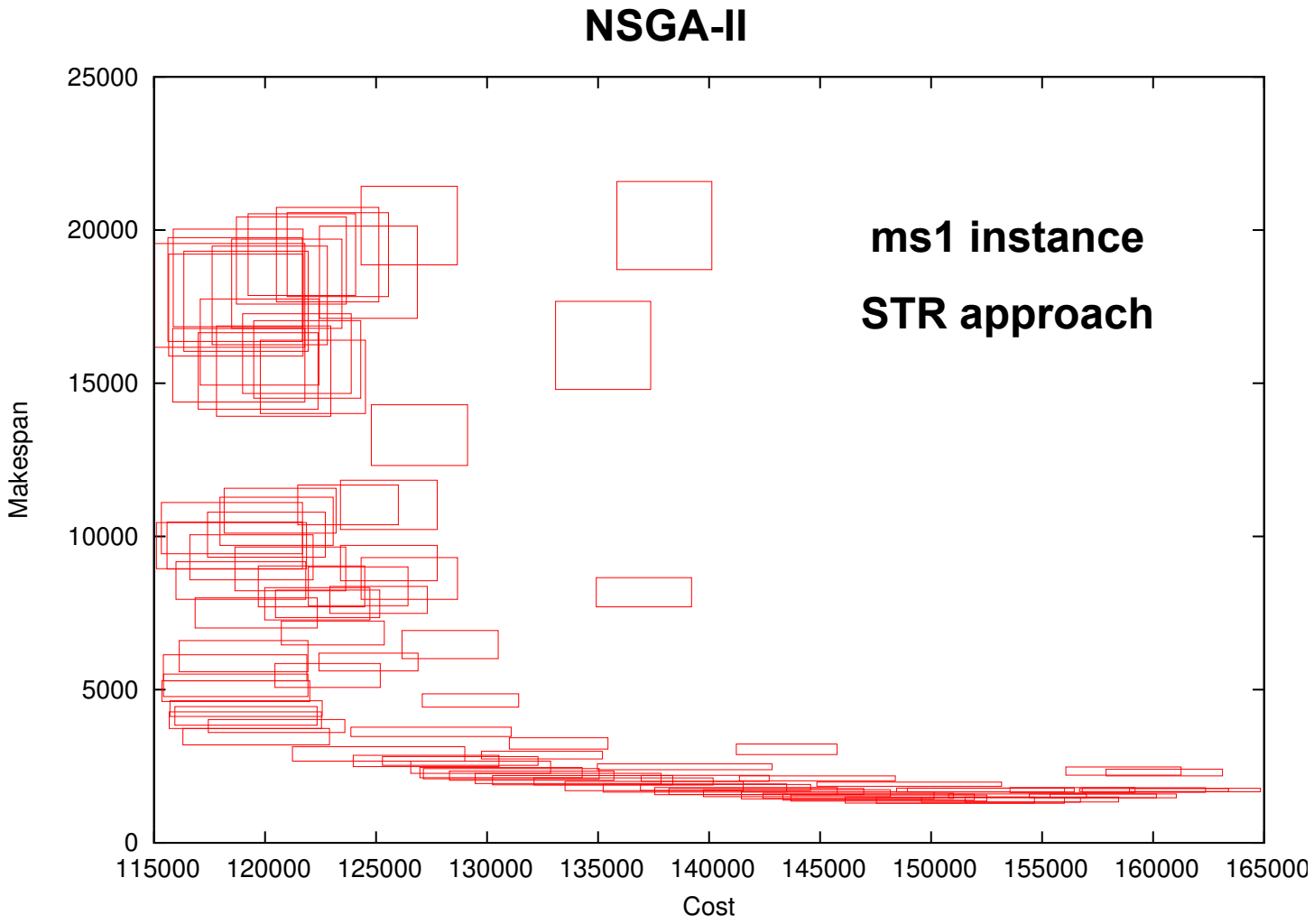
Median and interquartile range

	NSGAII	SPEA2	PAES	MOCell	NSGAII	SPEA2	PAES	MOCell
Rob.	ms1				ms2			
NR	$0.943^{*}_{0.000}$	$0.943^{*}_{0.000}$	$0.518^{*}_{0.065}$	$0.944_{0.000}$	$0.904^{*}_{\pm 0.000}$	$0.905^{*}_{\pm 0.001}$	$0.543^{*}_{\pm 0.031}$	$0.905_{\pm 0.000}$
OTR	$0.829^{*}_{0.027}$	$0.807^{*}_{0.030}$	$0.328^{*}_{0.039}$	$0.816_{0.032}$	$0.738_{\pm 0.025}$	$0.730_{\pm 0.018}$	$0.287^{*}_{\pm 0.020}$	$0.695^{*}_{\pm 0.043}$
STR	$0.746_{0.028}$	$0.688^{*}_{0.063}$	$0.345^{*}_{0.036}$	$0.742_{0.025}$	$0.764_{\pm 0.025}$	$0.717^{*}_{\pm 0.030}$	$0.387^{*}_{\pm 0.032}$	$0.769_{\pm 0.022}$

Results: Comparison with a (Human) Base Solution

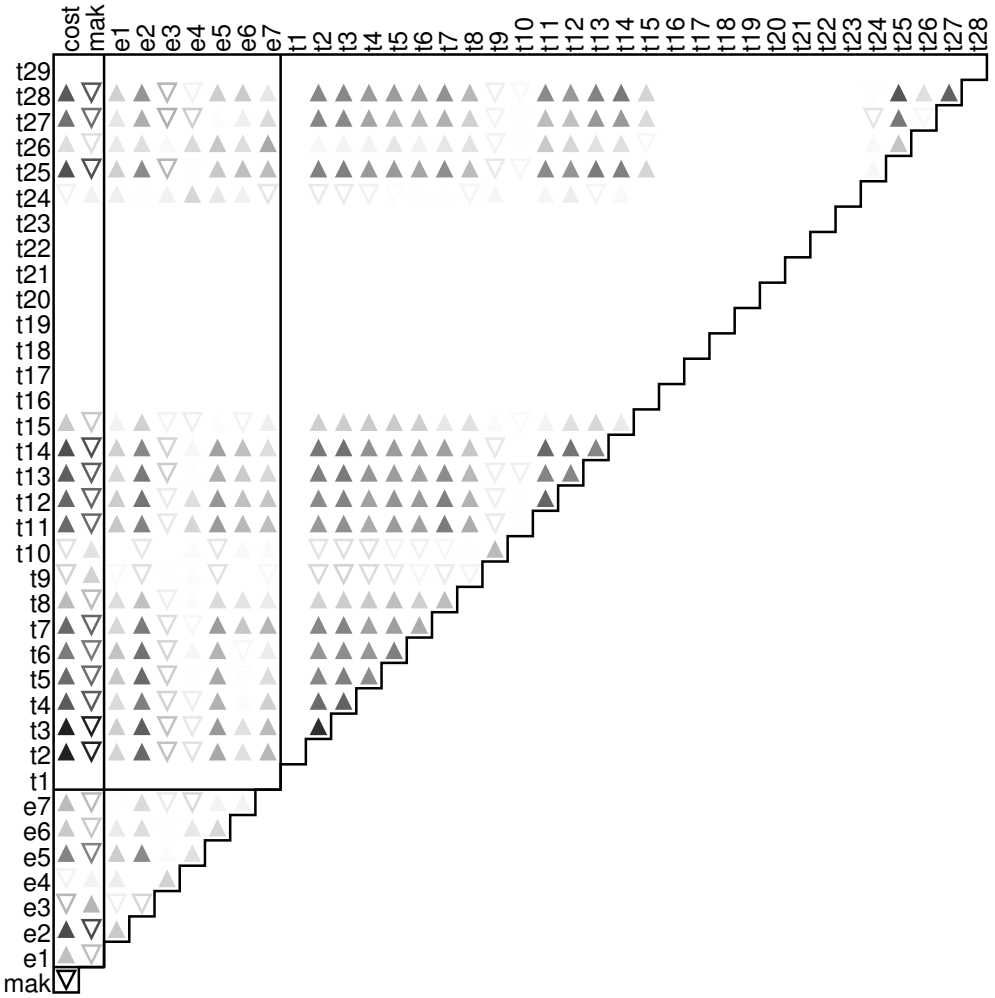


Results: 50%-Attainment Surface



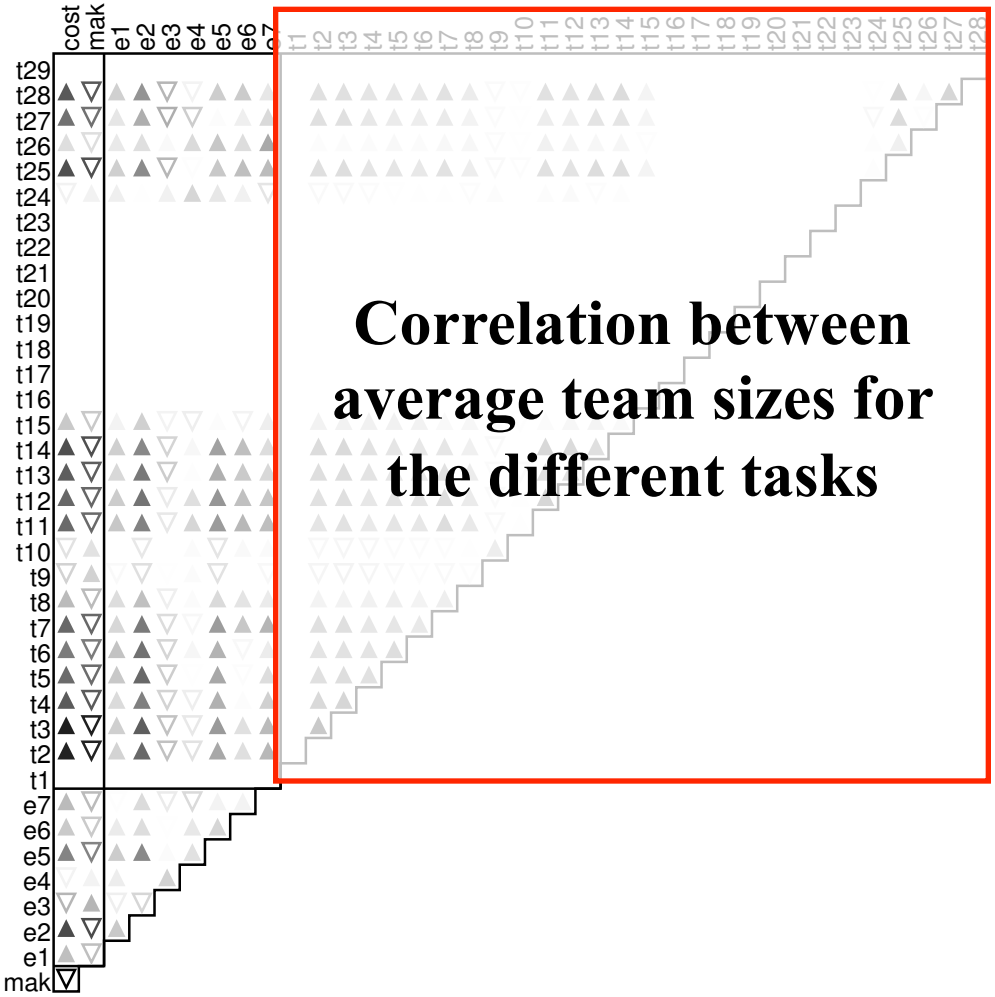
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients of the solutions in an approximated Front**
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- **An example for an approximated Pareto front of MOCell using the NR approach in the ms2 instance**



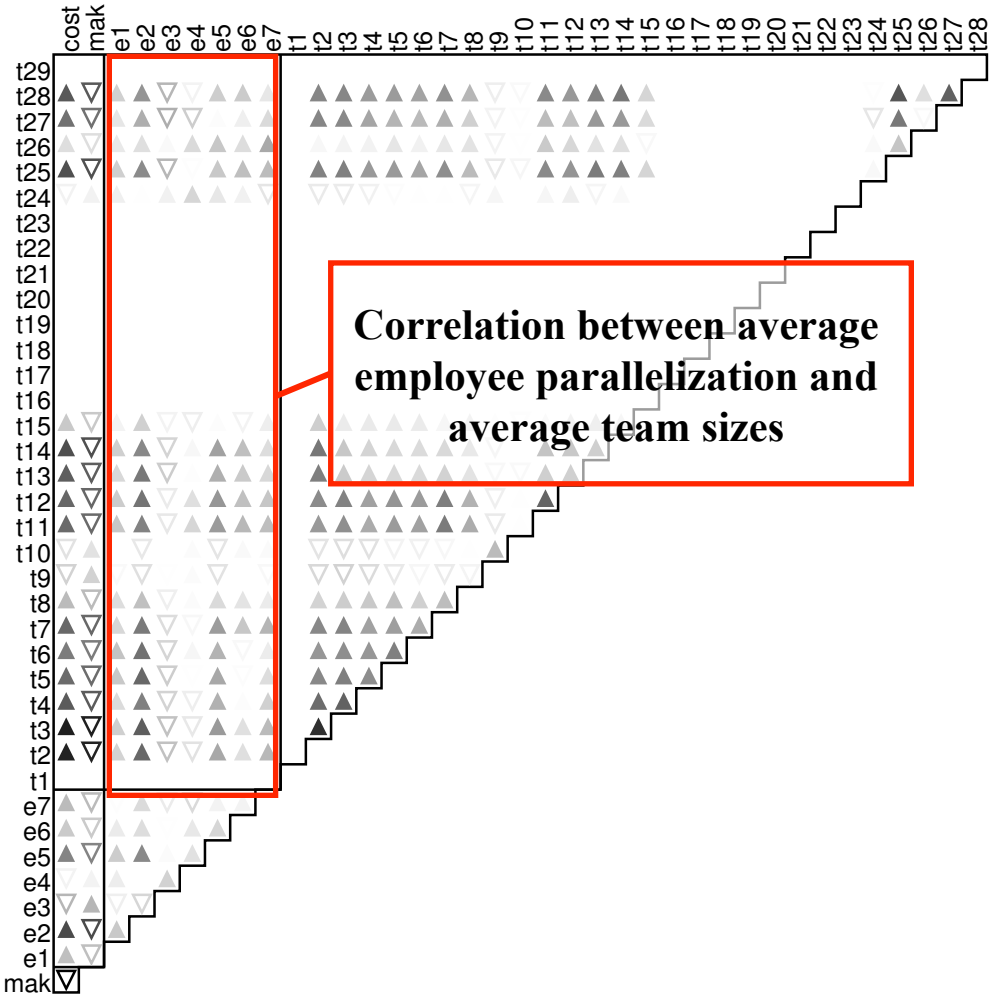
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients** of the solutions in an approximated Front
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- An example for an approximated Pareto front of **MOCell** using the NR approach in the **ms2** instance



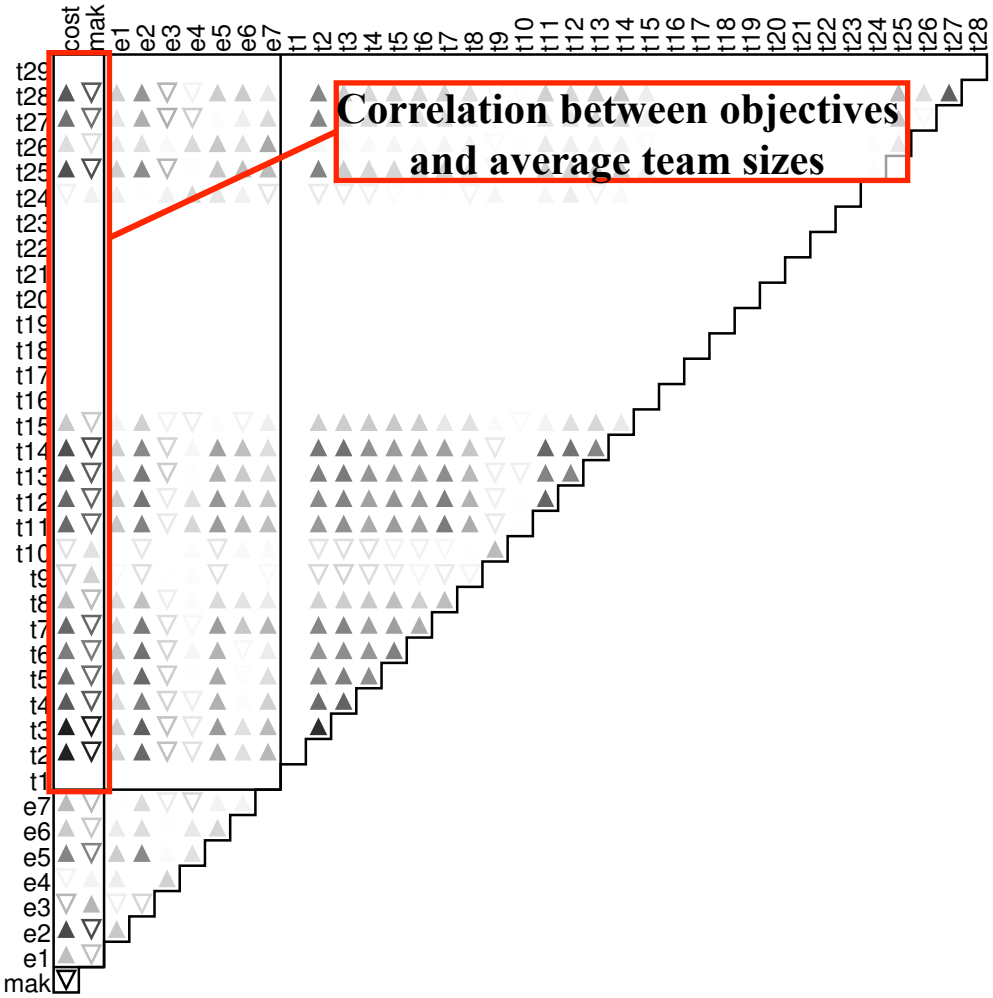
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients** of the solutions in an approximated Front
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- An example for an approximated Pareto front of **MOCell** using the NR approach in the **ms2** instance



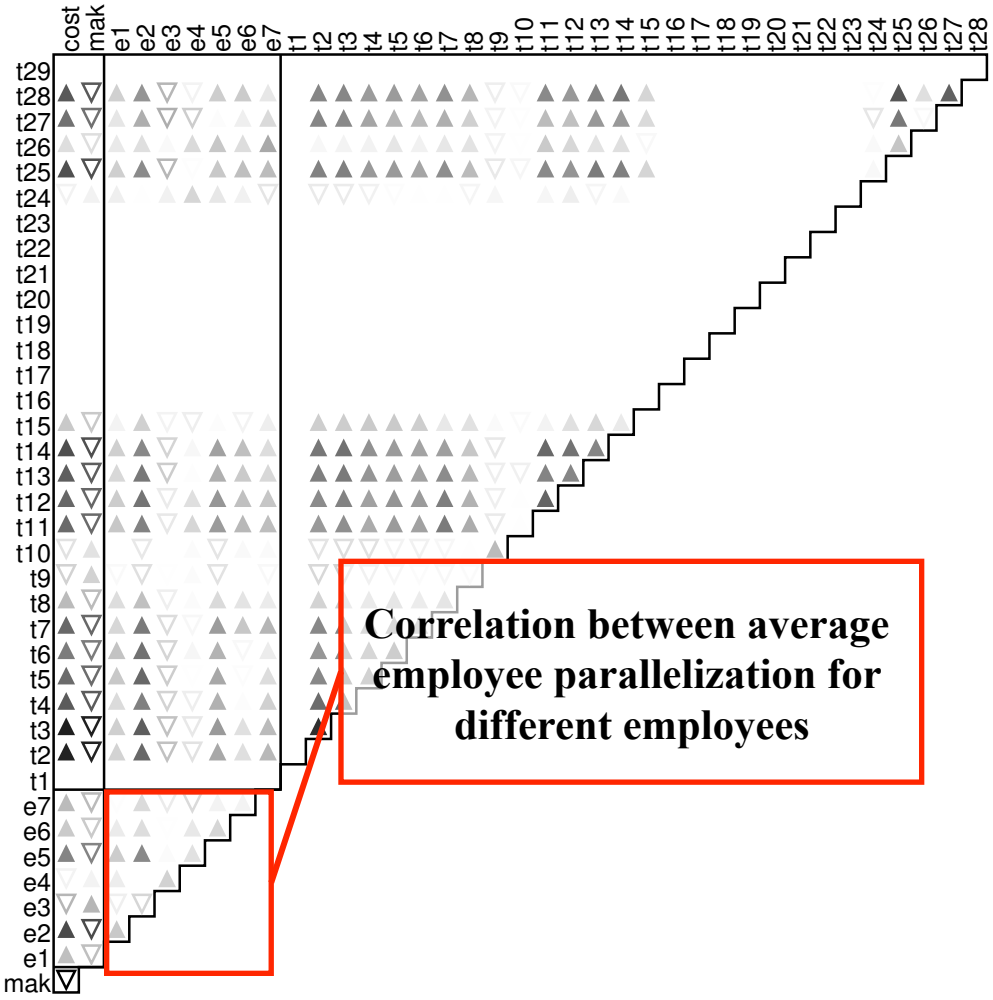
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients** of the solutions in an approximated Front
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- An example for an approximated Pareto front of **MOCell** using the NR approach in the **ms2** instance



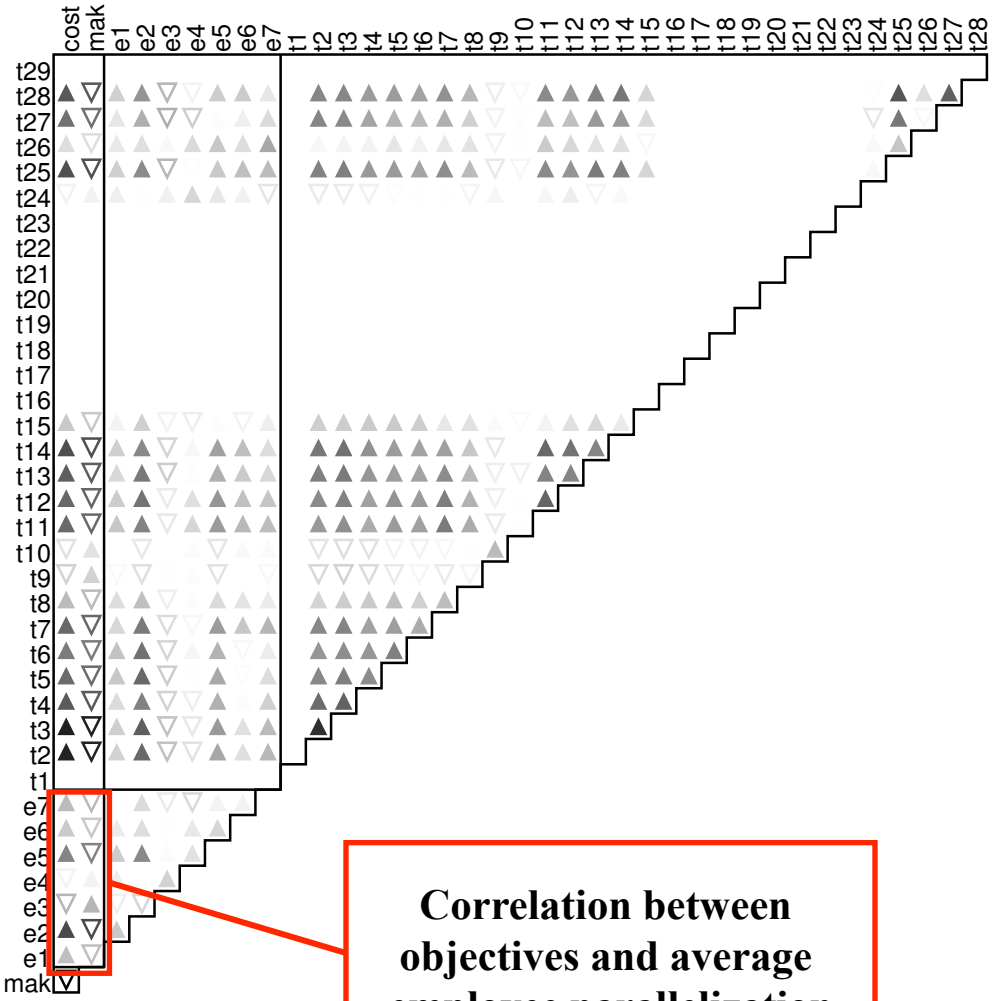
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients** of the solutions in an approximated Front
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- An example for an approximated Pareto front of **MOCell** using the NR approach in the **ms2** instance



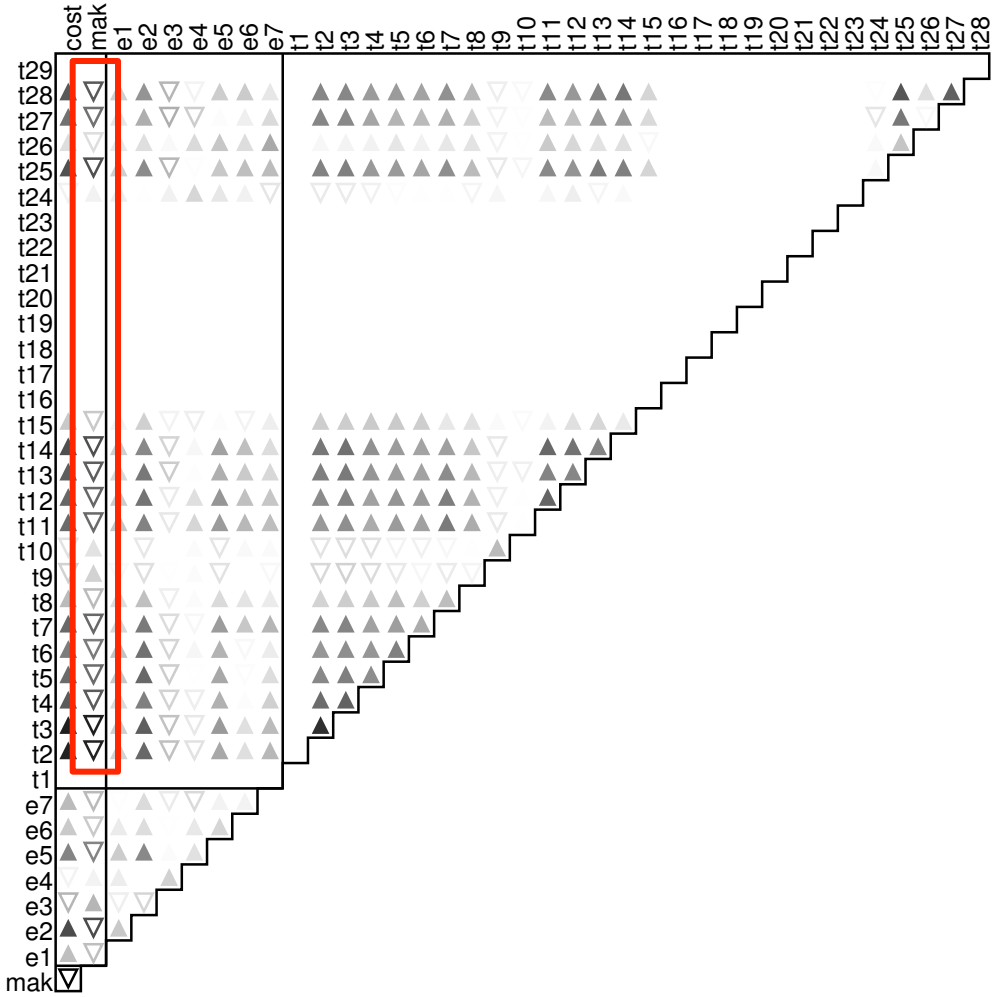
Results: Analysis of the Solution Features

- **Spearman rank correlation coefficients** of the solutions in an approximated Front
  - ▲ : positive correlation
  - ▽ : negative correlation
  - Gray scale: absolute value of correlation
- An example for an approximated Pareto front of **MOCell** using the NR approach in the **ms2** instance



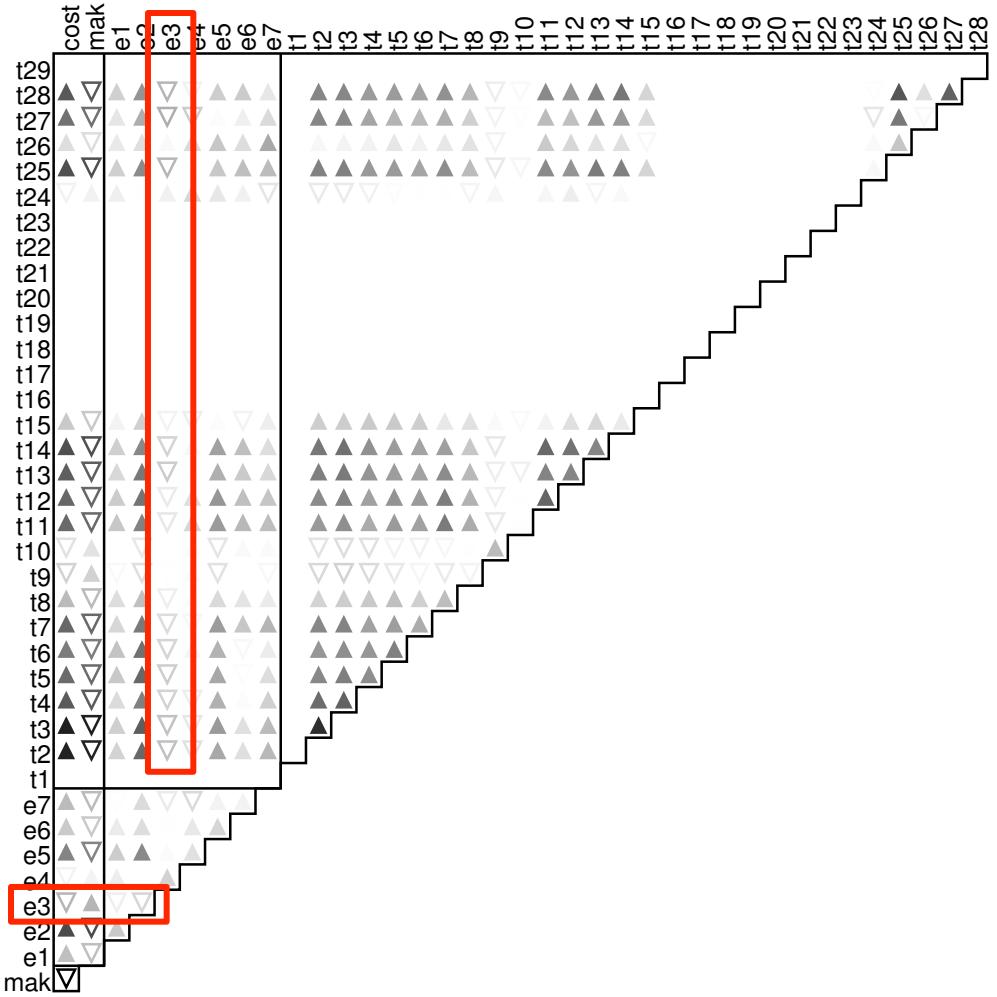
Results: Analysis of the Solution Features

- **Increasing the size of the working teams the makespan is reduced**



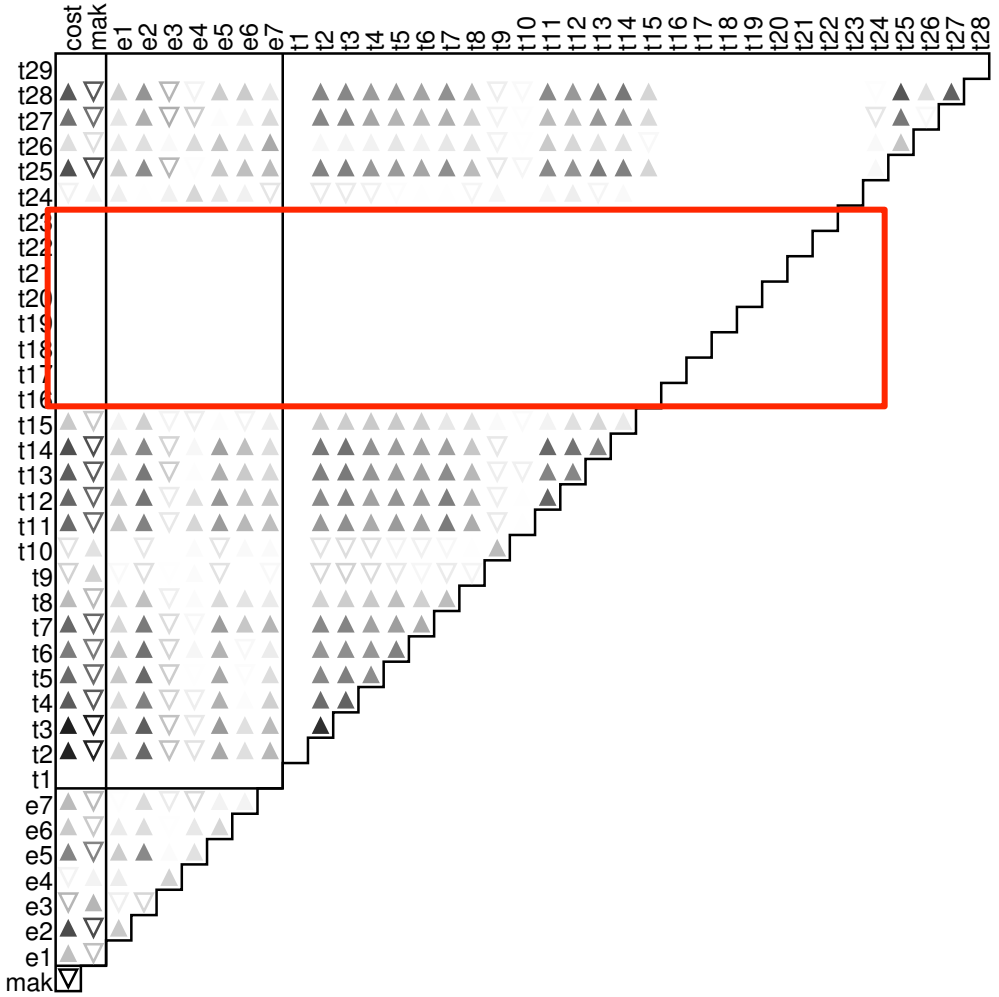
Results: Analysis of the Solution Features

- **Increasing the size of the working teams the makespan is reduced**
- **Employee e_3 is the only one able to perform a task in the critical path**



Results: Analysis of the Solution Features

- **Increasing the size of the working teams** the makespan is reduced
- **Employee e_3** is the only one able to perform a task in the critical path
- **No correlation** is observed in tasks for which only one employee can do the work



Conclusions & Future Work

Conclusions

- A new formulation for the SPS problem is presented
- The new formulation includes robustness
- We compare four metaheuristics
- NSGA-II and MOCell are the best algorithm according to HV
- The algorithms outperform human made solutions
- The analysis of the solutions reveals relevant information for the project manager

Future Work

- Use real-world instances of the problem
- Different robustness approaches
- New operators and search methods for the problem

Thanks for your attention !!!

| M*: Multidisciplinar Multiobjective Metaheuristics
| TIN2008-06491-C04-01 Ministerio de Ciencia e Innovación (Spain)
| UMA - University of Málaga