FACULTY OF ECONOMICS AND
BUSINESS ADMINISTRATION

DEPARTMENT OF BUSINESS INFORMATICS AND OPERATIONS MANAGEMENT

Sports scheduling: from consulting to science

Dries Goossens (Ghent University)

Contact
dries.goossens@ugent.be
www.sportscheduling.ugent.be

GHENT
UNIVERSITY

Scheduling seminar, May 10 2023

SPORTS SCHEDULING

The problem:

given a set of matches,

given a set of rounds (time slots),

decide which matches are scheduled on which rounds.

Each match A – B is played at the venue of the home team (A), with the opponent (B) being the away team.

We focus on double round robin tournament (2RR):

each team plays twice against each other team: once at home, once away

GHENT UNIVERSITY

SPORTS SCHEDULING

The problem:

given a set of matches,

given a set of rounds (time slots),

decide which matches are scheduled on which rounds.

Assumptions:

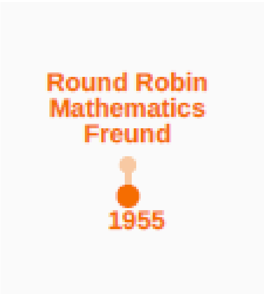
- a team can play at most one match per round
- time-constrained schedules

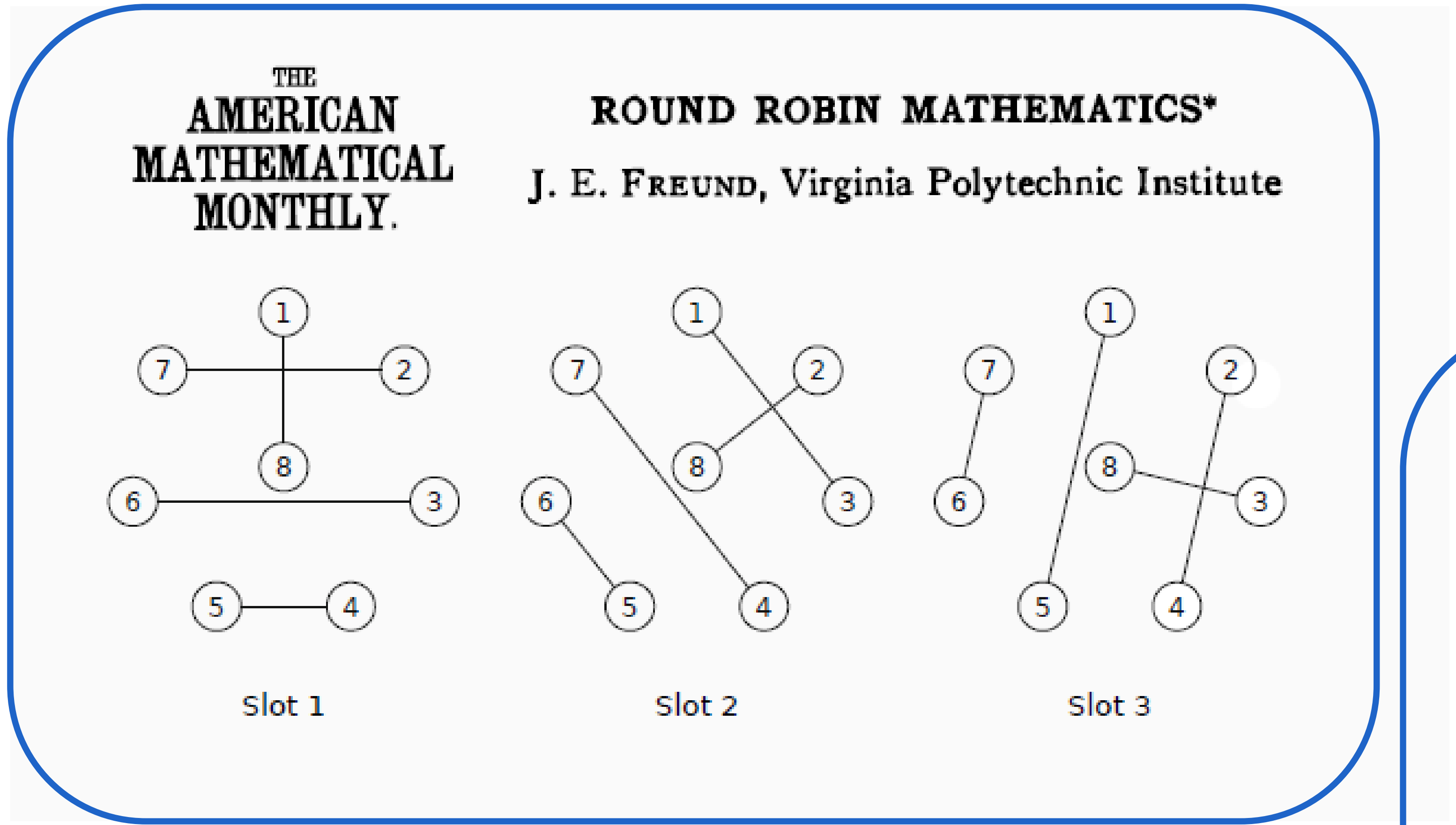
R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
A-B	B-E	B-D	B-C	F-B	B-A	E-B	D-B	C-B	B-F
C-D	D-A	A-F	E-A	D-E	D-C	A-D	F-A	A-E	E-D
E-F	F-C	E-C	F-D	A-C	F-E	C-F	C-E	D-F	C-A

A time-constrained schedule for a 2RR tournament with 6 teams (A-F)

A BRIEF HISTORY

Scheduling round robin tournaments





Thomas P. Kirkman (1806 – 1895)

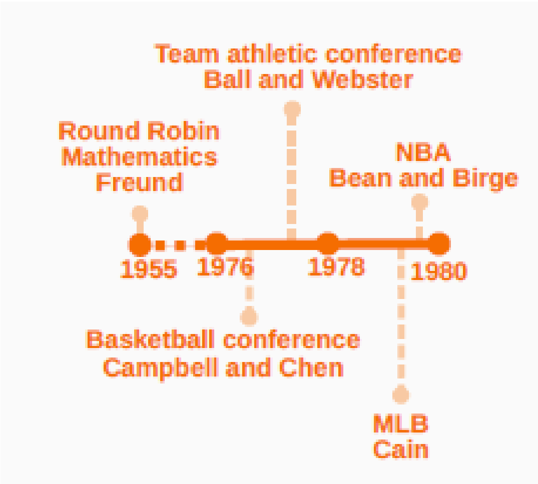
Constructs a feasible schedule (for any number of teams) with the clock method (1847)

The resulting schedule is called a “canonical schedule”.

GHENT UNIVERSITY

A BRIEF HISTORY

The 60's and 70's: a lack of computational tools



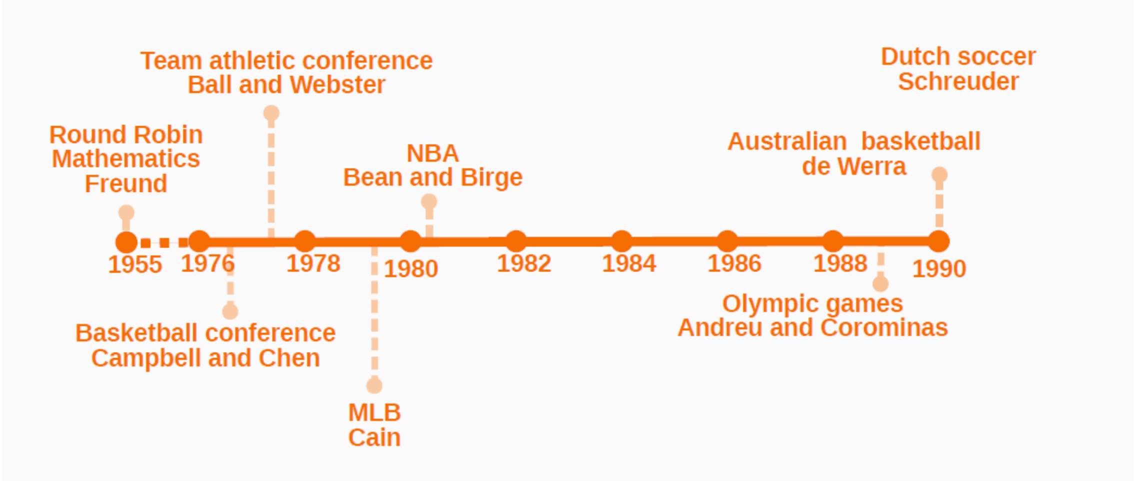
Although the above formulation shows that zero-one programming is a method for solving the problem, the number of decision variables and constraints required makes existing algorithms too large to handle on the available computer facilities. [1, p. 66-91].

1975: IBM 5100

This program has 41976 constraints and 873,136 variables. Suggested formulations of this type for the real problem are at least twice this size. The cost of solving such a program by standard methods would most likely outweigh the savings derived from the improved schedule.

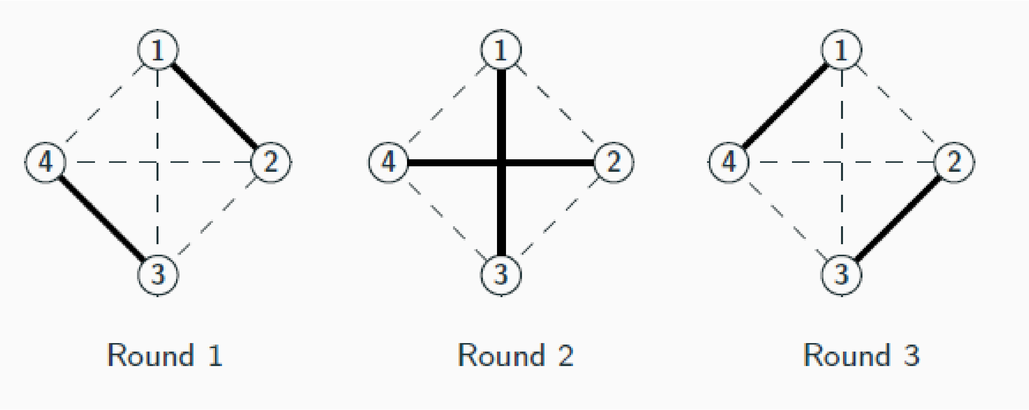
A BRIEF HISTORY

The 80's: how to minimize breaks

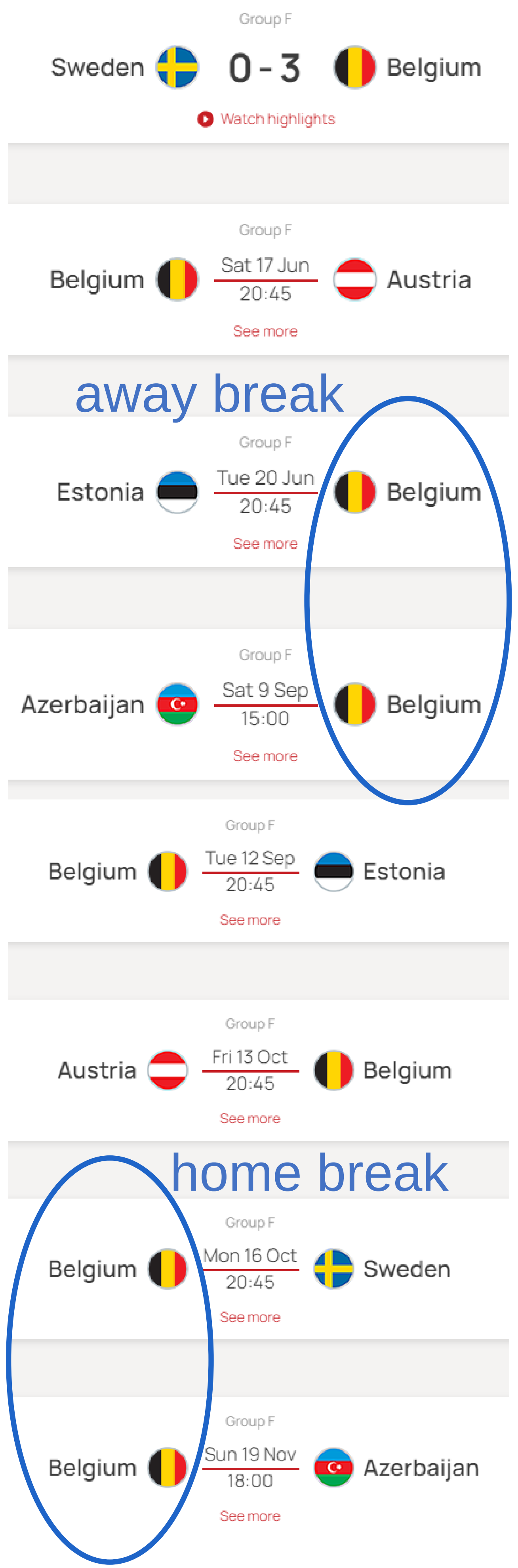


Theoretical work by de Werra

- 1RR ~ 1-factorization of K_{2n}



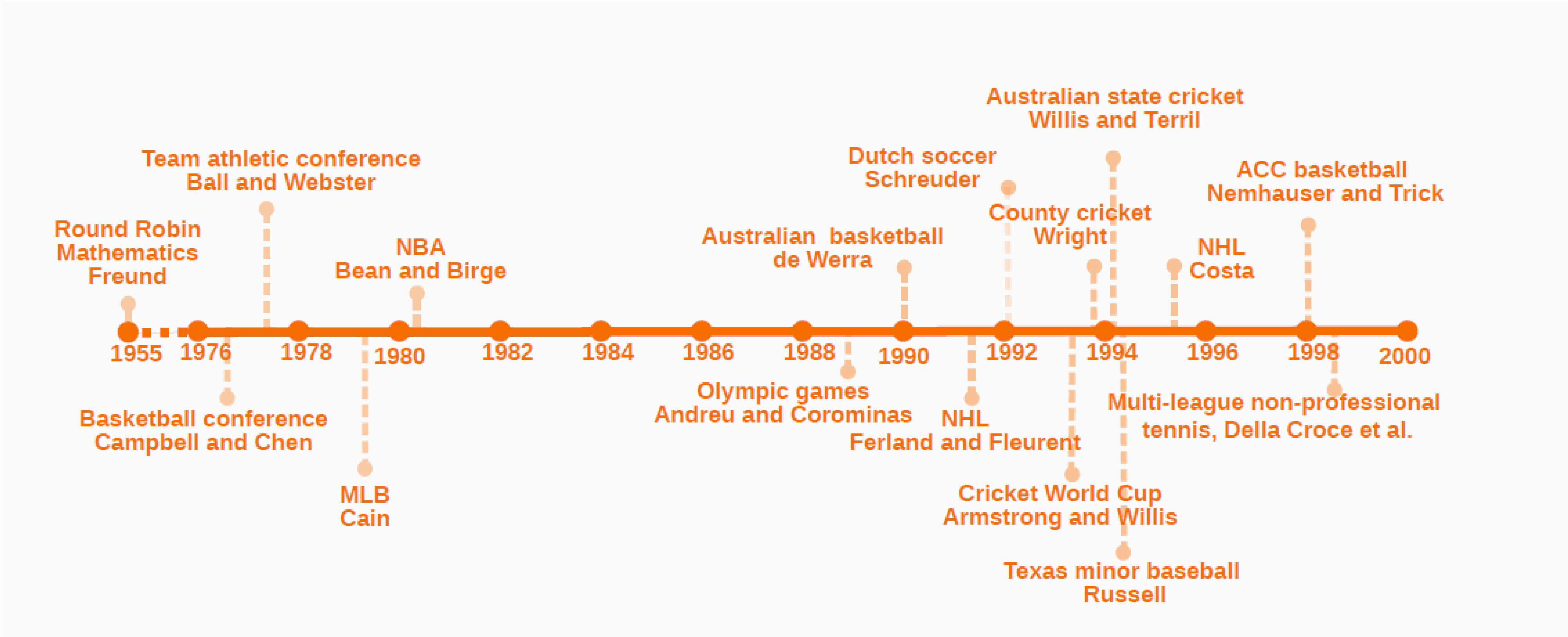
- Orientation of edges to minimize breaks



GHENT UNIVERSITY

A BRIEF HISTORY

The (late) 90's: first real-life applications

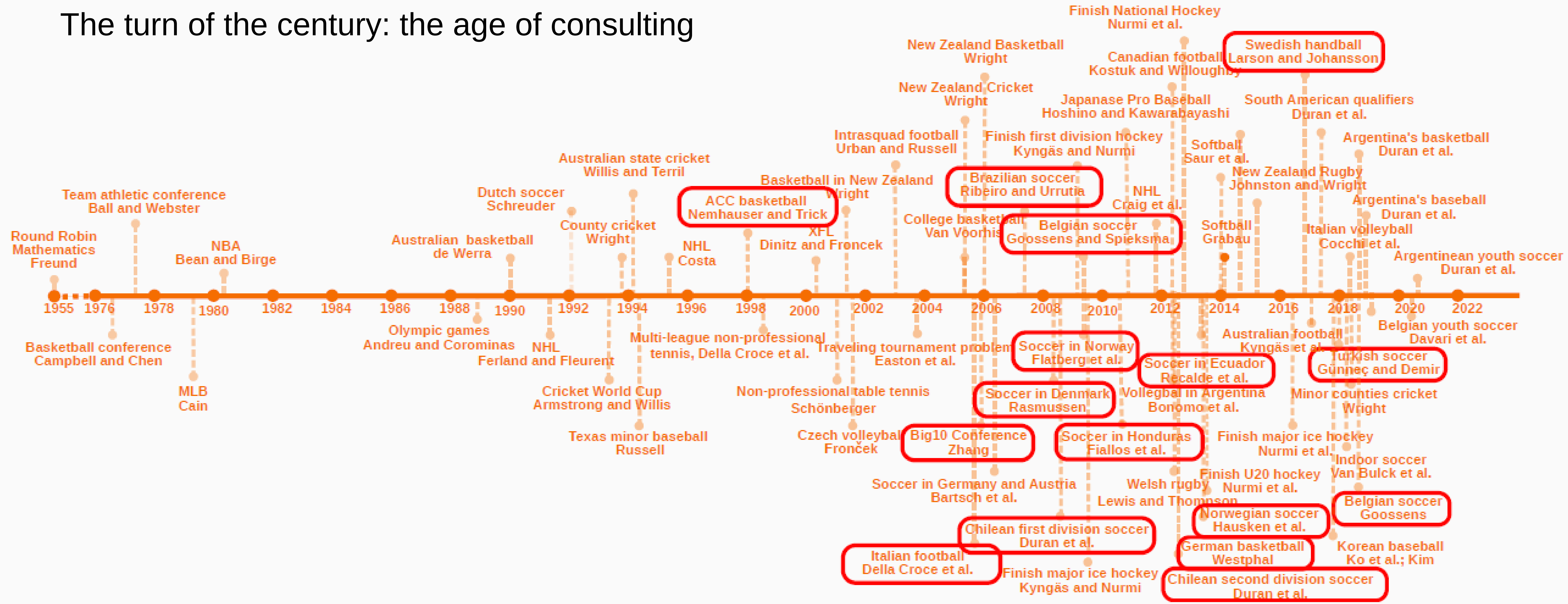


- Computing power increases
- Methodological breakthroughs
- Manual solutions still dominant
- Craftspeople, e.g. Henri & Holly Stephenson (MLB)

GHENT UNIVERSITY

A BRIEF HISTORY

The turn of the century: the age of consulting



A BRIEF HISTORY

The turn of the century: the age of consulting

Lack of benchmarking

- The literature consists mostly of specific case studies, with tailor-made algorithms.
- Problem instances are rarely shared.
- Algorithms are almost never benchmarked.

Lack of generality & understanding

- To what extent do approaches work well on other sports scheduling problems?
- When do algorithms work well, and why?
- Can we develop a general solver that can handle a wide variety of constraints?

GHENT UNIVERSITY

FROM CONSULTING TO SCIENCE

1. **A classification scheme for round robin tournament timetabling problems**
2. A standard problem instance data format
3. A benchmark instance set
4. General sport scheduling solvers
5. Algorithm selection & insights

1. CLASSIFICATION SCHEME

Each sport scheduling problem can be classified using 3 fields:

- α : competition format
- β : constraints in use
- γ : objective function

[inspired by the notation for machine scheduling problems by Graham et al. (1979)]

1. CLASSIFICATION SCHEME

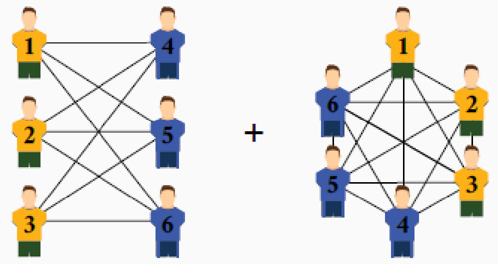
α : competition format

Alpha 1

Tournament structure

- k round-robin (kRR)
- k bipartite round-robin (kBRR)
- Non-round robin (NRR)

Example of a NRR: a 1BRR combined with a 1RR



Alpha 2

Compactness

- Time constrained (C)
- Time relaxed (R)

Team	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	+2	+4	-6	+5	-3	-2	-4	+6	-5	+3
2	-1	-5	-4	+3	-6	+1	+5	+4	-3	+6
3	+4	+6	+5	-2	+1	-4	-6	-5	+2	-1
4	-3	-1	+2	-6	+5	+3	+1	-2	+6	-5
5	-6	+2	-3	-1	-4	+6	-2	+3	+1	+4
6	+5	-3	+1	+4	+2	-5	+3	-1	-4	-2

Team	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
1	+2		+4	-6	+5			-3	-2		-4	+6	-5			+3
2	-1	-5			-4	+3	-6		+1	+5			+4	-3	+6	
3		+4	+6	+5		-2		+1		-4	-6	-5		+2		-1
4		-3	-1		+2	-6	+5			+3	+1		-2	+6	-5	
5	-6	+2		-3	+1		-4		+6	-2		+3	-1		+4	
6	+5		-3	+1		-4	+2		-5		+3	-1		+4	-2	

Alpha 3

Symmetry

- None (∅)
- Phased (P)
- Mirrored (M)
- Inverse (I)
- English (E)
- French (F)

1. CLASSIFICATION SCHEME

β : constraints in use

5 constraint groups:

- Capacity constraint (CA1 – CA5)
- Game constraints (GA1 – GA2)
- Break constraints (BR1 – BR4)
- Fairness constraints (FA1 – FA6)
- Separation constraints (SE1 – SE2)

1. CLASSIFICATION SCHEME

β : constraints in use

For each constraint c, we define:

- hard or soft constraint
- a deviation vector D_c
- cost function f_c
- constraint weight w_c

Five cost functions:

- Sum: $\sum_{i=1}^{q} d_i$
- Sum-squares: $\sum_{i=1}^{q} d_i^2$
- Square-sum: $(\sum_{i=1}^{q} d_i)^2$
- Min: $\min_{d_i \in D_c} d_i$
- Max: $\max_{d_i \in D_c} d_i$

$$p_c = w_c f_c(D_c)$$

GHENT UNIVERSITY

1. CLASSIFICATION SCHEME

γ : objective function

- No objective (∅)
- Minimum (weighted) breaks (BR)
- Travel distance minimization (TR)
- Cost minimization (CR)
- Minimum (weighted) carry-over effect value (CO)
- Minimum soft constraint violation (SC)

GHENT UNIVERSITY

1. CLASSIFICATION SCHEME

Notation overview

α	α_1	kRR, kBRR, NRR
	α_2	C,R
	α_3	$\emptyset$, M, I, E, F, P
β	β_1	CA1, CA2, CA3, CA4, CA5
	β_2	GA1, GA2
	β_3	BR1, BR2, BR3, BR4
	β_4	FA1, FA2, FA3, FA4, FA5, FA6
	β_5	SE1, SE2
γ	γ_1	$\emptyset$, BR, TR, CR, CO, SC

GHENT UNIVERSITY

1. CLASSIFICATION SCHEME

Classifying the literature

Paper reference	Description
Bean and Birge (1980)	NRR, R, ∅ | CA1, CA3 | TR
Ball and Webster (1977)	2RR, R, P | GA2, CA3, | TR
Bao and Trick (2010)	2RR, R, ∅ | CA3, SE1| TR
Bartsch et al. (2006)	2RR, C, M | $BR1^{H,S}$, $BR2^{H}$, $CA1^{H}$, $CA3^{H,S}$, $CA4^{H}$, $GA1^{H}$, $SE1^{H}$ | SC
Bartsch et al. (2006)	2RR, C, E | $BR1^{H}$, $CA1^{H,S}$, $CA2^{H}$, $CA3^{S}$, $CA4^{H}$, $GA1^{H}$, $SE1^{H}$ | SC
Bonomo et al. (2012)	2RR, C, M | BR1, CA1, CA3 | TR
Briskorn and Drexl (2009a)	1RR, C, ∅ | BR1, BR2, CA1, CA3, CA4, GA1| CR
...	...
Westphal (2014)	2RR, C, P | $BR2^{S}$, $CA1^{H,S}$, $FA6^{S}$, $GA1^{S}$ | SC
Wright (2006)	2RR, C, ∅ | $CA1^{S}$, $CA2^{S}$, $CA4^{S}$, $FA2^{S}$, $FA5^{S}$, $GA1^{S}$ | SC
Zhang (2002)	NRR, C, ∅ | CA1, CA2, CA3, CA4 | ∅

1. CLASSIFICATION SCHEME

Online query tool

Alpha: competition format

Alpha 1: 1RR 2RR 3RR 4RR BRR NRR ML

Alpha 2: C R

Alpha 3: M I E F P ø

Constraints in use

Capacity: CA1 CA2 CA3 CA4 CA5

Game: GA1 GA2

Break: BR1 BR2 BR3 BR4

Fairness: FA1 FA2 FA3 FA4 FA5 FA6

Separation: SE1 SE2

Objective function

Objective: BM TR CR SC CO ø

Search

Data set	Ref.	No. Teams	No. Slots	Classification
XFL American Football	Dinitz & Froncek [1]	8	10	NRR, C, ø | BR1, BR2, CA1, CA2, CA3, CA4, SE1 | SC
Finish U20 hockey	Nurmi et al. [2]	15	62	3RR, R, ø | BR2, CA1, CA2, CA4, FA2, GA1, SE1 | SC
Belgian soccer	Goossens & Spieksma [3]	18	34	2RR, C, M | BR1, BR2, CA1, CA2, CA3, CA4, GA1 | SC
Brazilian soccer	Ribeiro & Urrutia [4]	20 - 22	38 - 42	2RR, C, M | BR1, BR2, CA1, CA2, CA4, FA1 | SC
Danish soccer	Rasmussen [5]	12	3 - 33	3RR, C, P | BR1, BR2, CA1, CA2, CA3, CA4, GA1, SE1 | SC
Swedish handball	Larson & Johansson [6]	14	33	NRR, C, ø | BR2, CA1, CA2, CA4, FA1, FA3, SE2 | SC
Italian volleyball	Cocchi et al. [7]	14	26	2RR, C, M | BR1, BR2, CA1, CA2, CA3, CA4, FA1, FA6, GA1 | SC
Non-professional table tennis 3	Knust [8]	10	100	1RR, R, ø | BR2, CA1, CA2, CA3, FA1, GA1 | SC
English soccer holidays	Kendall [9]	92	2	ML, C, ø | BR2, CA1, CA2, CA4, FA6 | SC
Australian Football	yng\a [10]	18	138	NRR, R, ø | BR1, BR2, CA1, CA2, CA3, CA4, FA5, GA1, SE1 | SC

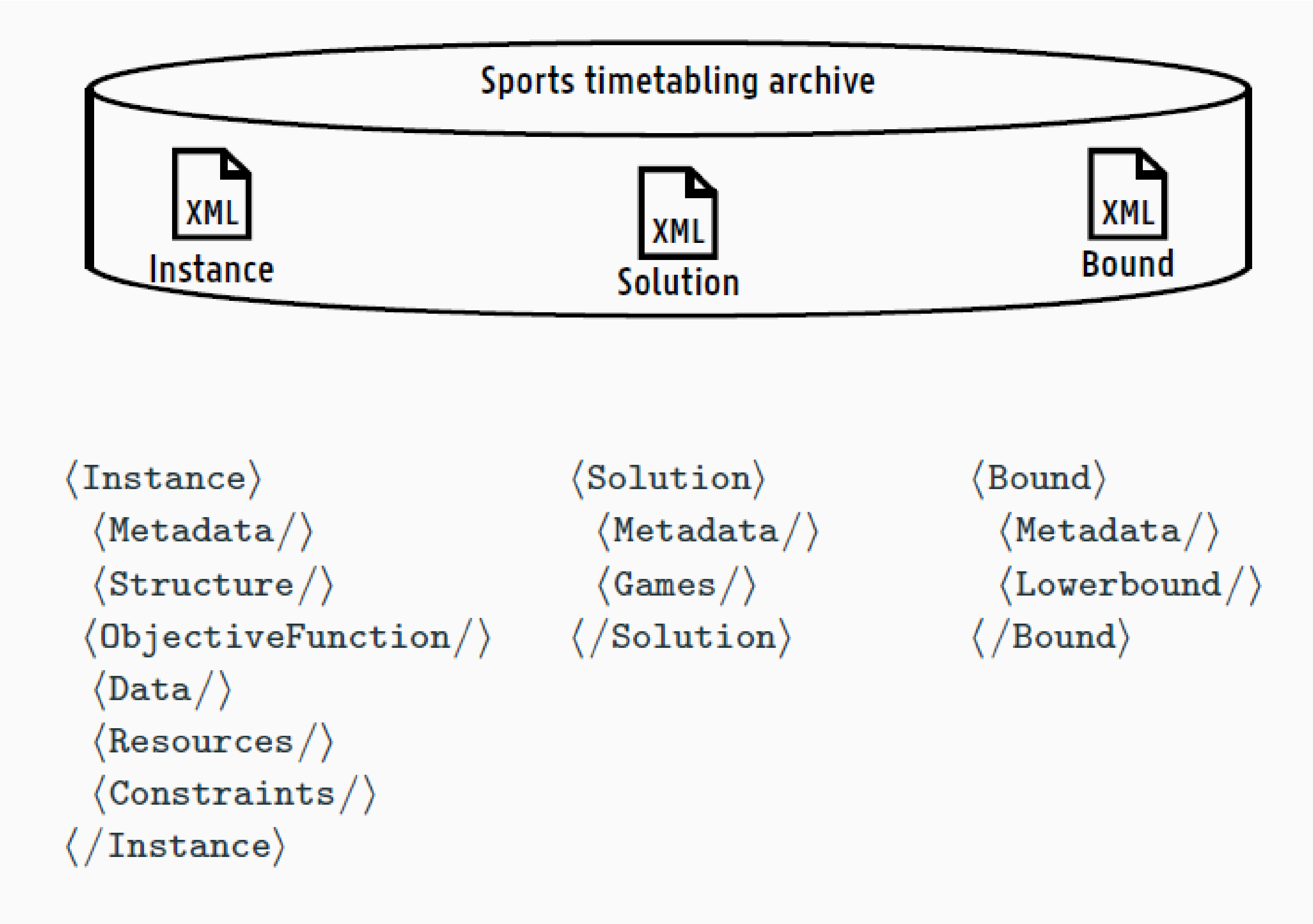
www.sportscheduling.ugent.be/RobinX/query.php

FROM CONSULTING TO SCIENCE

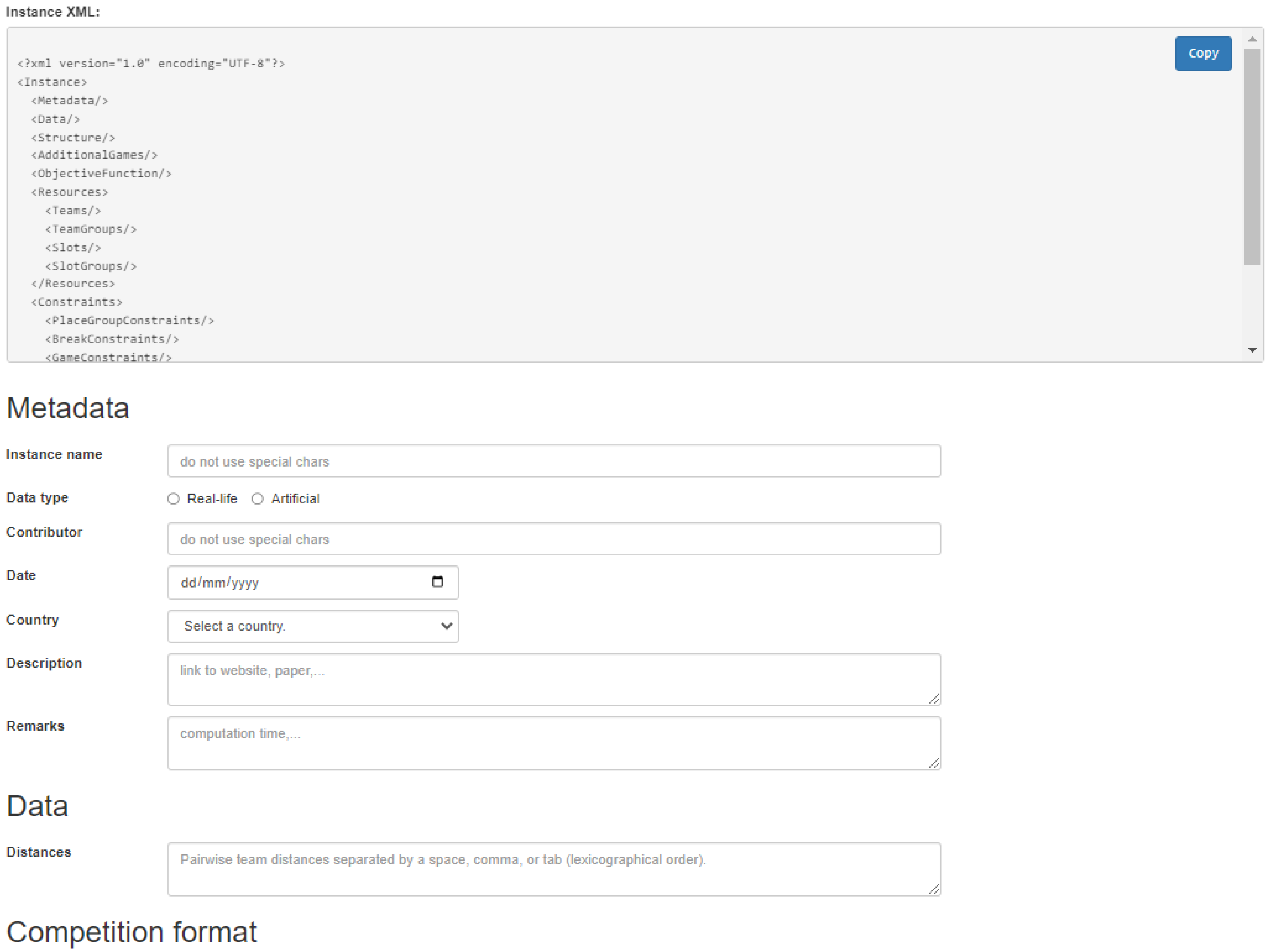
1. A classification scheme for round robin tournament timetabling problems
2. **A standard problem instance data format**
3. A benchmark instance set
4. General sport scheduling solvers
5. Algorithm selection & insights

2. A STANDARD DATA FORMAT

RobinX: an XML-driven classification for round-robin sports timetabling



- User-friendly web-application



- Open-source C++ library to read, write, validate XML files

www.sportscheduling.ugent.be/RobinX

GHENT UNIVERSITY

2. A STANDARD DATA FORMAT

Problem instance repository

Soft constraints objective repository

Specification | XML files | Downloads

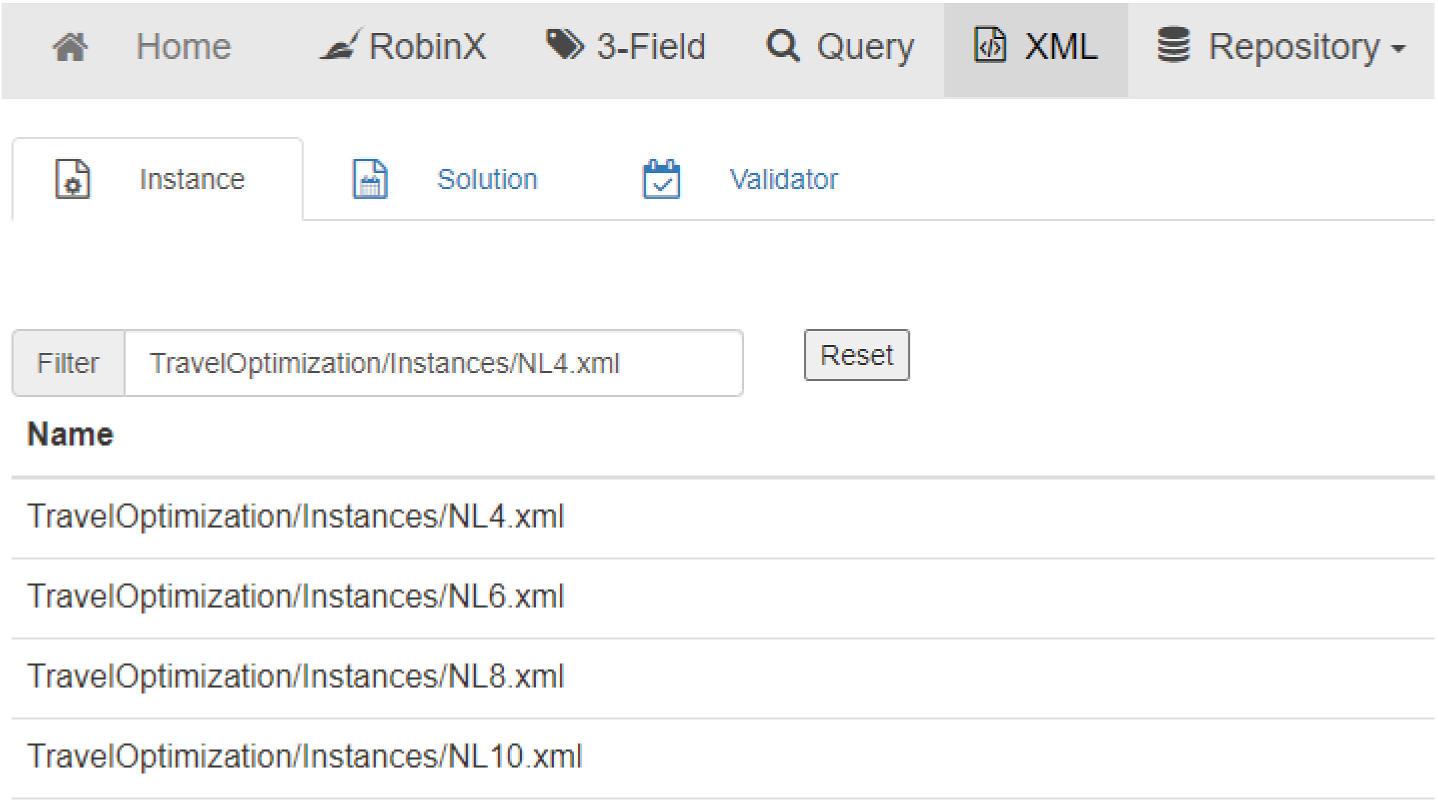
Instance name	Contributor	Teams	Slots	Classification	Best LB	Best UB	History
AUS1	Bartsch, Drexl, Kroger	10	18	2RR, C, | BR2, CA1, CA4, GA1, SE1 | SC	(/, /)	(0, 89)	history
GER1	Bartsch, Drexl, Kroger	18	34	2RR, C, M | BR2, CA1, CA4, GA1, SE1 | SC	(/, /)	(/, /)	history
GER2	Bartsch, Drexl, Kroger	18	34	2RR, C, M | BR1, BR2, CA1, CA4, GA1 | SC	(/, /)	(/, /)	history
BEL1	Goossens, Spieksma	18	34	2RR, C, M | BR1, BR2, CA1, CA2, CA3, CA4, GA1 | SC	(/, /)	(0, 122)	history
BEL2	Goossens, Spieksma	18	34	2RR, C, M | BR1, BR2, CA1, CA2, CA3, CA4, GA1 | SC	(/, /)	(0, 102)	history
BEL3	Goossens, Spieksma	18	34	2RR, C, M | BR1, BR2, CA1, CA2, CA3, CA4, GA1 | SC	(/, /)	(0, 150)	history
CHI1	Duran, Guajardo, Miranda, Saure, Souyris, Weintraub, Wolf	20	19	1RR, C, NULL | BR1, CA1, CA3, CA4, GA1 | SC	(/, /)	(/, /)	history
DanishFootball	Rasmussen	12	33	3RR, C, P | BR1, BR2, CA1, CA2, CA3, SE1 | SC	(/, /)	(0, 49)	history
FIN1	Kyngas, Nurmi	14	30	2RR, R, NULL | BR1, BR2, CA1, CA3, CA4, FA1, FA2, FA3, GA1, SE1 | SC	(/, /)	(/, /)	history
FIN2	Kyngas, Nurmi	12	22	2RR, C, NULL | BR1, BR2, CA1, CA3, CA4, FA1, GA1, SE1 | SC	(/, /)	(0, 87)	history
FootballEcuador_2012	Recalde, Torres, Vaca	12	22	2RR, C, I | BR1, CA2, CA3, CA4 | SC	(/, /)	(0, 12)	history
FootballSouthAmerica	Duran, Guajardo, Saure	10	18	2RR, C, F | BR1, CA1, CA3 | SC	(/, /)	(0, 0)	history
NorwegianFootball_1_2009	Hausken, Andersson, Fagerholt, Flatberg	6	10	2RR, C, P | BR1, BR2, CA1, CA3, CA4, GA1, SE1 | SC	(/, /)	(0, 14)	history

www.sportscheduling.ugent.be/RobinX

2. A STANDARD DATA FORMAT

Online validator

1. Select (or generate) your instance

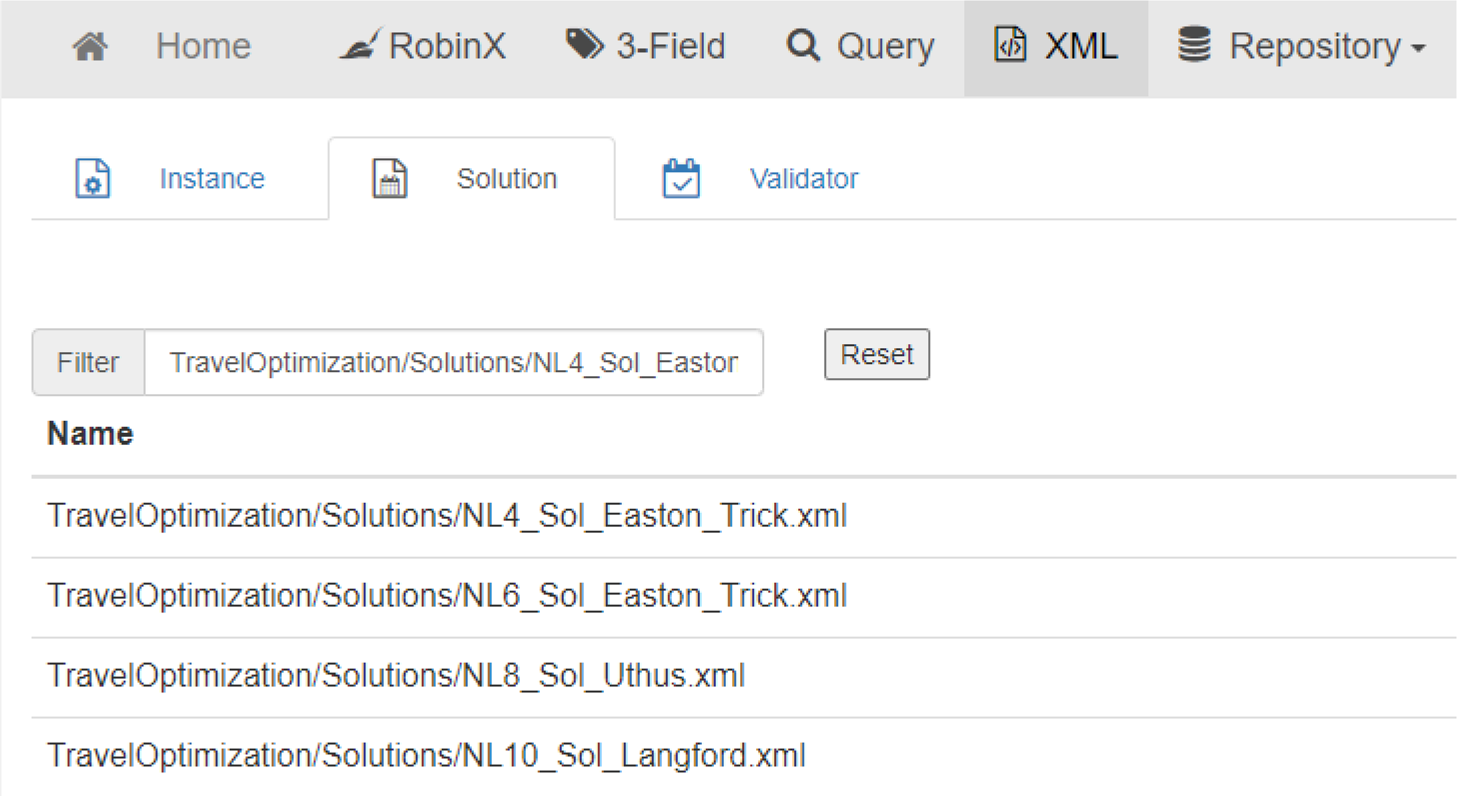


GHENT UNIVERSITY

www.sportscheduling.ugent.be/RobinX/validator.php

2. A STANDARD DATA FORMAT

Online validator



1. Select (or generate) your instance
2. Generate / upload your solution

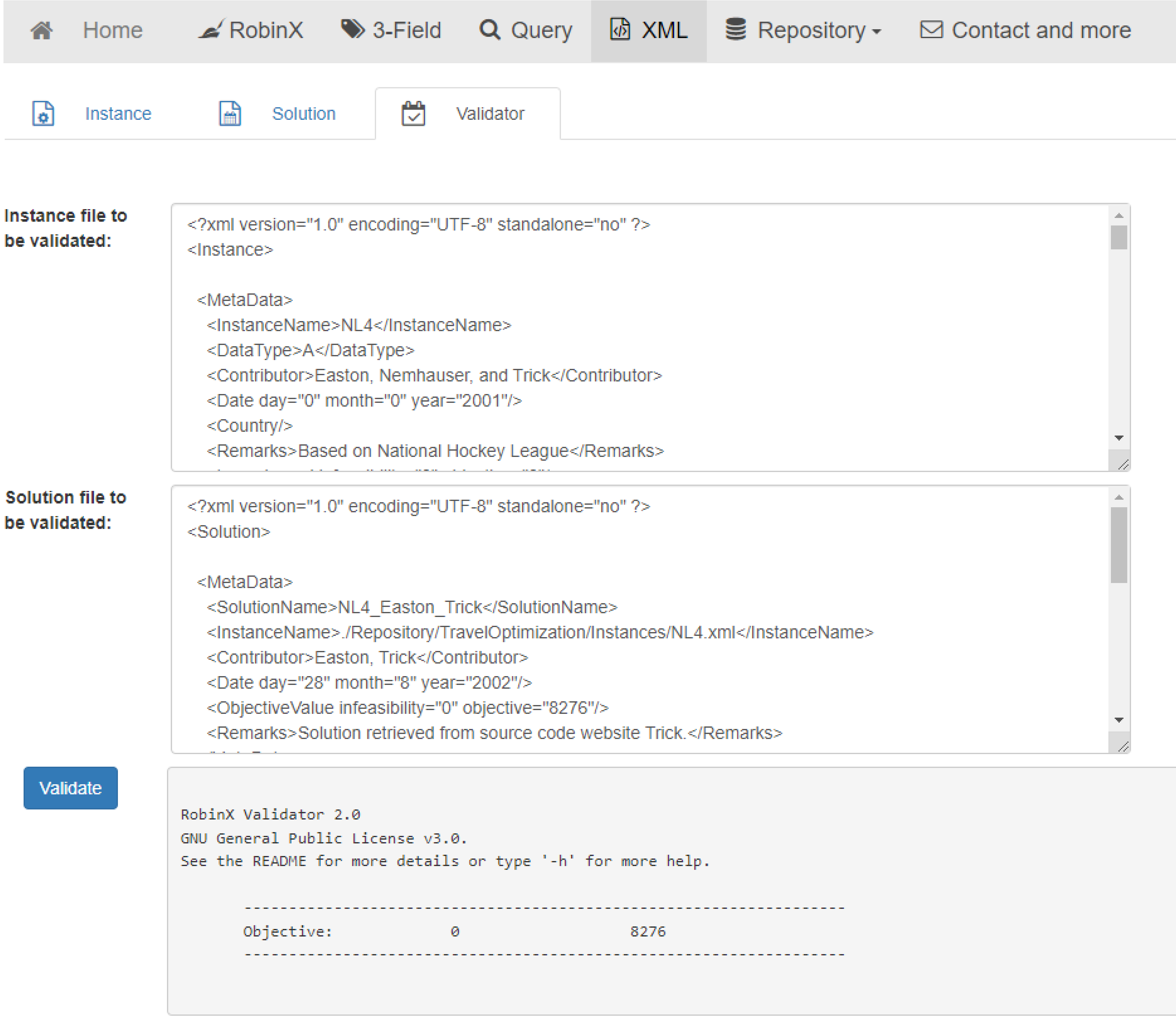
GHENT UNIVERSITY

www.sportscheduling.ugent.be/RobinX/validator.php

2. A STANDARD DATA FORMAT

Online validator

1. Select (or generate) your instance
2. Generate / upload your solution
3. Press validate



GHENT UNIVERSITY

www.sportscheduling.ugent.be/RobinX/validator.php

FROM CONSULTING TO SCIENCE

1. A classification scheme for round robin tournament timetabling problems
2. A standard problem instance data format
3. **A benchmark instance set**
4. General sport scheduling solvers
5. Algorithm selection & insights

3. A BENCHMARK INSTANCE SET

Goal: develop a diverse set of challenging and realistic problem instances

Setting the scope

Tournament structure (α)

2RR

Time-constrained

Phased, or no symmetry

Constraints (β)

Capacity constraints (CA1-4)

Break constraints (BR1-2)

Fairness constraints (FA2)

Game constraints (GA1)

Separation constraints (SE1)

Objective function (γ)

Minimize sum of soft constraint penalties (SC)

[satisfy all hard constraints]

Starting point

<table>
<tr><th>Name</th><th>Contributor</th><th>No.</th><th>Teams</th><th>Description</th></tr>
<tr><td>BEL</td><td>Goossens & Spieksma (2009)</td><td>3</td><td>18</td><td>2RR, C, P | BR1^{H}, BR2^{H}, CA1^{H,S}, CA2^{S}, CA3^{H,S}, CA4^{H,S}, GA1^{S}, SE1^{S}| SC</td></tr>
<tr><td>PRIN</td><td>Lewis & Thompson (2011)</td><td>10</td><td>12-18</td><td>2RR, C, Ø | CA1^{H}, CA2^{H,S}, CA3^{H,S}, CA4^{H}, SE1^{S}| SC</td></tr>
<tr><td>ECUA</td><td>Recalde, Torres, & Vaca (2013)</td><td>1</td><td>12</td><td>2RR, C, P | BR1^{H,S}, CA2^{H}, CA3^{H,S}, CA4^{H}, SE1^{S}| SC</td></tr>
<tr><td>FIN</td><td>Kyngäs & Nurmi (2009)</td><td>1</td><td>14</td><td>2RR, C, P | BR1^{S}, BR2^{S}, CA1^{H,S}, CA3^{S}, CA4^{S}, FA2^{S}, GA1^{H}, SE1^{S}| SC</td></tr>
<tr><td>GER</td><td>Bartsch, Drexl, & Kröger (2006)</td><td>3</td><td>18</td><td>2RR, C, P | BR1^{H}, BR2^{H}, CA1^{H,S}, CA4^{H}, GA1^{H}, SE1^{S}| SC</td></tr>
<tr><td>ART</td><td>Nurmi et al. (2010)</td><td>16</td><td>10-16</td><td>2RR, C, {P,Ø} | BR1^{H}, BR2^{S}, CA1^{H,S}, CA3^{S}, CA4^{H,S}, GA1^{S}, SE1^{S}| SC</td></tr>
<tr><td>SOUTHA</td><td>Durán, Guajardo, & Sauré (2017)</td><td>1</td><td>10</td><td>2RR, C, P | BR1^{S}, CA1^{H}, CA3^{H}, SE1^{S}| SC</td></tr>
<tr><td>ITA</td><td>Cocchi et al. (2018)</td><td>1</td><td>14</td><td>2RR, C, P | BR1^{H}, BR2^{S}, CA1^{H,S}, CA2^{H,S}, CA3^{H}, CA4^{H}, FA2^{S}, GA1^{S}, SE1^{S}| SC</td></tr>
<tr><td>RRT</td><td>Horbach, Bartsch, & Briskorn (2012)</td><td>33</td><td>10-22</td><td>2RR, C, P | BR1^{H}, CA1^{S}, CA4^{H}, GA1^{S}, SE1^{S}| SC</td></tr>
<tr><td>NOR</td><td>Hausken, Andersson, Fagerholt, & Flatberg (2012)</td><td>8</td><td>14-16</td><td>2RR, C, P | BR1^{H}, BR2^{S}, CA1^{S}, CA3^{S}, CA4^{H,S}, GA1^{H,S}, SE1^{S}| SC</td></tr>
</table>

GHENT UNIVERSITY

3. A BENCHMARK INSTANCE SET

Goal: develop a diverse set of challenging and realistic problem instances

Setting the scope

Tournament structure (α)

2RR

Time-constrained

Phased, or no symmetry

Constraints (β)

Capacity constraints (CA1-4)

Break constraints (BR1-2)

Fairness constraints (FA2)

Game constraints (GA1)

Separation constraints (SE1)

Objective function (γ)

Minimize sum of soft constraint penalties (SC)

[satisfy all hard constraints]

How to characterize problem instances?

Using problem "features" (= measurable properties of a problem instance)

- Number of teams
- Symmetry (phased or not)
- For each hard constraint type: number
- For each soft constraint type: number

GHENT UNIVERSITY

3. A BENCHMARK INSTANCE SET

Goal: develop a diverse set of challenging and realistic problem instances

Setting the scope

Tournament structure (α)

2RR

Time-constrained

Phased, or no symmetry

Constraints (β)

Capacity constraints (CA1-4)

Break constraints (BR1-2)

Fairness constraints (FA2)

Game constraints (GA1)

Separation constraints (SE1)

Objective function (γ)

Minimize sum of soft constraint penalties (SC) [satisfy all hard constraints]

How to characterize problem instances?

	$f_{\|T\|}$	f_P	f^H_{CA1}	f^S_{CA1}	f^H_{CA2}	f^S_{CA2}	f^H_{CA3}	f^S_{CA3}	f^H_{CA4}	f^S_{CA4}	f^H_{GA1}	f^S_{GA1}	f^H_{BR1}	f^S_{BR1}	f^H_{BR2}	f^S_{BR2}	f^S_{FA2}	f^S_{SE1}
min	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10% decile	10	0	5	15	6	12	1	12	18	18	2	1	12	10	1	1	1	1
mean	14.83	0.68	5.56	16.35	3.90	27.92	0.27	20.36	32.06	19.01	1.78	64.99	26.14	0.60	0.08	0.31	0.03	0.94
90% decile	20	1	42	32	72	620	2	112	85	340	34	126	44	24	1	1	1	1
max	22	1	116	60	112	620	2	126	198	374	34	4368	602	24	1	1	1	1

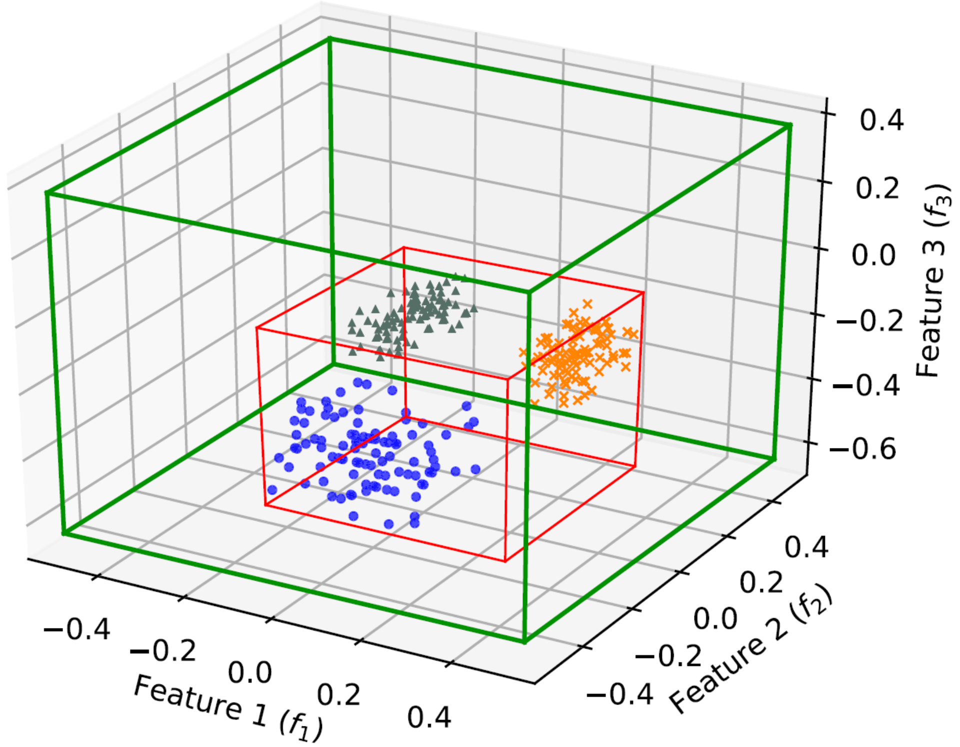
3. A BENCHMARK INSTANCE SET

Goal: develop a diverse set of challenging and realistic problem instances

How to visualize set of problem instances?

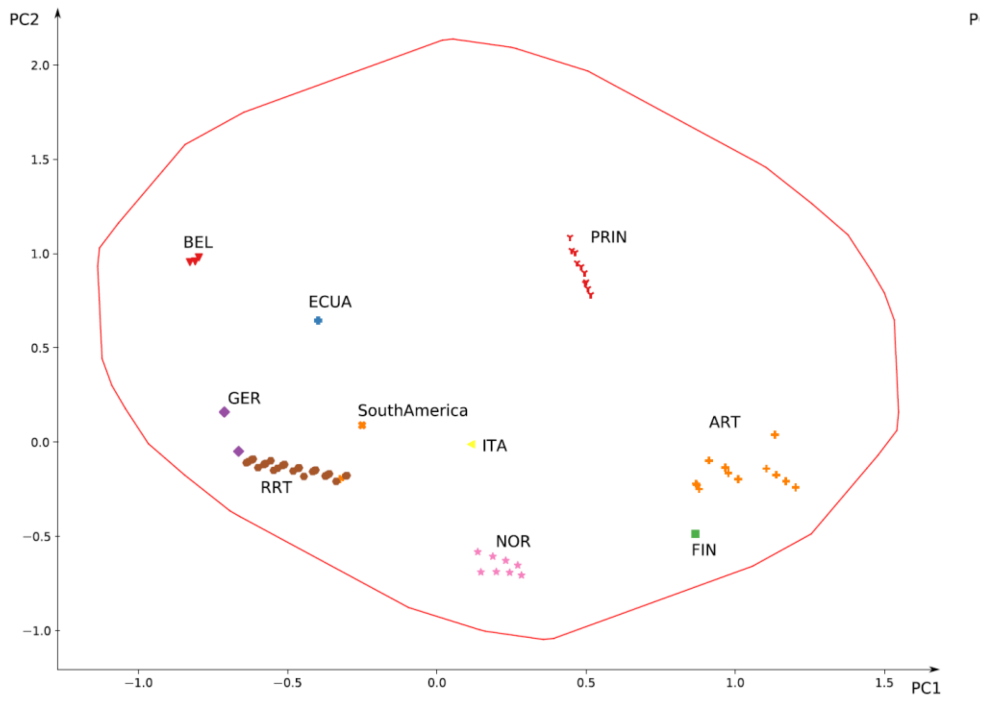
$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{18} \end{bmatrix}$$

Feature vector



Feature space (18D)

PCA



2D space

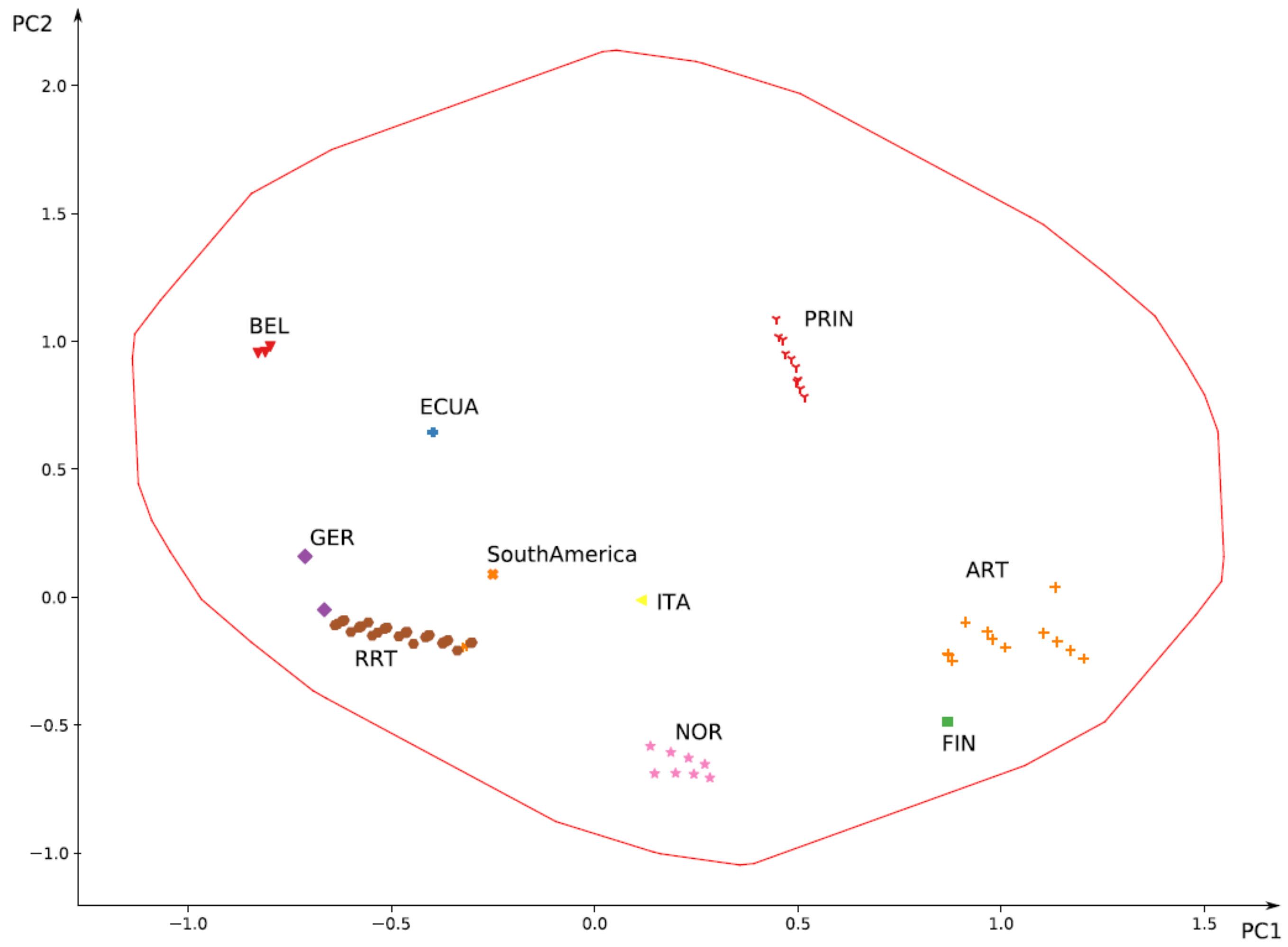
See e.g. Smith-Miles et al. (C&OR, 2014)

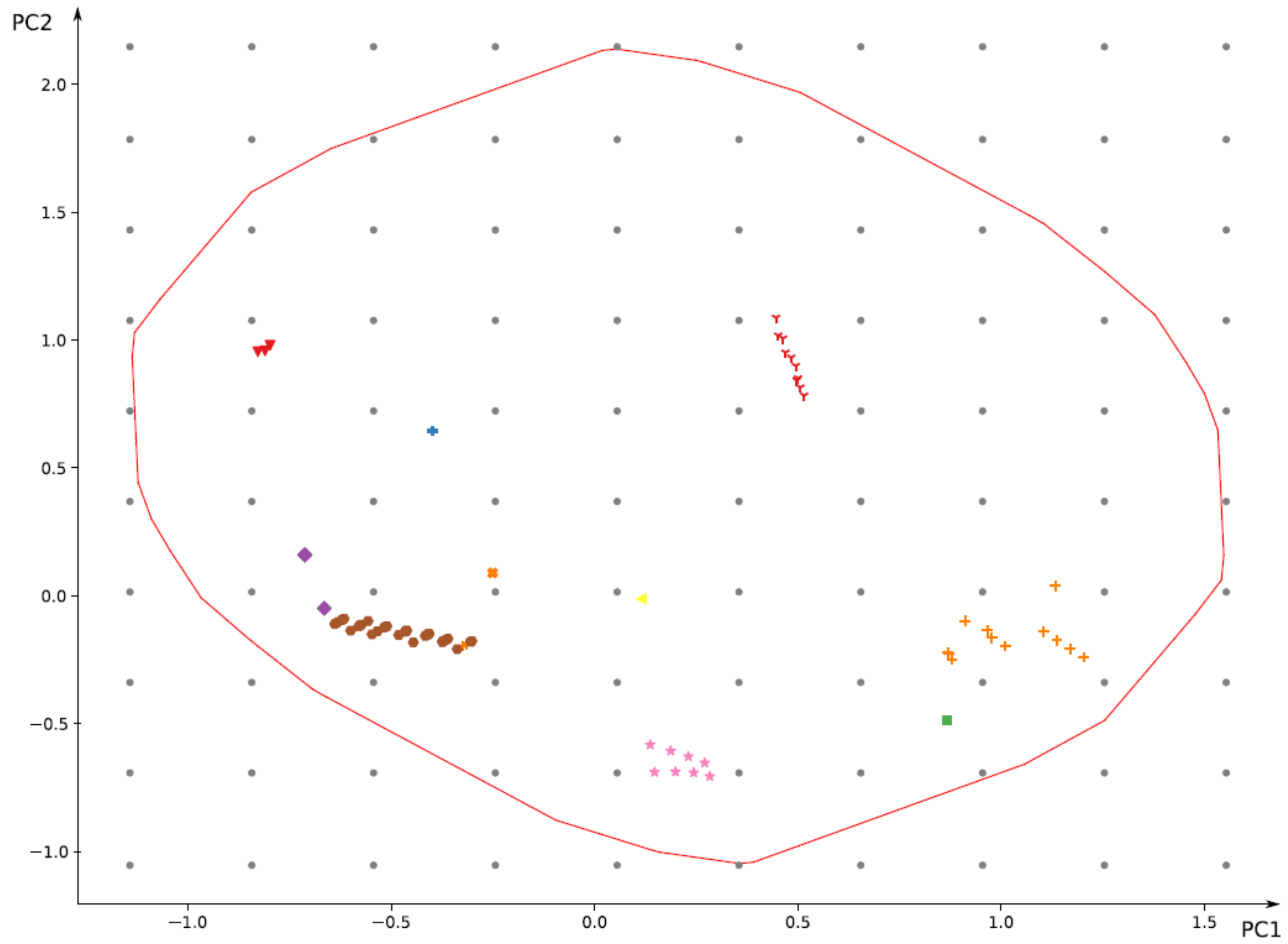
GHENT UNIVERSITY

3. A BENCHMARK INSTANCE SET

Goal: develop a diverse set of challenging and realistic problem instances

These instances may be challenging and realistic, but not diverse.





So let’s try to fill the gaps: target instances (within the red “convex hull” of realistic instances)

GHENT UNIVERSITY

3. A BENCHMARK INSTANCE SET

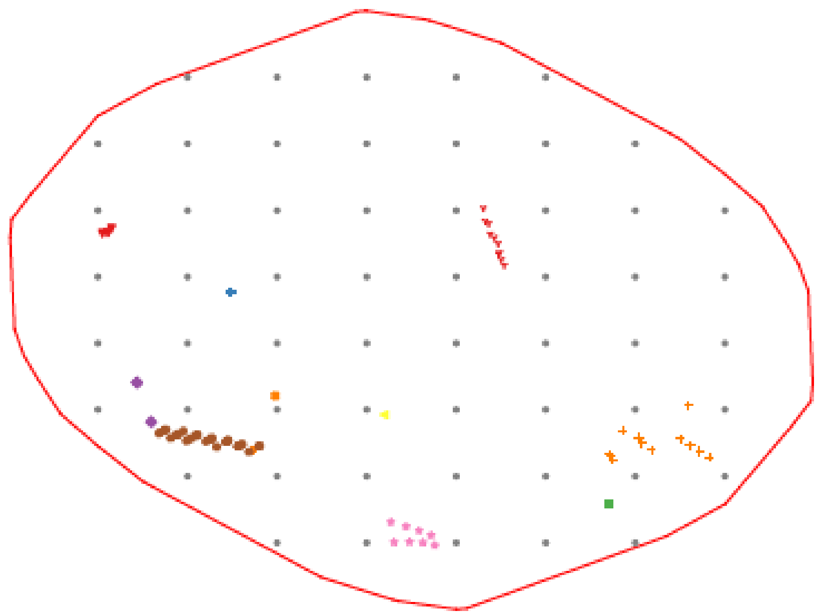
Goal: develop a diverse set of challenging and realistic problem instances

Real-life instances

Feature extraction with RobinX ↓

Feature vector
2RR, C, Ø | CA4=10, SE1=2 | TR 2RR, C, M | BR1=20 | CO 2RR, C, Ø | CA1=10, CA3=27 | BR

→ PCA



→ Undo PCA with IP

Feature vector
1RR, C, Ø | CA1=10, CA2=2 | TR 2RR, C, M | BR1=5, SE1=20 | CO 4RR, C, Ø | CA4=10, SE1=3 | BR

Instance generator ↑

Applicable synthetic instances

Parameters

F = set of all features

$w_{i,x}$ = PCA-coefficient of feature i for dimension x

x, y = target coordinates in 2D-space

Variables

f_i = value of feature i

g_i = value of normalized feature i

$s_{i,x}, e_{i,x}$ = slack (excess) on x-projection of feature i

minimize $M \cdot \sum_{i \in F} (s_{i,x} + e_{i,x} + s_{i,y} + e_{i,y})$

$$\sum_{i \in F} (g_i w_{i,x}) + s_{i,x} + e_{i,x} = x$$

$$\sum_{i \in F} (g_i w_{i,y}) + s_{i,y} + e_{i,y} = y$$

$$(\max_i - \min_i) g_i - f_i = -\min_i \qquad \forall i \in F$$

$$f_i \in \mathbb{N}^+, s_{i,x}, e_{i,x}, s_{i,y}, e_{i,y} \geq 0 \qquad \forall i \in F$$

GHENT UNIVERSITY

3. A BENCHMARK INSTANCE SET

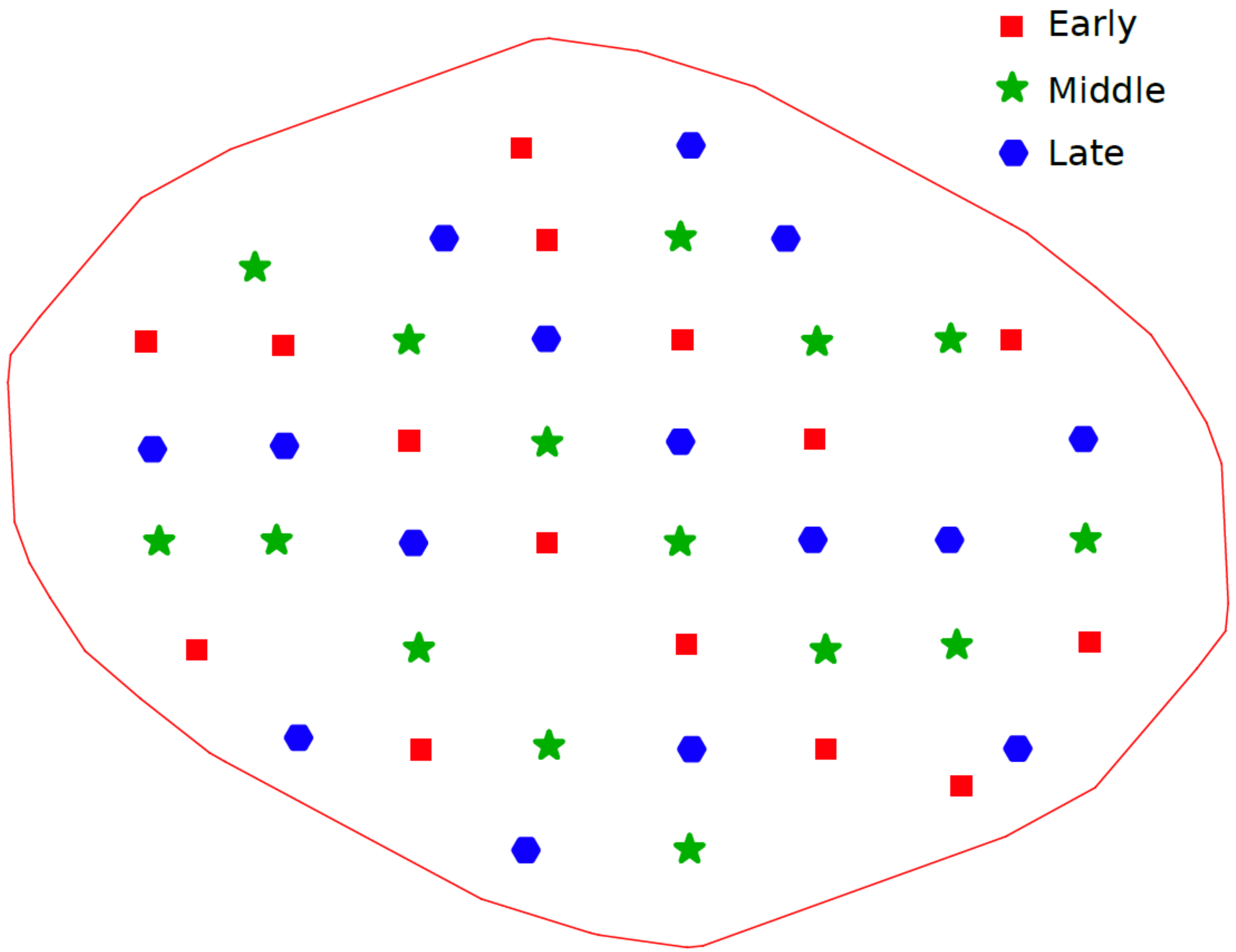
Goal: develop a diverse set of challenging and realistic problem instances

Final check: are these instances challenging?

Empirical hardness check

- Integer programming solver
- Constraint programming solver
- Fix-and-optimize matheuristic

Note: feasible solution exists by design.
Feasible solution found within 1 hour for 12 (IP), 16 (CP) and 15 (F&O) instances.
None solved with proven optimality.
Solutions with different objective values found.



GHENT UNIVERSITY

FROM CONSULTING TO SCIENCE

1. A classification scheme for round robin tournament timetabling problems
2. A standard problem instance data format
3. A benchmark instance set
4. **General sport scheduling solvers**
5. Algorithm selection & insights

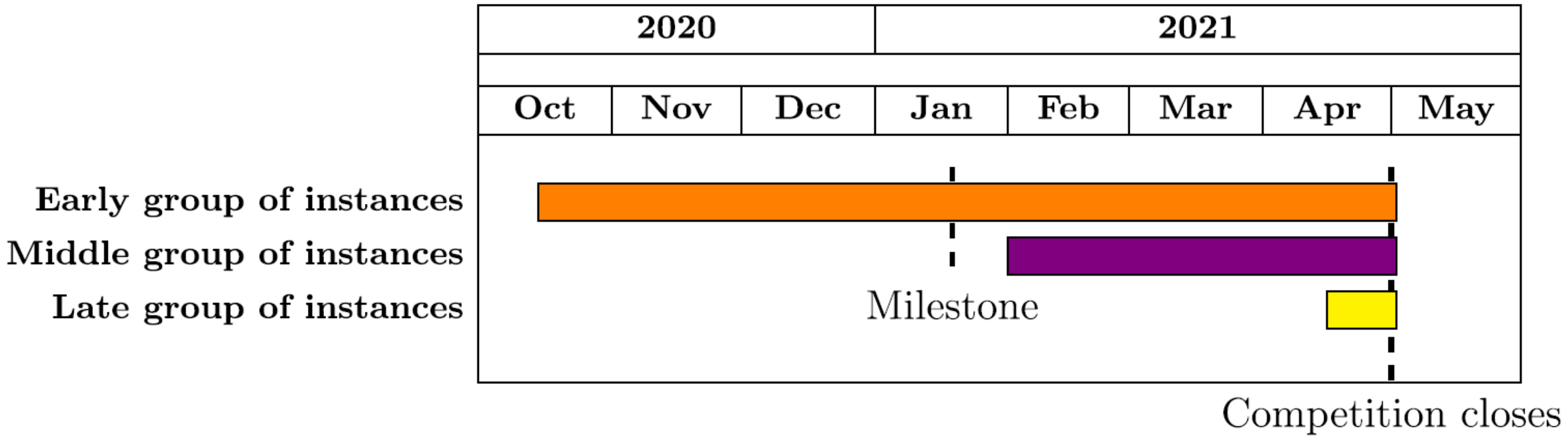
4. GENERAL SPORT SCHEDULING SOLVERS

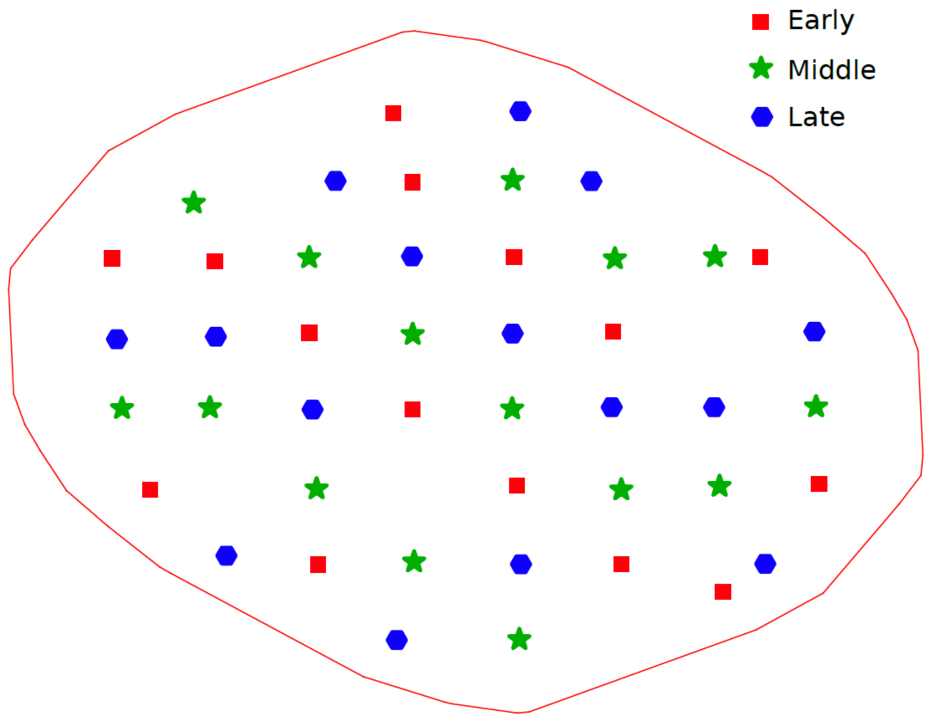
How to develop general sport scheduling solvers?

International Timetabling Competition (ITC2021)

History of ITC competitions

1st ITC 2002	University course timetabling
2nd ITC 2007	Examination & course timetabling
3rd ITC 2011	High-school timetabling
4th ITC 2019	University timetabling
5th ITC 2021	Sports timetabling





GHENT UNIVERSITY

4. GENERAL SPORT SCHEDULING SOLVERS

Main competition rules:

- No computation time or technology restrictions (one deadline for all problem instances)
- Organizers do not run algorithm code
- The same version of the algorithm must be used for all instances
- For each instance, points are awarded according to the position of the competitor
- Ordering of participants is based on weighted sum of points for all early, middle and late instances

	Instance		
Position	Early	Middle	Late
1st	10	15	25
2nd	7	11	18
3rd	5	8	15
4th	3	6	12
5th	2	4	10
6th	1	3	8
7th		2	6
8th		1	4
9th			2
10th			1

4. GENERAL SPORT SCHEDULING SOLVERS

Team	Early	Middle	Late	Total	Feas. sol	Best sol.
1. UoS	**121**	**178**	**297**	**596**	**45**	**21**
2. Udine	**75**	**114**	**235**	**424**	**44**	**4**
3. Saturn	**64**	**115**	**207**	**386**	**37**	**16**
4. GOAL	**38**	**72**	**133**	**243**	**37**	**4**
5. MODAL	**21**	**65**	**150**	**236**	**40**	**4**
6. TU/e	**41**	**47**	**136**	**224**	**38**	**2**
7. DES	8	42	72	122	37	3
8. Gionar	25	16	68	109	40	3
9. DITUoI Arta	4	29	68	101	37	2
10. NHH	5	13	70	88	40	1
11. Aures	0	1	12	13	31	1
12. UoR	0	0	10	10	29	1
13. Team zero	0	0	5	5	26	0

PATAT AUTOMATED TIMETABLING

OR IN SPORTS

GHENT UNIVERSITY

FROM CONSULTING TO SCIENCE

1. A classification scheme for round robin tournament timetabling problems
2. A standard problem instance data format
3. A benchmark instance set
4. General sport scheduling solvers
5. **Algorithm selection & insights**

5. ALGORITHM SELECTION AND INSIGHTS

Computational experiment

- 7 algorithms from ITC 2021
- 1 newly generated algorithm (FBHS)

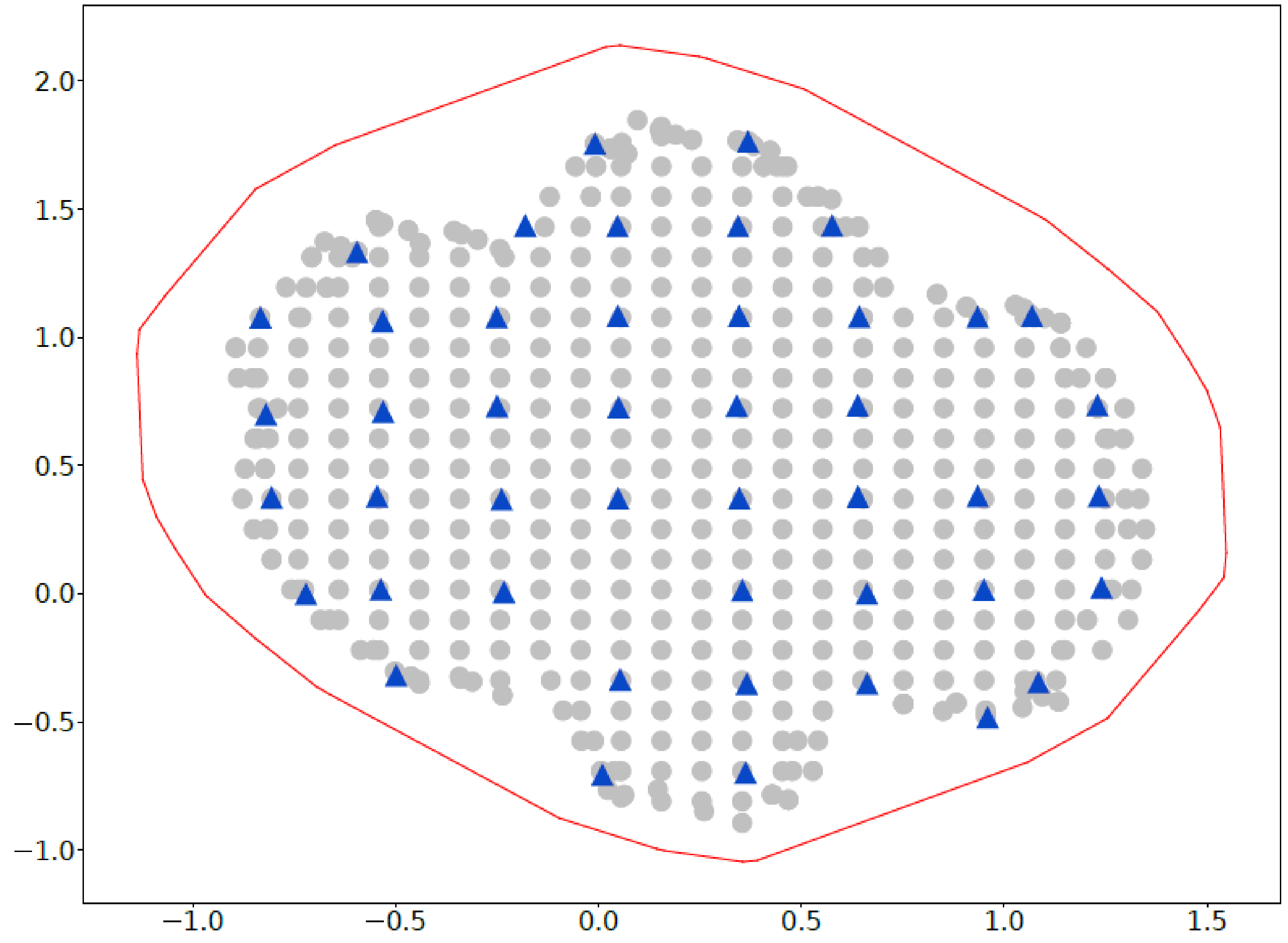
-> algorithm portfolio with a variety of approaches

Algorithm	Search method	Software details	Hardware details
MODAL	IP Branch & Cut	Python, Zimpl, C, Gurobi 10, Xpress	Per instance one thread was used and several instances where run at the same time on a machine with multiple Intel(R) Xeon(R) Gold Processors with an average clock speed of 2.4 GHz and enabled with 768 GB and 36 threads.
Goal	Fix-and-optimize matheuristic	Java 16, Gurobi 10.0	Intel Core i7-8700 3.2GHz with 12 GB RAM (single thread running for 24 hours per instance)
DES	Adaptive LNS matheuristic	Python 3.10, Gurobi 10.0	Intel Xeon 3.9GHz with 4 cores and 8 threads (Google Compute Engine "c2-standard-8") for 2.5 hours per instance
UoS	VND matheuristic	Python 3.10.4, Gurobi 9.0.2	Dual 2.0 GHz Intel Skylake with 4 or 20 cores
Udine	Simulated annealing	C++17	Intel Xeon Processor (Cascadelake) @ 2.4 GHz, 16 cores, (max one core per execution)
DITUoIArta	CP/SAT + Simulated annealing	Python 3.10, OR-Tools 9.4	6 x Intel Core i5-10505 @ 3.2 GHz with 8GB RAM (all cores activated for the solver only) for 1 hour each time for each instance (an instance may run multiple times though)
Reprobate	Pseudoboolean optimization	Perl, clasp 3.3.9, Sat4J 2.3.6, RoundingSat Git Nov 2022	Intel Core i7-8700 3.2GHz with 64 GB RAM (single core running for 2.5 hours per instance)
FBHS	IP Decomposition + matheuristic	C++, CPLEX 12.10	Intel Xeon E5-2660v3 (Haswell-EP @ 2.6 GHz) processor enabled with 8 cores

GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS

Using the same approach, we generated 518 additional problem instances



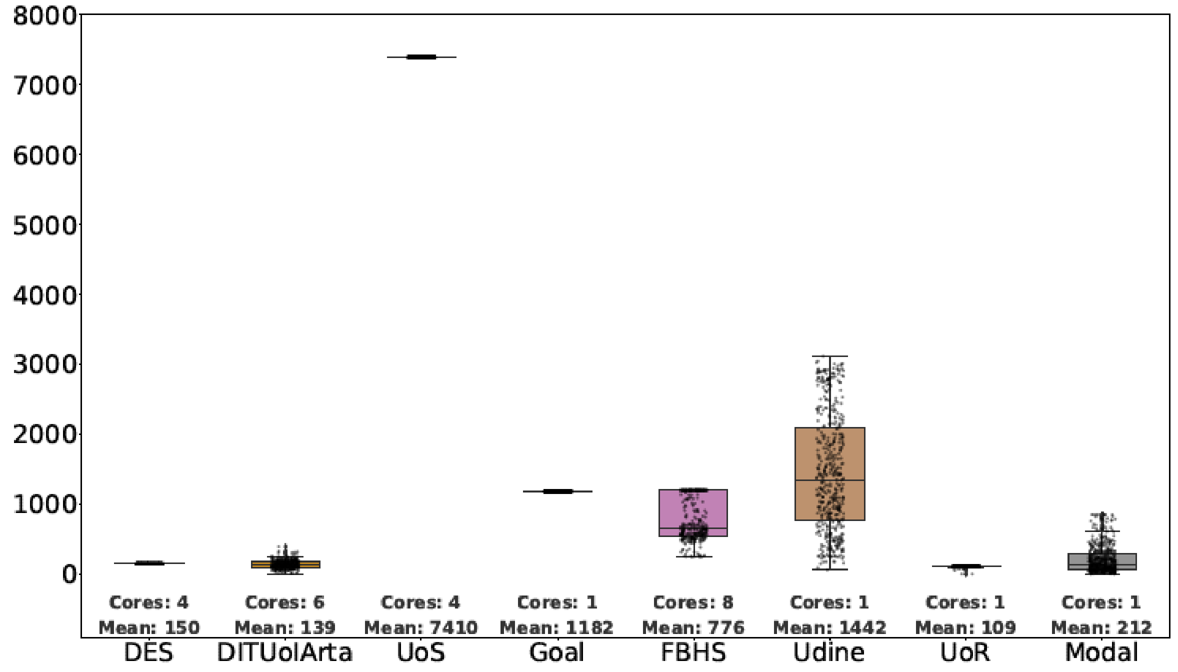
GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS

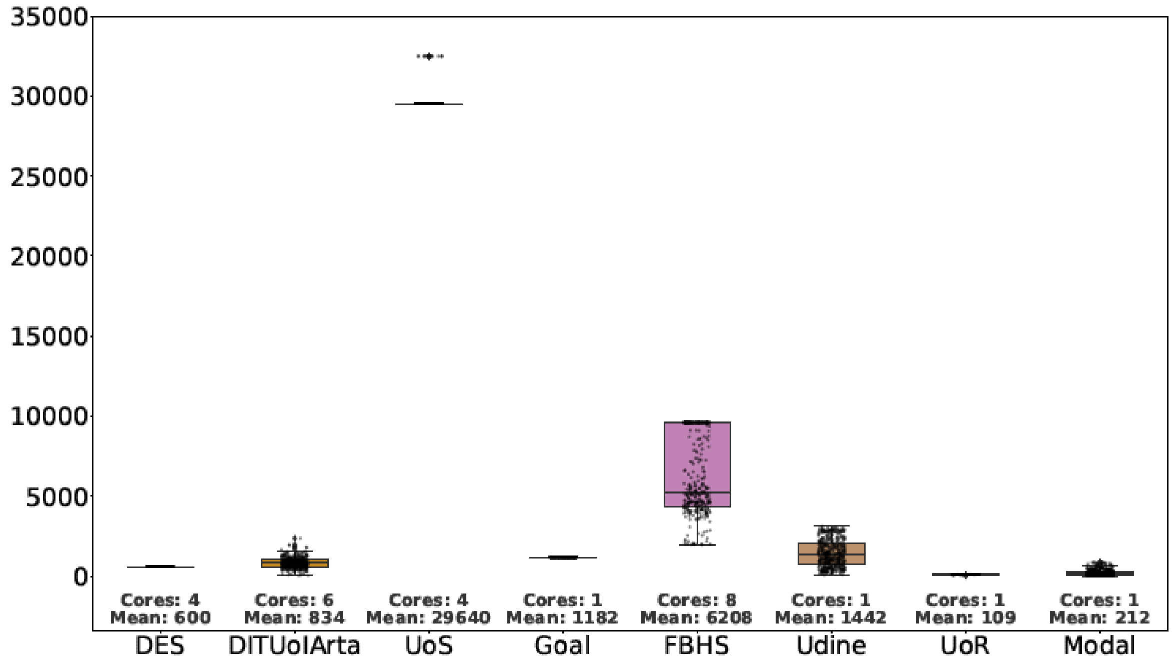
Experimental settings:

– Every algorithm is given 2 weeks time for the 518 instances

– Each algorithm is run on infrastructure available at the developing institution



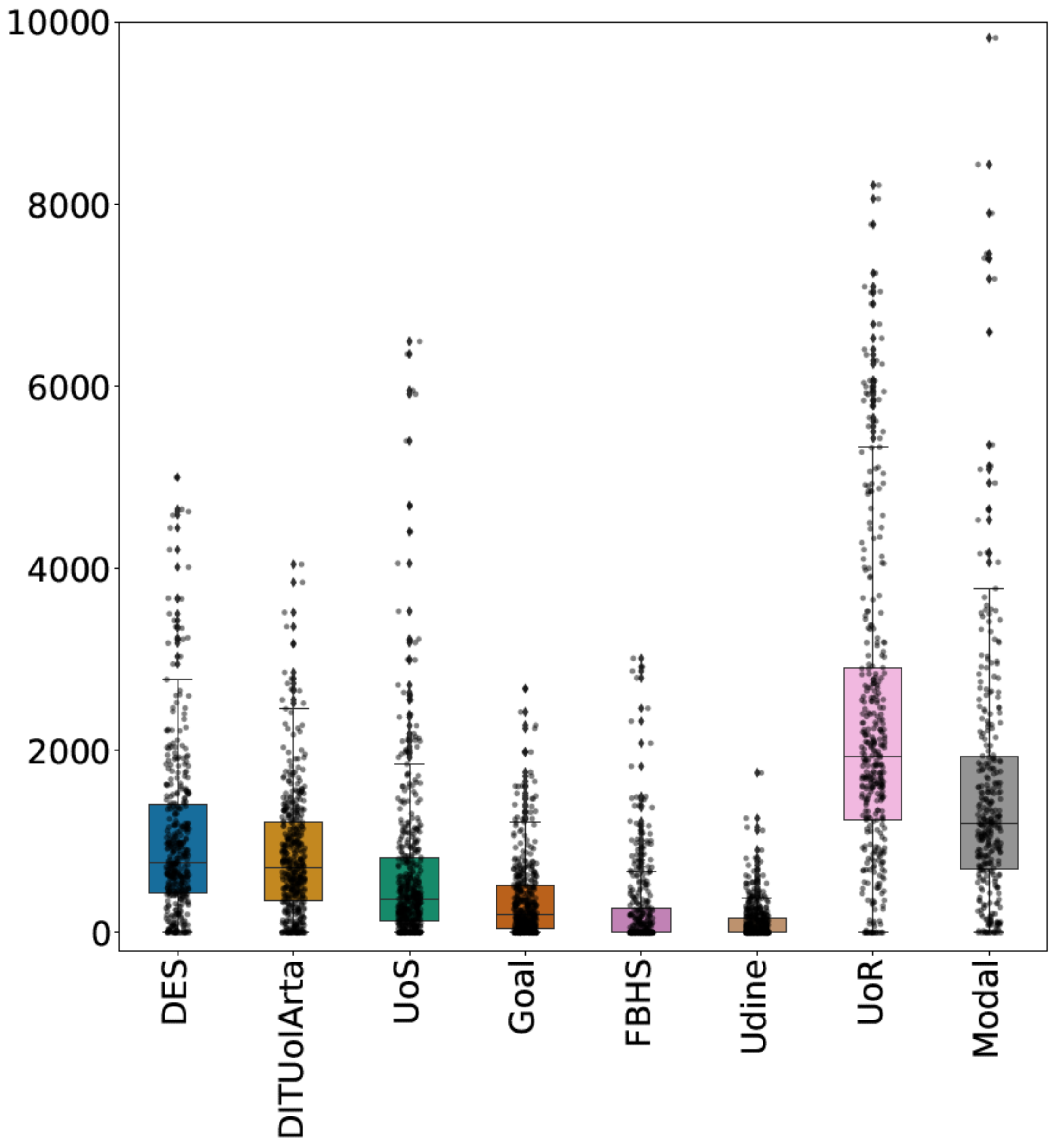
(a) Wall time (minutes)



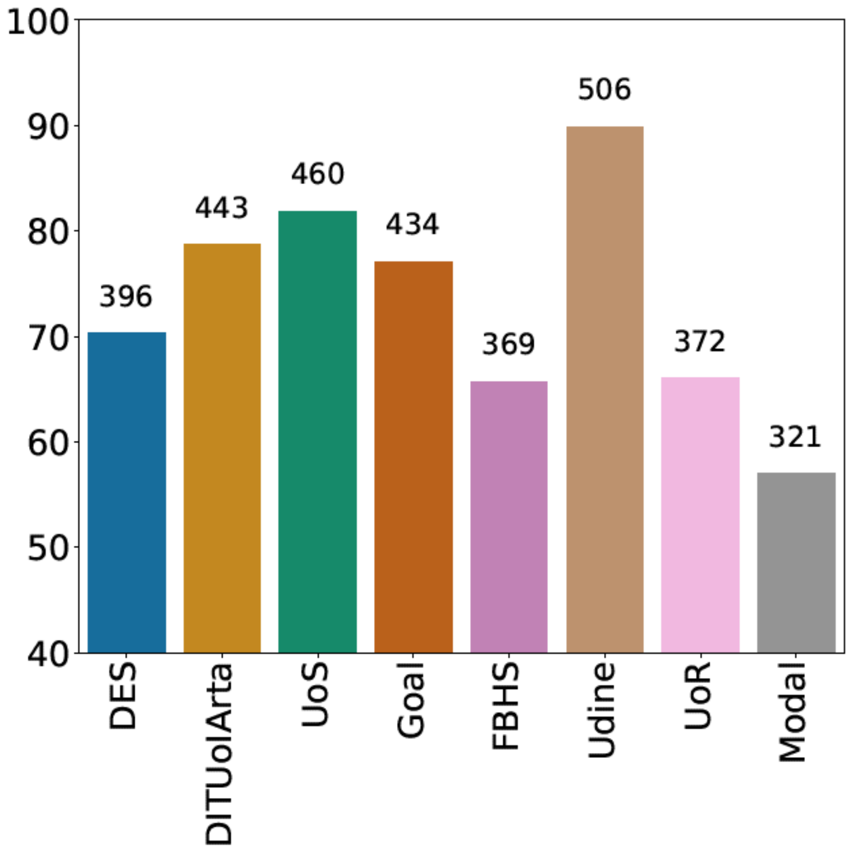
(b) CPU time (minutes)

GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS



absolute performance gap

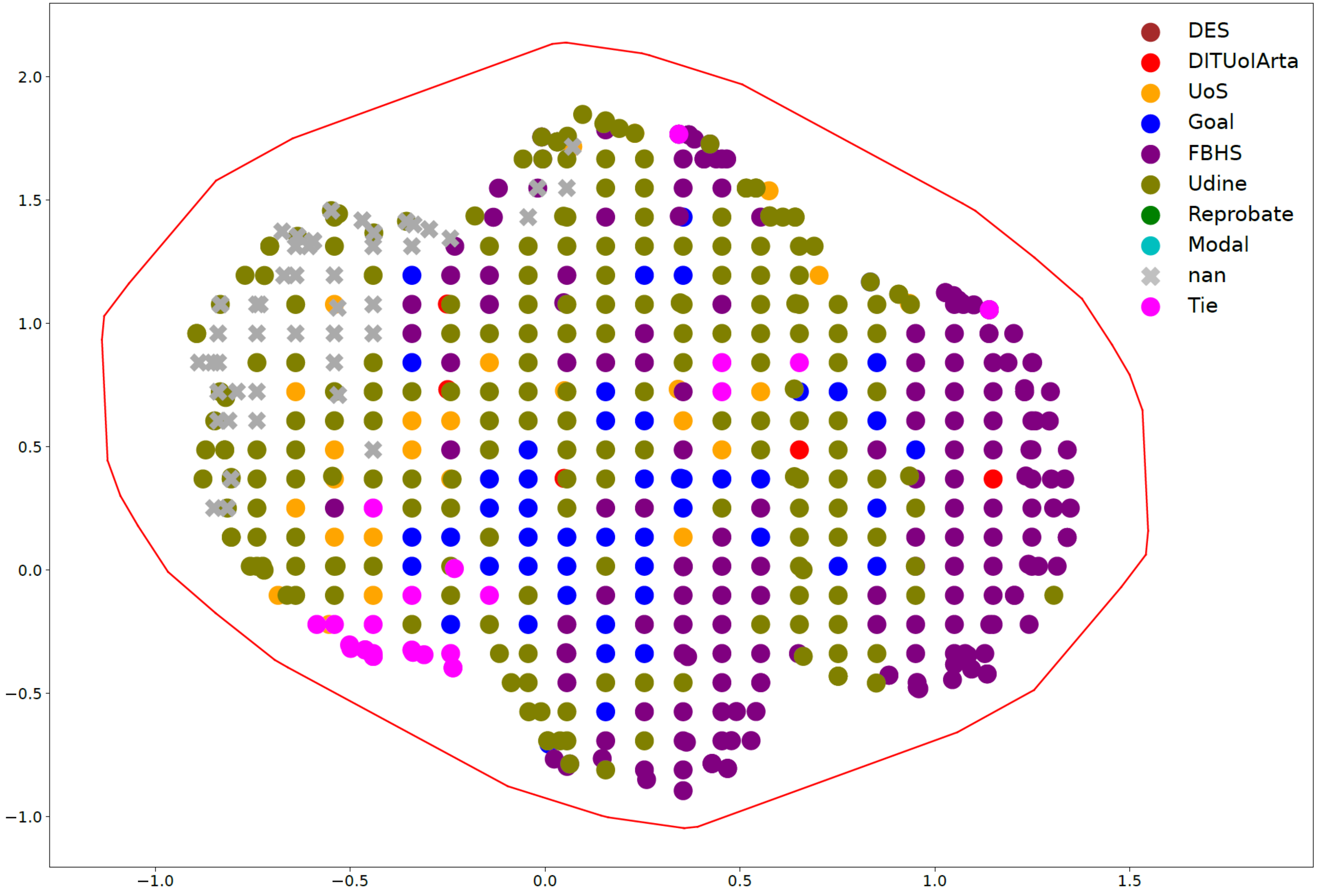


#instances for which a feasible solution was found

GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS

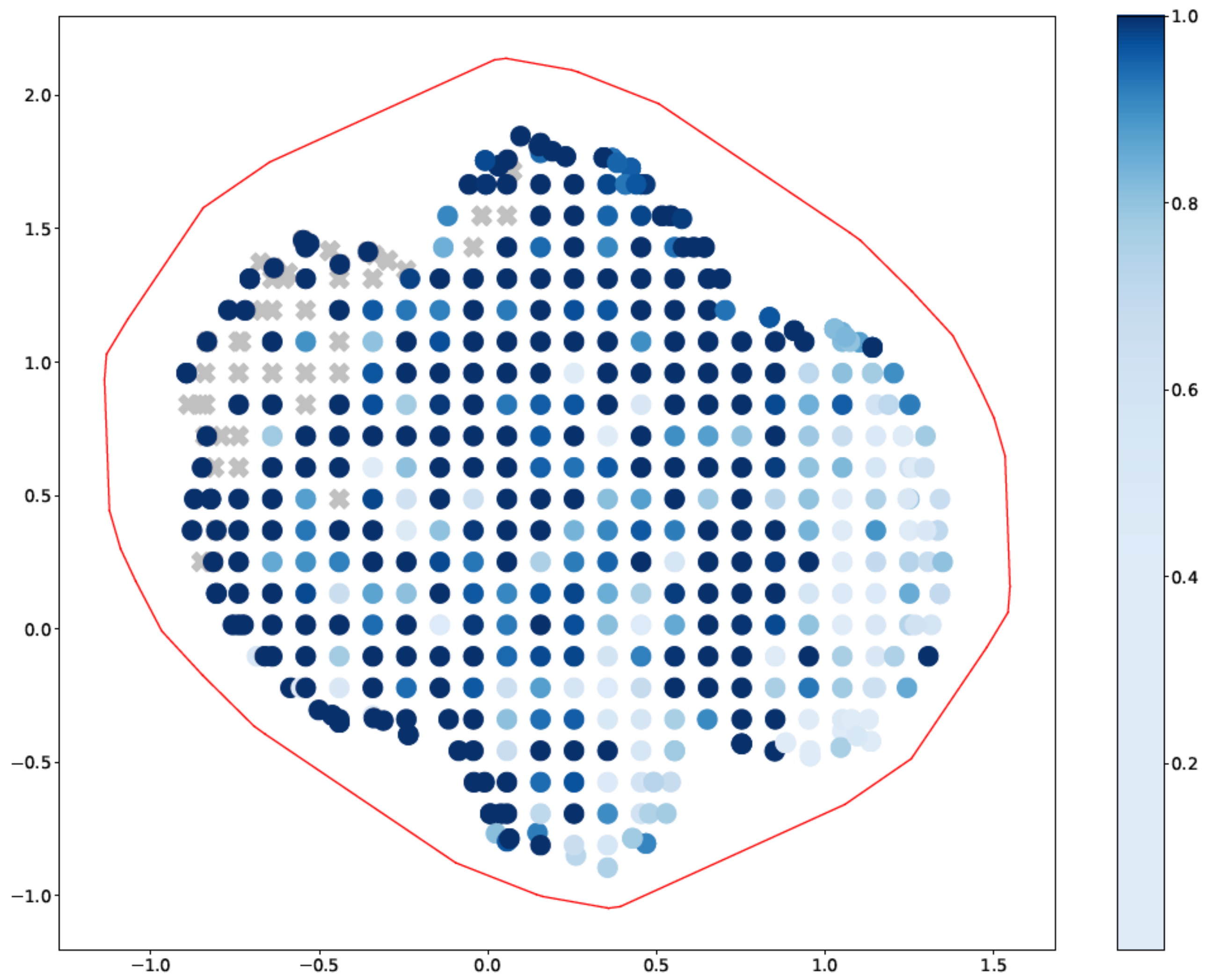
Overview of best performing algorithms

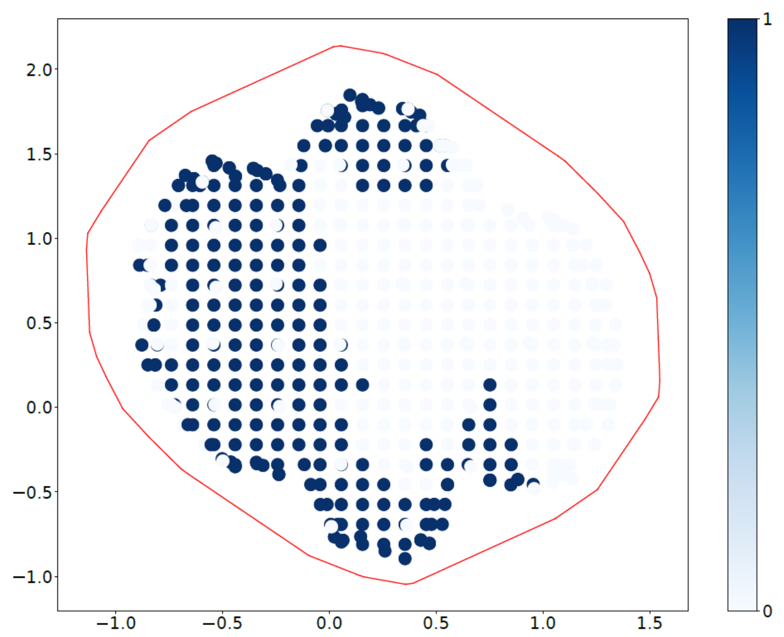


GHENT UNIVERSITY

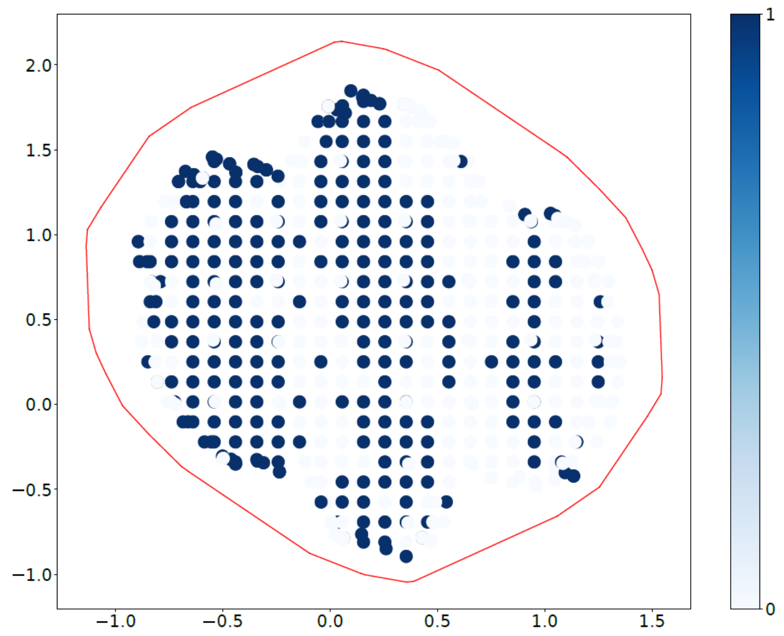
5. ALGORITHM SELECTION AND INSIGHTS

Algorithm footprint: Udine





Phased

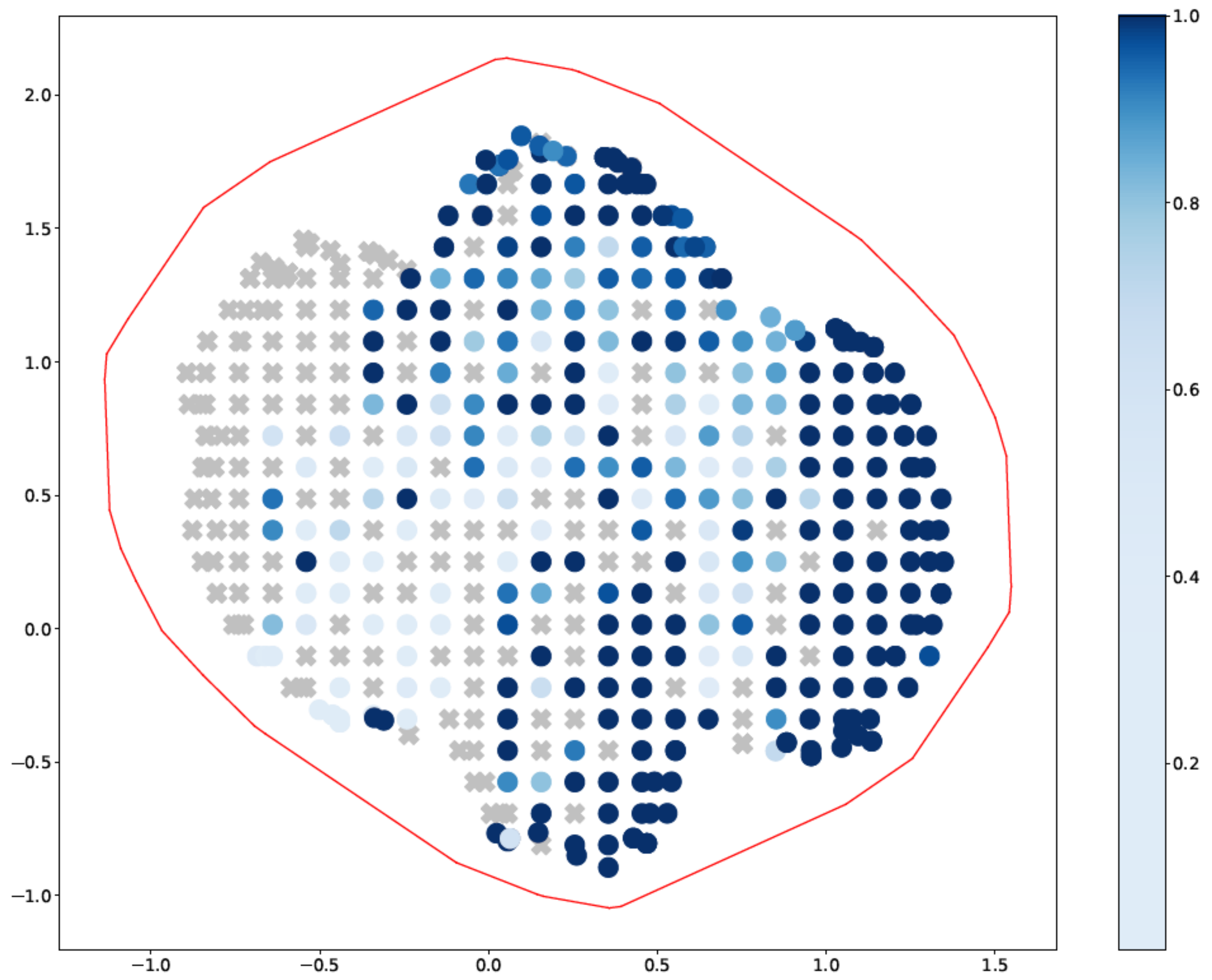


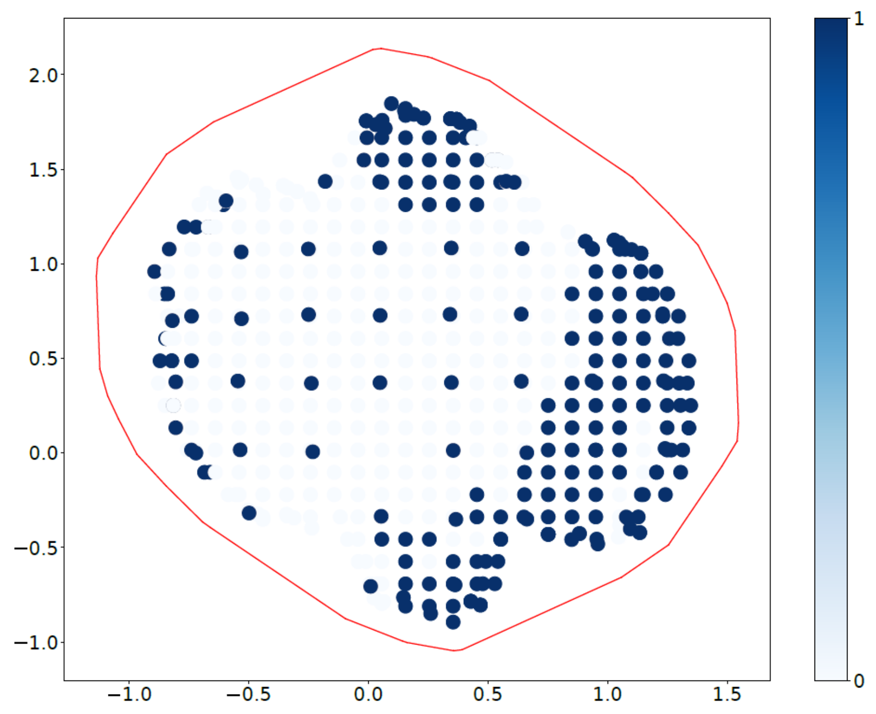
SE1 soft

GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS

Algorithm footprint: FBHS

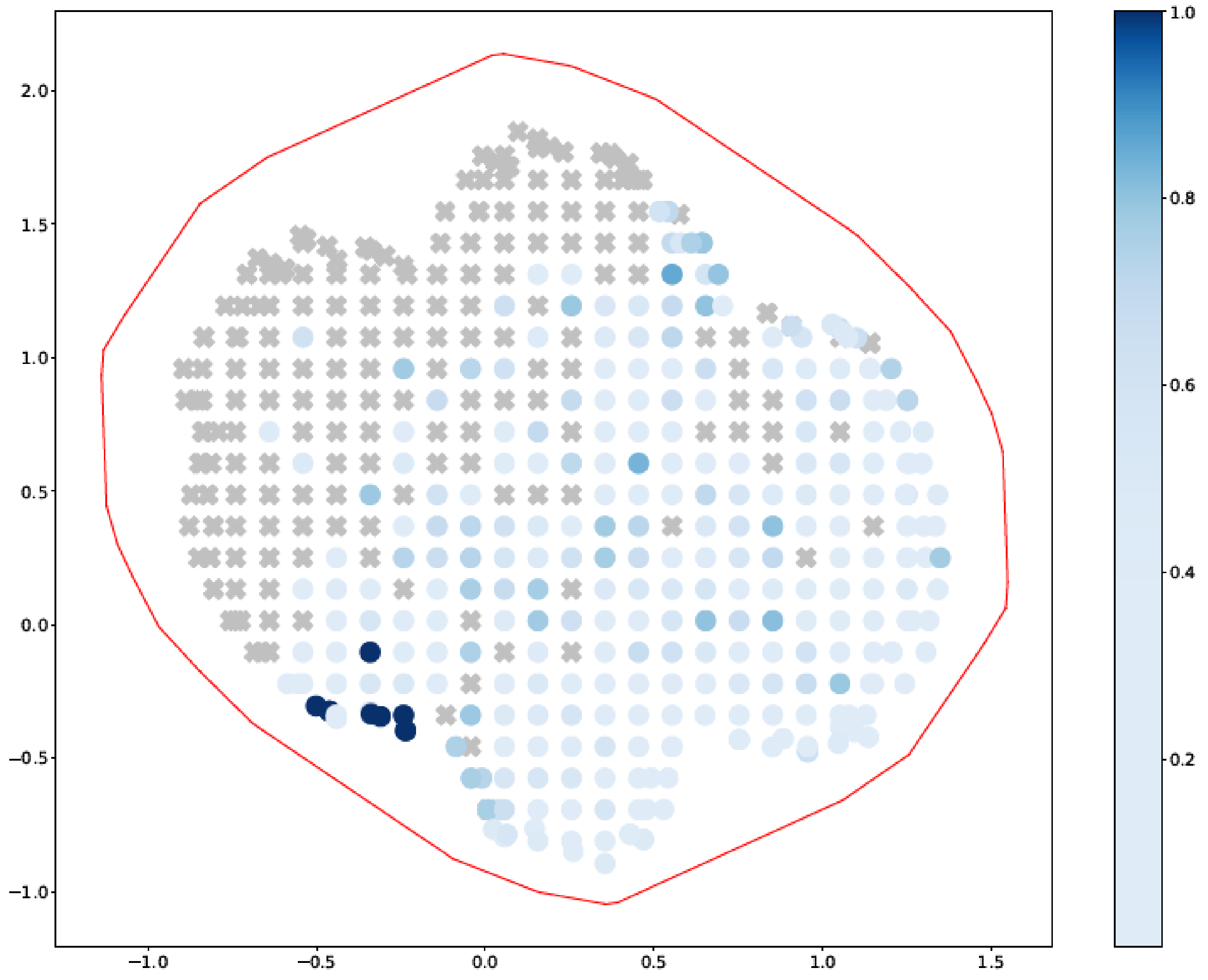


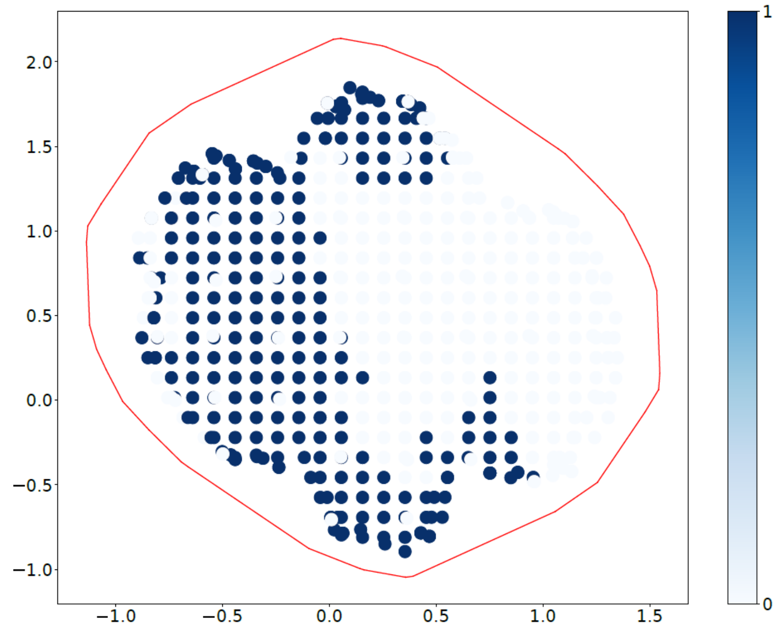


BR2 Soft

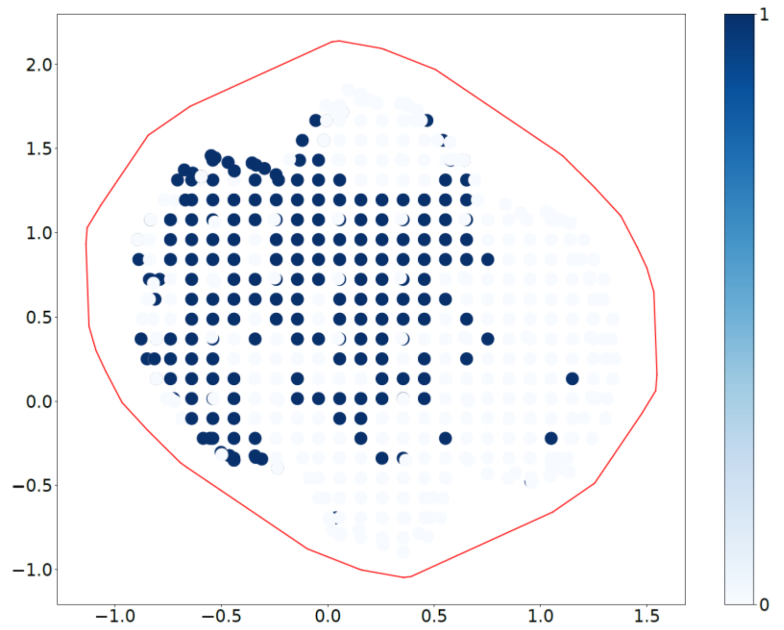
5. ALGORITHM SELECTION AND INSIGHTS

Algorithm footprint: Modal





Phased



BR2 Hard

5. ALGORITHM SELECTION AND INSIGHTS

Can we predict which algorithm will work well on which instance?

Classification models

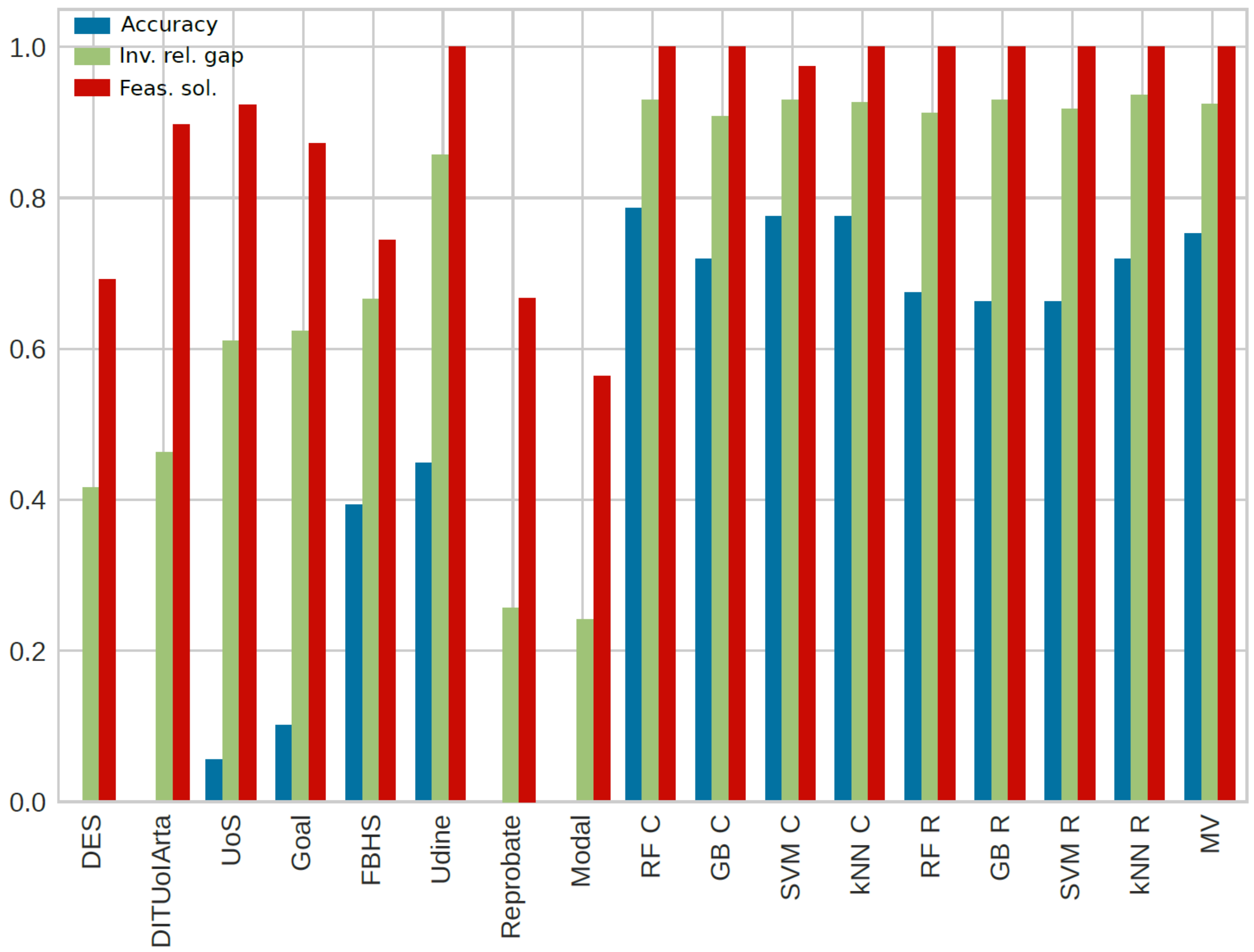
- K-nearest neighbours (kNN)
- Random forest (RF)
- Gradient-boosted trees (GB)
- Support vector machines (SVM)

Regression variants

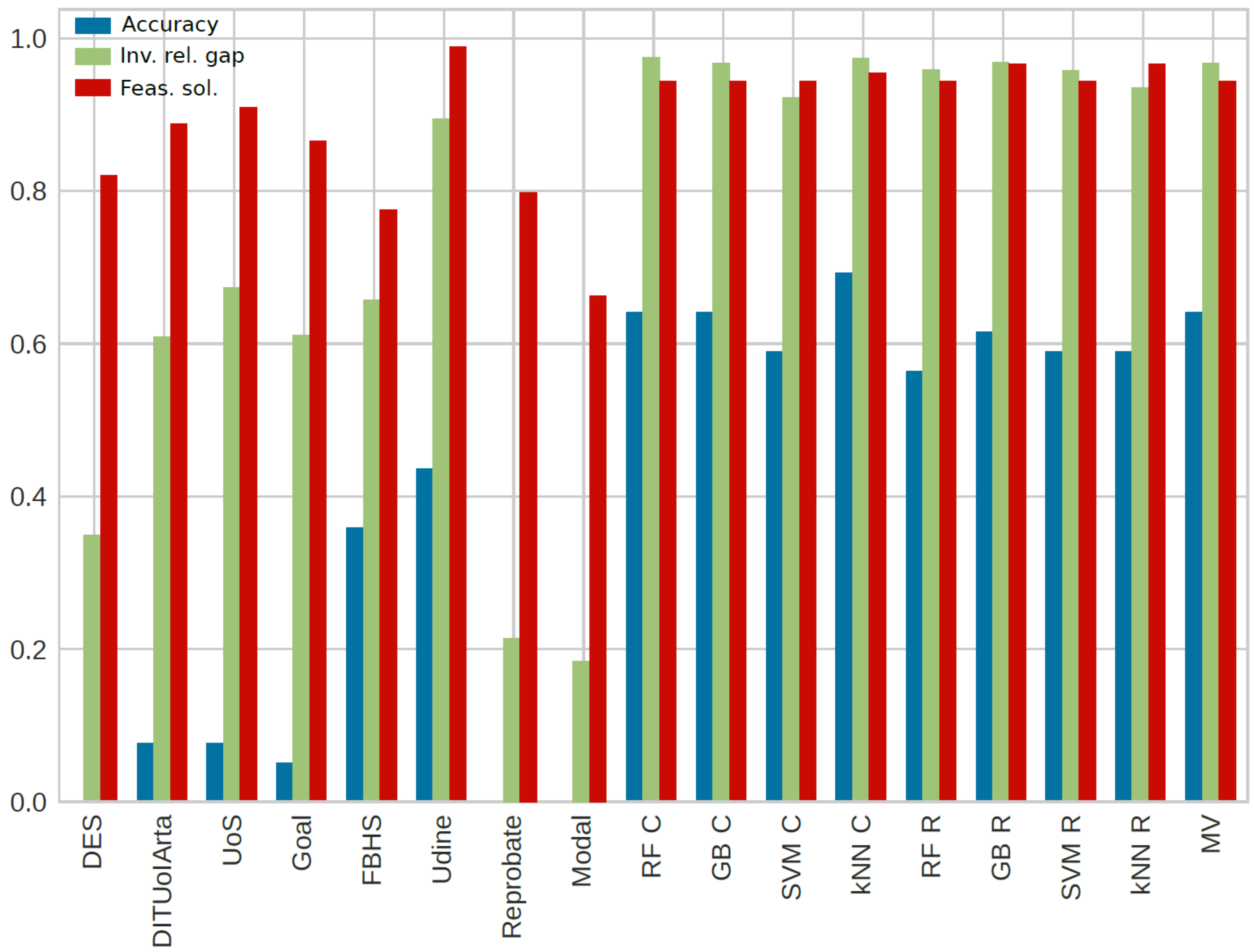
- K-nearest neighbours (kNN)
- Random forest (RF)
- Gradient-boosted trees (GB)
- Support vector machines (SVM)

- Majority voting (MV)

5. ALGORITHM SELECTION AND INSIGHTS



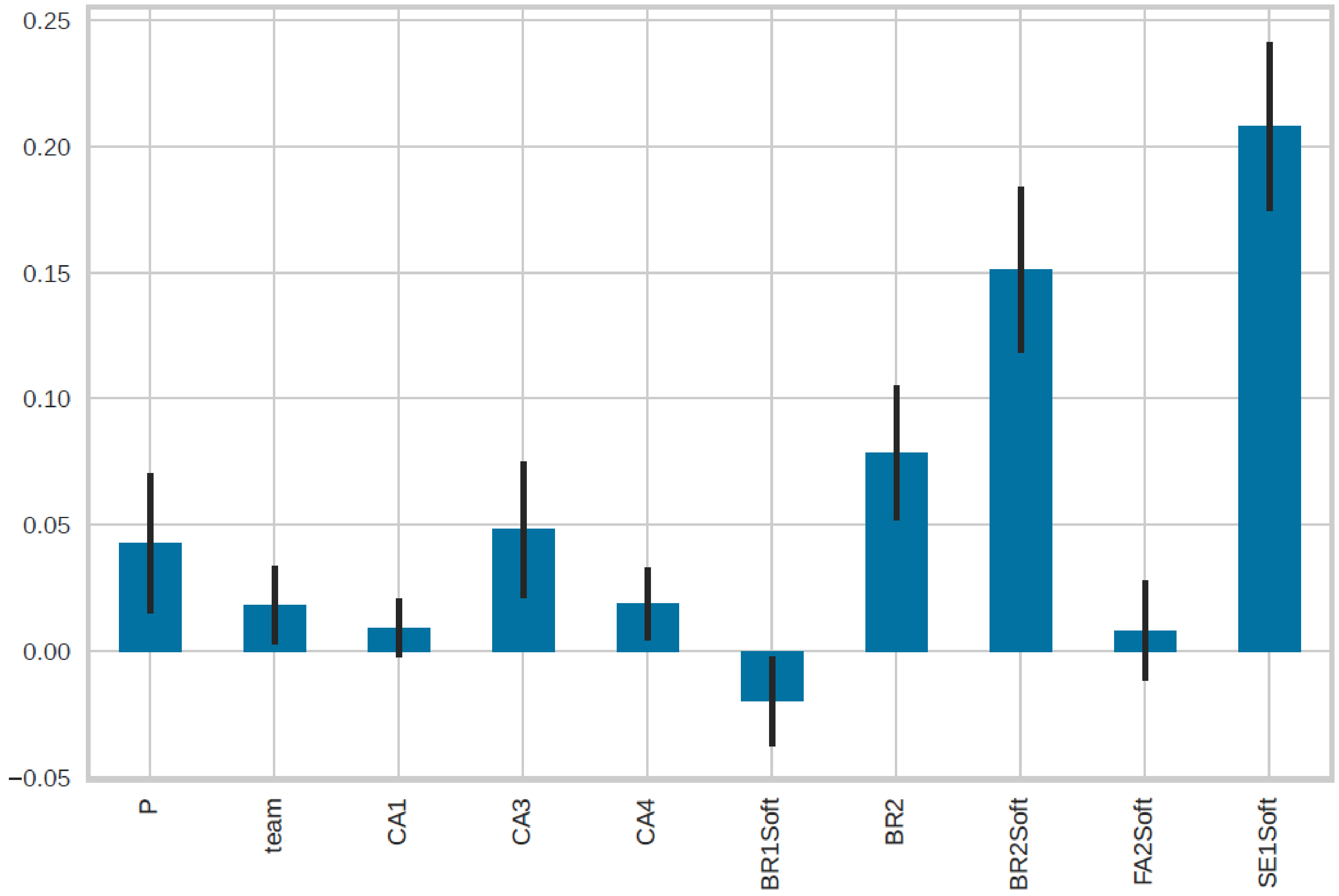
Validation set



Test set

GHENT UNIVERSITY

5. ALGORITHM SELECTION AND INSIGHTS



Feature permutation based on the kNN classifier, showing for each feature the mean decrease in classification accuracy when disabling the feature.

GHENT UNIVERSITY

CONCLUSION

From consulting

- find a customer with a particular problem
- develop a clever tailor-made method
- celebrate your success of beating a manual solution
- do not disclose the problem specifics

to science

- share your problem instance (RobinX XML instance repository)
- test your approach on other instances (benchmark instance set)
- learn what the strengths and weaknesses of your algorithm are (footprints)
- understand why, and improve

GHENT UNIVERSITY

REFERENCES

- Van Bulck, D & Goossens D. (2023) The international timetabling competition on sports timetabling (ITC2021). EJOR, 308(3), p.1249-1267
- Van Bulck D, Goossens D, Schönberger J & Guajardo M (2020) RobinX: a three-field classification and unified data format for round-robin sports timetabling (2020). EJOR, 280(2), p.568-580
- Van Bulck, D, Goossens D, et al. (2023) Which algorithm to select in sports timetabling? Working paper.

- Berthold T, Koch T, Shinano Y (2021) MILP. Try. Repeat. In: De Causmaecker P, Özcan E, Vanden Berghe G (eds) Proc. 13th Int. Conf. Pract. Theory Autom. Timetabling, PATAT, vol 2, pp 403–411
- Dimitsas A, Gogos C, Valouxis C, Tzallas A, Alefragis P (2022) A pragmatic approach for solving the sports scheduling problem. In: Proc. 13th Int. Conf. Pract. Theory Autom. Timetabling, PATAT, vol 3, pp 195–207
- Fonseca GHG, Toffolo TAM (2022) A fix-and-optimize heuristic for the ITC2021 sports timetabling problem. J Sched 25:273–286
- Lester MM (2022) Pseudo-boolean optimisation for robinx sports timetabling. J Sched, 25:287–299
- Lamas-Fernandez C, Martinez-Sykora A, Potts CN (2021) Scheduling double round robin sports tournaments. In: De Causmaecker P, Özcan E, Vanden Berghe G (eds) Proc. 13th Int. Conf. Pract. Theory Autom. Timetabling, PATAT, vol 2, pp 435 – 448
- Phillips AE, O'Sullivan M, Walker C (2021) An adaptive large neighbourhood search matheuristic for the ITC2021 sports timetabling competition. In: De Causmaecker P, Özcan E, Vanden Berghe G (eds) Proc. 13th Int. Conf. Pract. Theory Autom. Timetabling, PATAT, vol 2, pp 426 – 430
- Rosati RM, Petris M, Di Gaspero L, Schaerf A (2022) Multi-neighborhood simulated annealing for the sports timetabling competition ITC2021. J Sched 25:301–319
- Van Bulck D, Goossens D (2023) First-break-heuristically-schedule: Constructing highly-constrained sports timetables. Operations Research Letters 51:326–331

- Smith-Miles K, Baatar D, Wreford B, Lewis R (2014) Towards objective measures of algorithm performance across instance space. C&OR, 45:12-24
- Graham RL, Lawler EL, Lenstra JK, Rinnooy Kan AHG (1979). Optimization and Approximation in Deterministic Sequencing and Scheduling: a survey. Annals of Discrete Mathematics,Vol. 5, p. 287-326,

- My co-authors: David Van Bulck (Ghent U), Mario Guajardo (NHH), Jörn Schönberger (TU Dresden), Jan-Patrick Clarner (Zuse Institute Berlin), Angelos Dimitsas (U of Ioannina), George Fonseca (U Ouro Preto), Carlos Lamas-Fernandez (U Southampton), Martin Lester (U Reading), Jaap Pedersen (Zuse Institute Berlin), Antony Phillips (7bridges), Roberto Rosati (U Udine)
- My sports team: Xiajie Yi, Miao Li, Karel Devriesere, David Van Bulck

- All participants to ITC 2021
- EURO working groups OR in Sports & PATAT

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

DEPARTMENT OF BUSINESS INFORMATICS AND OPERATIONS MANAGEMENT

Sports scheduling: from consulting to science

Dries Goossens (Ghent University)

Contact
dries.goossens@ugent.be
www.sportscheduling.ugent.be

Scheduling seminar, May 10 2023

GHENT UNIVERSITY