

Max-SAT Evaluation 2007

Josep Argelich (Universitat de Lleida)

Chu Min Li (Université de Picardie)

Felip Manyà (Universitat de Lleida)

Jordi Planes (University of Southampton)



Objectives

- Evaluate the performance of existing solvers
- Identify challenging benchmarks
- Identify successful solving techniques
- Stimulate researchers to improve their solvers
and do more research in Max-SAT
- Gain new insights for future evaluations

Categories

- Max-SAT
- Weighted Max-SAT
- Partial Max-SAT
- Weighted Partial Max-SAT

Computational resources

Experiments performed on a Linux cluster with 2Ghz AMD Opteron processors with 1Gb of RAM provided by the Universitat de Lleida

S O L V E R S

- Clone: Pipatsrisawat, Chavira, Choi & Darwiche
- ChaffBS & ChaffLS: Fu & Malik
- LB-SAT: Lin & Su
- MaxSatz: Li, Manyà & Planes
- MaxSatz14: Darras, Dequen, Devendeville & Li
- MiniMaxSat: Heras, Larrosa, Oliveras & de Givry
- PMS: Argelich & Manyà
- Sat4jmaxsat: Le Berre
- SR(w): Ramírez & Geffner
- Toolbar: de Givry, Heras, Larrosa & Schiex
- W-MaxStaz: Argelich, Li & Manyà

Max-SAT Category

Solvers

- Clone
- LB-SAT
- MaxSatz14
- MaxSatz
- MiniMaxSat
- PMS
- Sat4jmaxsat
- SR(w)
- Toolbar
- W-MaxSatz

Max-SAT Category

Benchmarks

- **Max-SAT Instances**

- ▶ 815 instances with a timeout of 30 minutes
- ▶ Random Max-2-SAT and Max-3-SAT without timeout

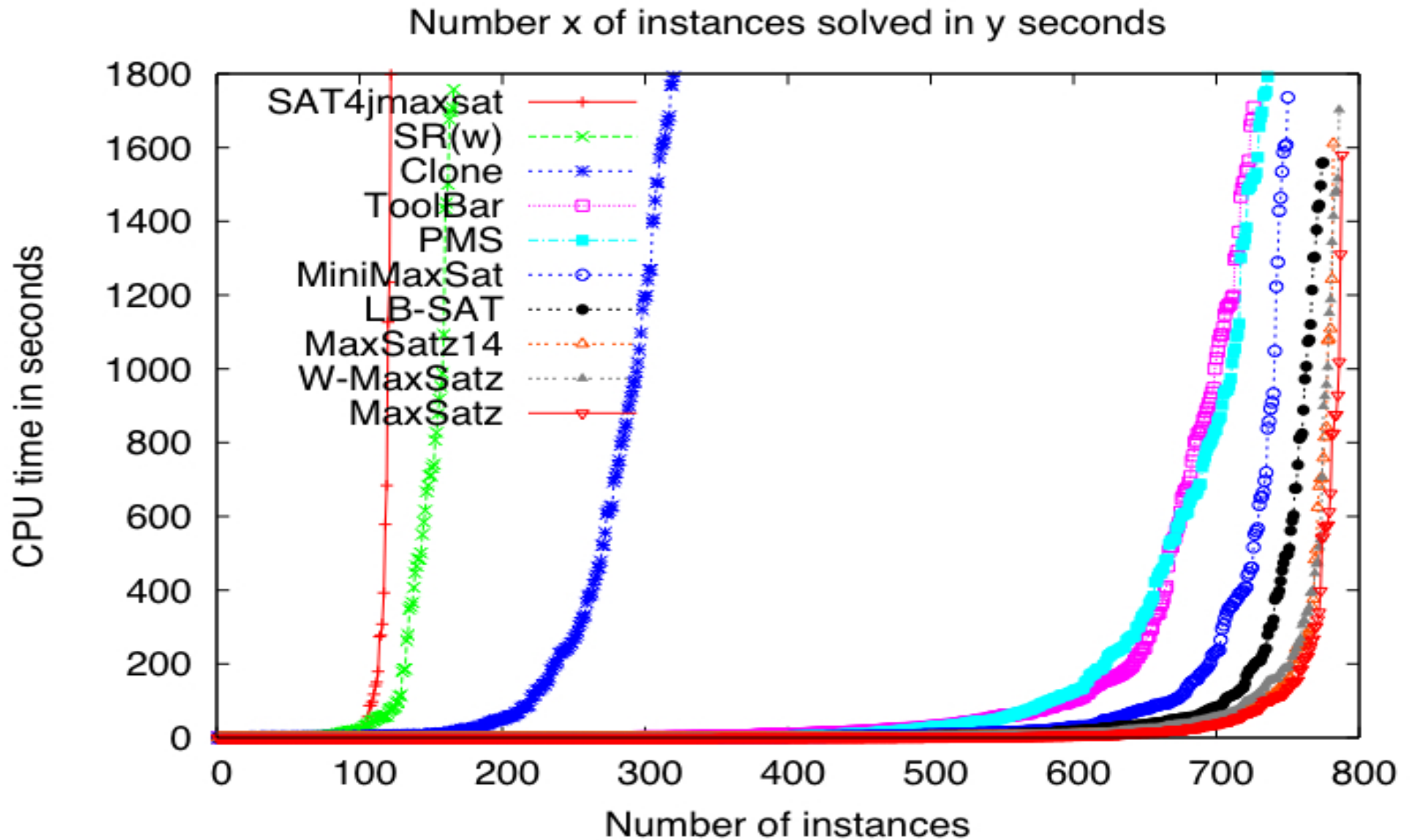
Max-SAT Category

Set Name	NI	Clone	LB-SAT	MaxSatz14	MaxSatz	MiniMaxSat	PMS	SAT4jmaxsat	SR(w)	ToolBar	W-MaxSatz
MAX3SAT/40VARS	40	376.02(28)	1.23 (40)	1.02 (40)	1.05 (40)	3.34 (40)	9.48 (40)	1462.17(2)	629.93(9)	7.20 (40)	1.48 (40)
MAX3SAT/50VARS	40	492.35(16)	7.84 (40)	6.04 (40)	5.90 (40)	25.79(40)	58.45 (40)	480.68(3)	1287.91(2)	57.64 (40)	8.60 (40)
MAX3SAT/60VARS	40	356.90(13)	24.13(40)	15.61(40)	14.24(40)	77.53(38)	128.38 (40)	68.05(9)	650.58(4)	272.90 (40)	21.63 (40)
MAX3SAT/70VARS	40	7.79 (10)	124.68(40)	57.85(40)	48.82(40)	207.90(35)	191.93 (37)	2.24 (10)	891.70(8)	334.33 (29)	77.88 (40)
SPINGLASS	20	6.19 (10)	11.83(20)	43.01(20)	69.40(20)	4.56 (20)	3.29 (10)	-	24.51(10)	24.02 (10)	80.76 (20)
RAMSEY	48	103.20(33)	21.15(35)	12.27(29)	8.99 (34)	29.81(34)	29.99 (35)	2.88 (33)	55.88(23)	20.40 (35)	16.57 (34)
MAX2SAT/100VARS	110	138.34(31)	10.53(110)	1.84 (110)	1.78 (110)	9.62 (110)	40.82 (110)	17.83(10)	97.45(20)	29.02 (110)	2.54 (110)
MAX2SAT/140VARS	110	112.22(31)	156.54(103)	26.83(110)	29.57(110)	121.54(99)	155.06 (93)	37.74(15)	4.77 (20)	235.40 (96)	39.48 (110)
MAX2SAT/60VARS	110	329.83(51)	0.11 (110)	0.03 (110)	0.03 (110)	0.19 (110)	0.23 (110)	-	140.84(21)	0.69 (110)	0.04 (110)
MAX3SAT/40VARS	50	373.87(34)	1.74 (50)	1.43 (50)	1.50 (50)	5.53 (46)	15.09 (50)	5.40 (10)	17.02(10)	9.54 (50)	2.13 (50)
MAX3SAT/60VARS	50	134.16(20)	36.05(50)	25.22(50)	23.33(50)	111.81(50)	214.28 (50)	1.61 (10)	5.48 (10)	339.96 (48)	35.40 (50)
MAX3SAT/80VARS	50	151.15(20)	170.41(42)	210.89(48)	197.58(49)	230.82(37)	253.57 (41)	111.81(18)	0.45 (10)	241.94 (28)	245.23 (47)
MAXCUT/DIMACS_MOD	62	123.42(21)	156.01(52)	83.66(52)	83.86(52)	100.06(48)	333.28(44)	0.93 (2)	305.10(16)	127.82(48)	145.06(52)
MAXCUT/RANDOM	40	-	10.66(40)	5.43 (40)	5.58 (40)	15.88(40)	683.22 (34)	-	-	55.54(40)	8.43 (40)
MAXCUT/SPINGLASS	5	2.67 (2)	7.60 (3)	25.99(3)	44.96(3)	1.62 (3)	0.41 (2)	-	9.96 (2)	4.75 (2)	54.07 (3)

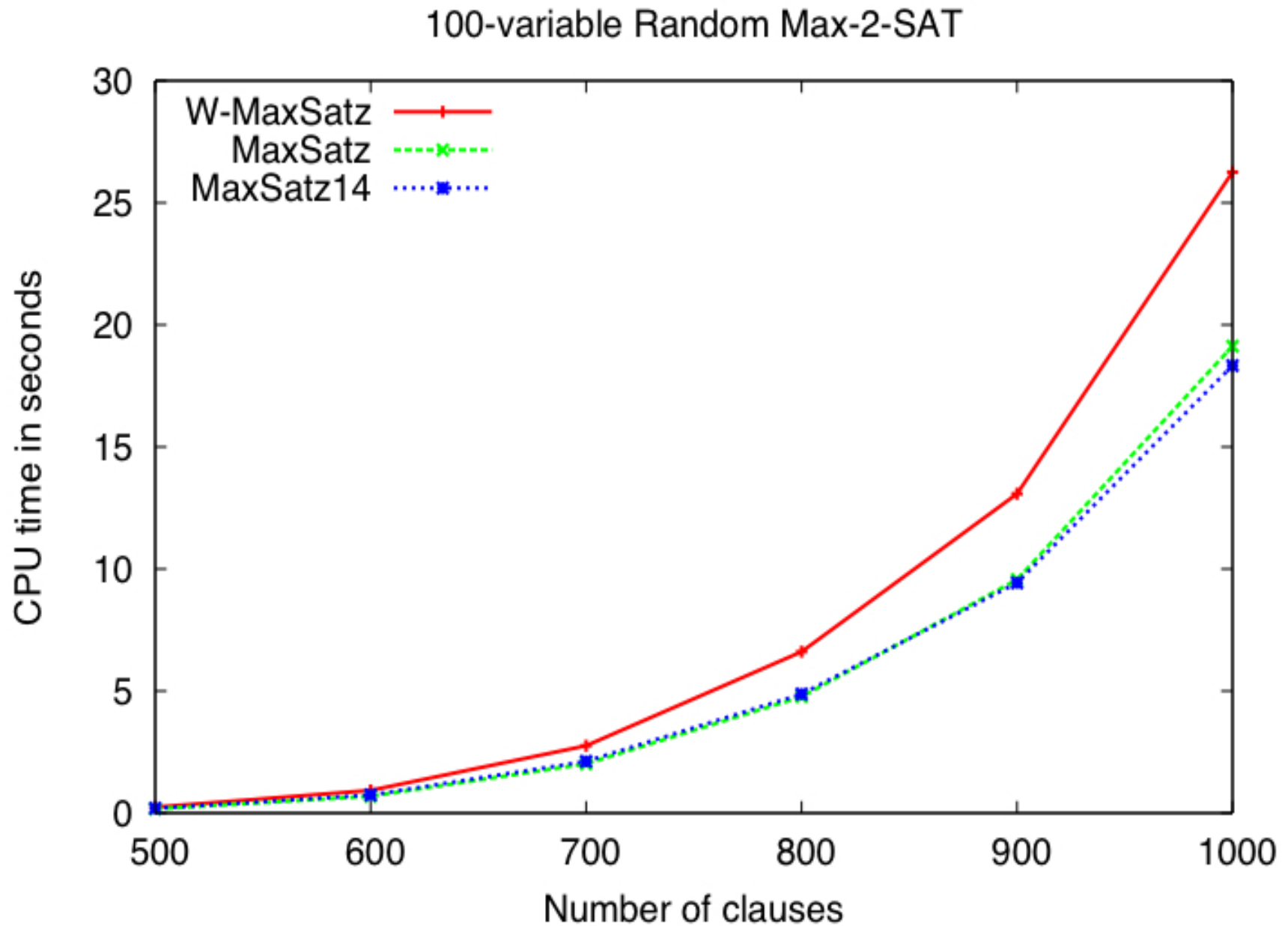
- For each solver and for each set of instances, we display the mean time of the solved instances and the number of solved instances (in brackets). Time in seconds. Timeout: 30 minutes.

- The best performing solver: it solves the maximum number of instances in minimum time
- Solvers solving the same number of instances as the best performing solver

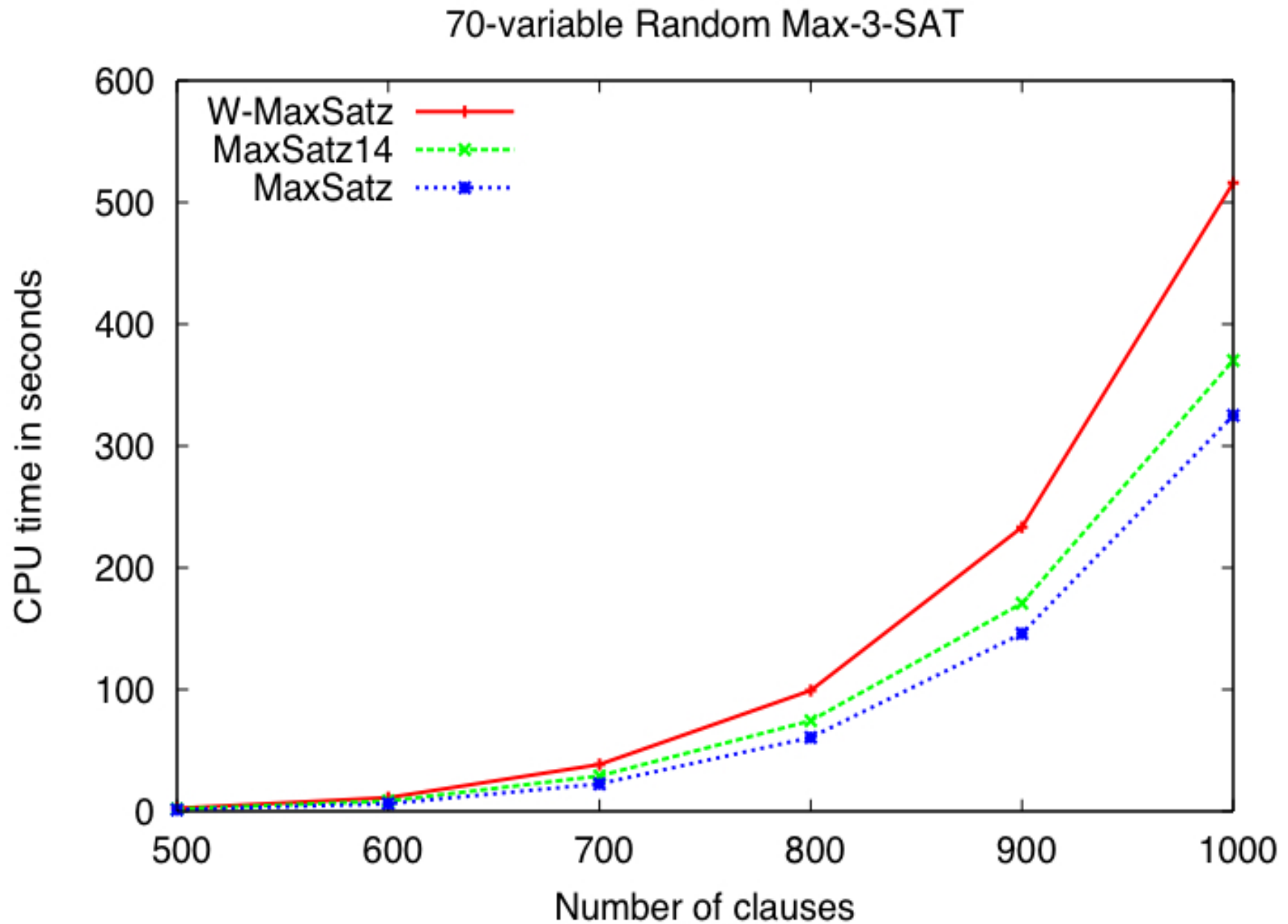
Max-SAT Category



Max-SAT Category



Max-SAT Category



Weighted Max-SAT Category

Solvers

- Clone
- LB-PSAT
- MiniMaxSat
- Sat4jmaxsat
- SR(w)
- Toolbar
- W-MaxSatz

Weighted Max-SAT Category

Benchmarks

- **Weighted Max-SAT Instances**

- ▶ 325 instances with a timeout of 30 minutes

- ▶ Random Weighted Max-2-SAT and Weighted Max-3-SAT

without timeout

Weighted Max-SAT Category

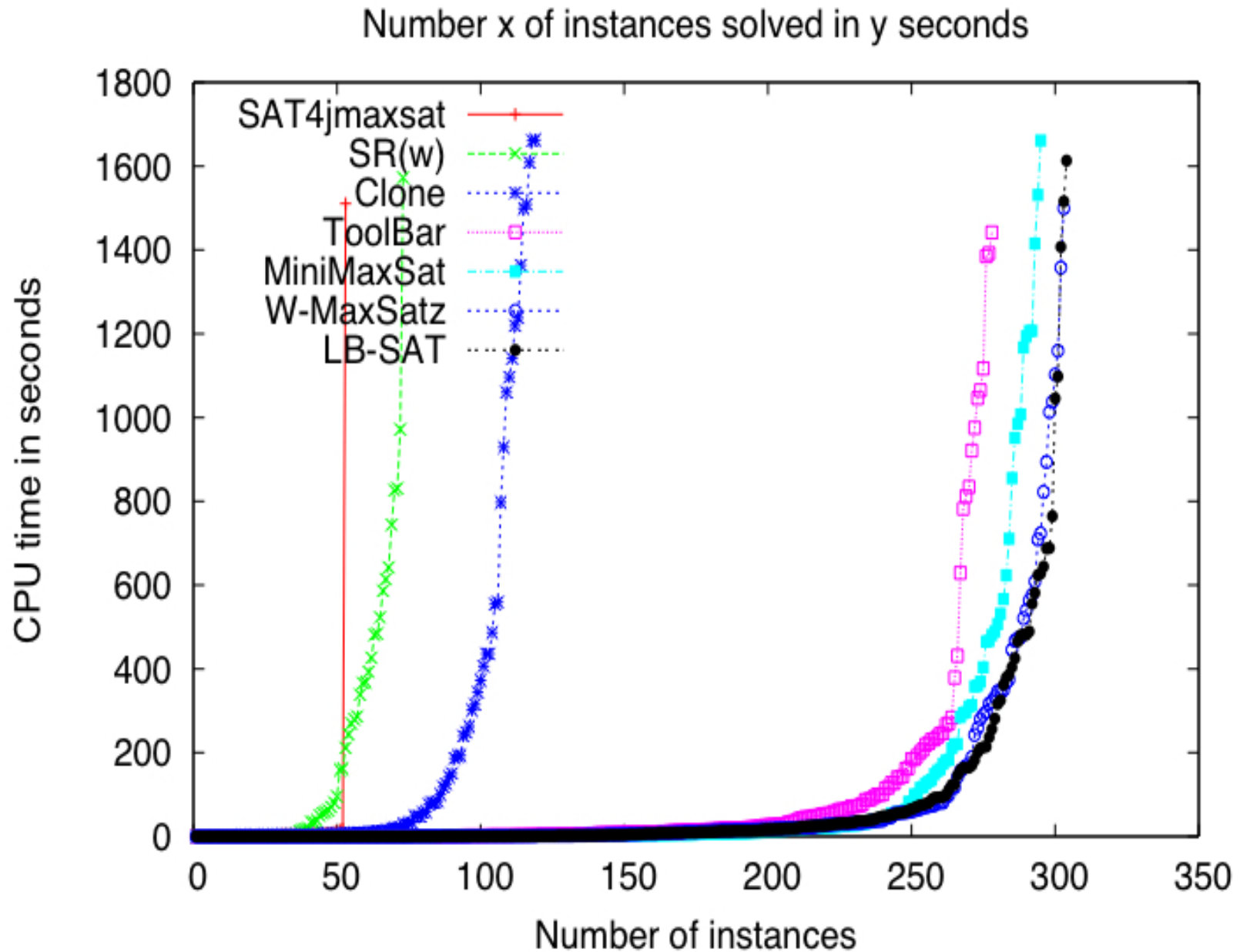
Set Name	NI	Clone	LB-SAT	MiniMaxSat	SAT4jmaxsat	SR(w)	ToolBar	W-MaxSatz
RAMSEY	48	98.37(35)	3.59 (36)	7.08 (36)	3.60 (32)	82.89 (25)	5.48 (35)	44.85(36)
WMAX2SAT	90	197.41(34)	18.79 (90)	10.59(90)	156.50(10)	95.18 (20)	34.70 (90)	7.95 (90)
WMAX3SAT	80	248.65(23)	207.84 (80)	280.49(70)	6.52 (9)	414.35 (9)	242.76 (51)	191.40(80)
WMAXCUT/DIMACS_MOD	62	325.43 (25)	75.77 (55)	75.46 (55)	1.32 (2)	200.11 (16)	81.48 (57)	93.79 (55)
WMAXCUT/RANDOM	40	-	16.39 (40)	5.42 (40)	-	-	15.91 (40)	19.22(40)
WMAXCUT/SPINGLASS	5	2.57 (2)	2.50 (3)	45.50(4)	-	7.38 (2)	89.23 (3)	35.77(2)

- For each solver and for each set of instances, we display the mean time of the solved instances and the number of solved instances (in brackets). Time in seconds. Timeout: 30 minutes.

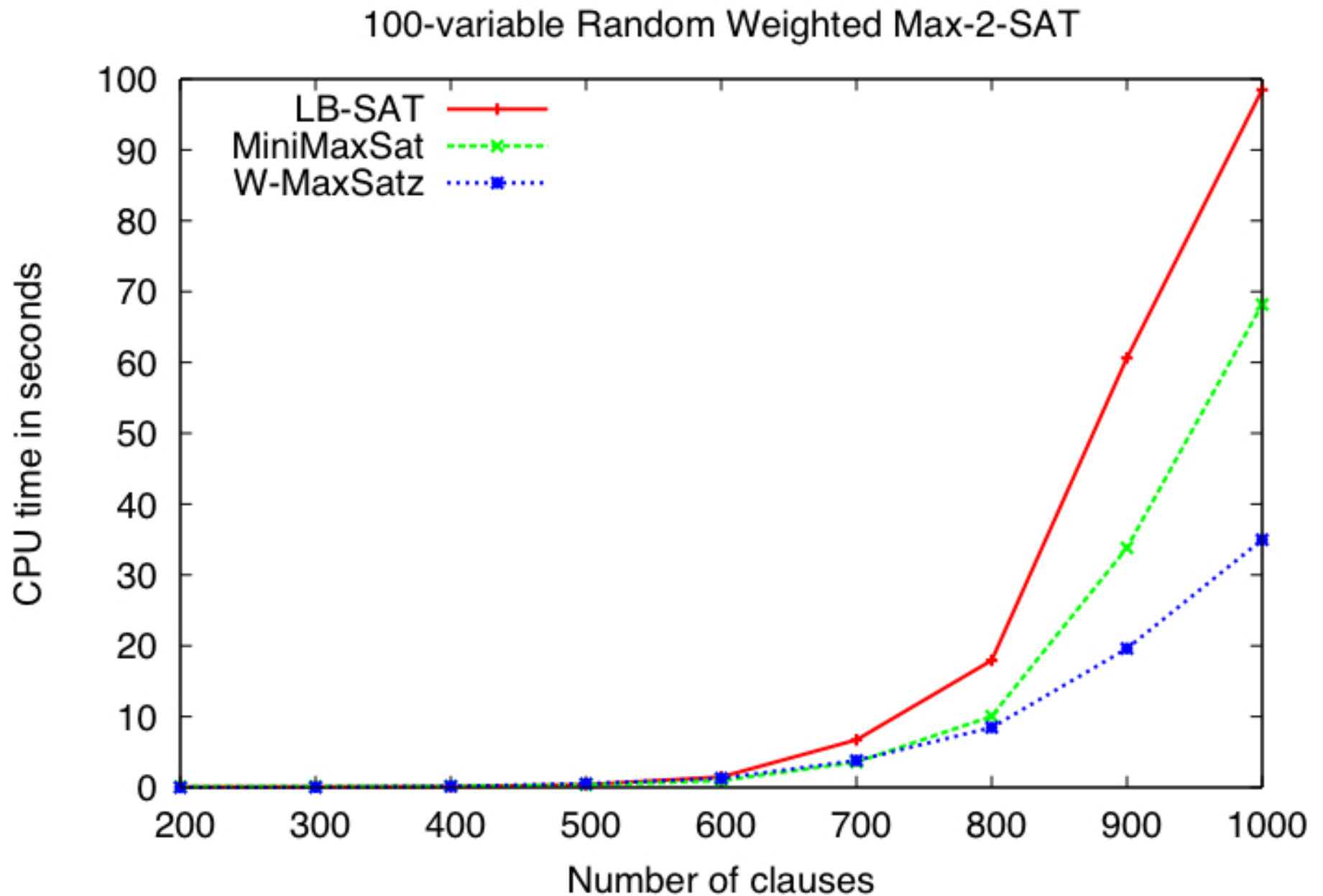
- The best performing solver: it solves the maximum number of instances in minimum time

- Solvers solving the same number of instances as the best performing solver

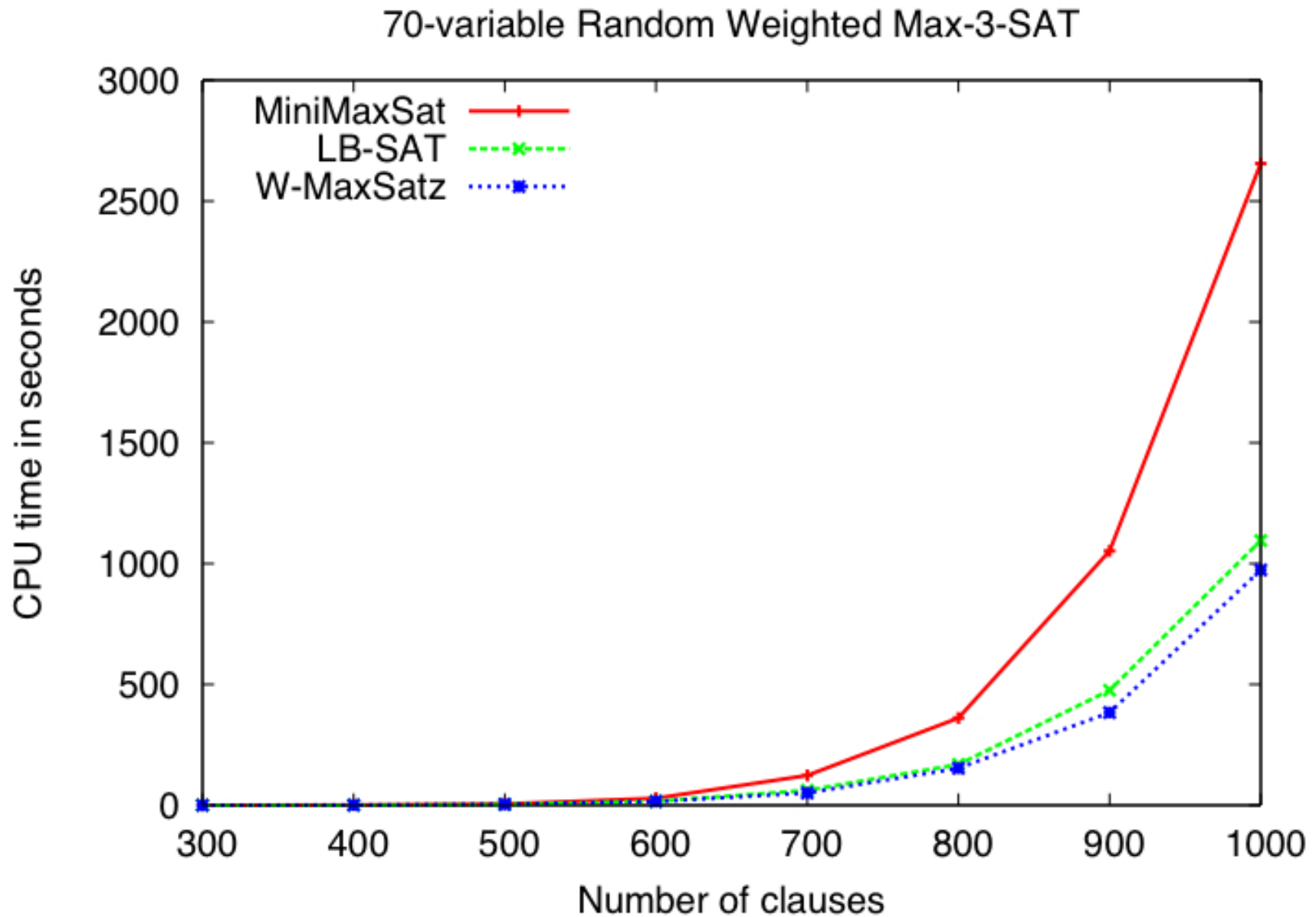
Weighted Max-SAT Category



Weighted Max-SAT Category



Weighted Max-SAT Category



Partial Max-SAT Category

Solvers

- Chaff_BS
- Chaff_LC
- Clone
- LB-PSAT
- MiniMaxSat
- PMS
- Sat4jmaxsat
- SR(w)
- Toolbar
- W-MaxSatz

Partial Max-SAT Category

Benchmarks

- **Partial Max-SAT Instances**

- ▶ 722 instances with a timeout of 30 minutes
- ▶ Random Partial Max-2-SAT and Partial Max-3-SAT
without timeout

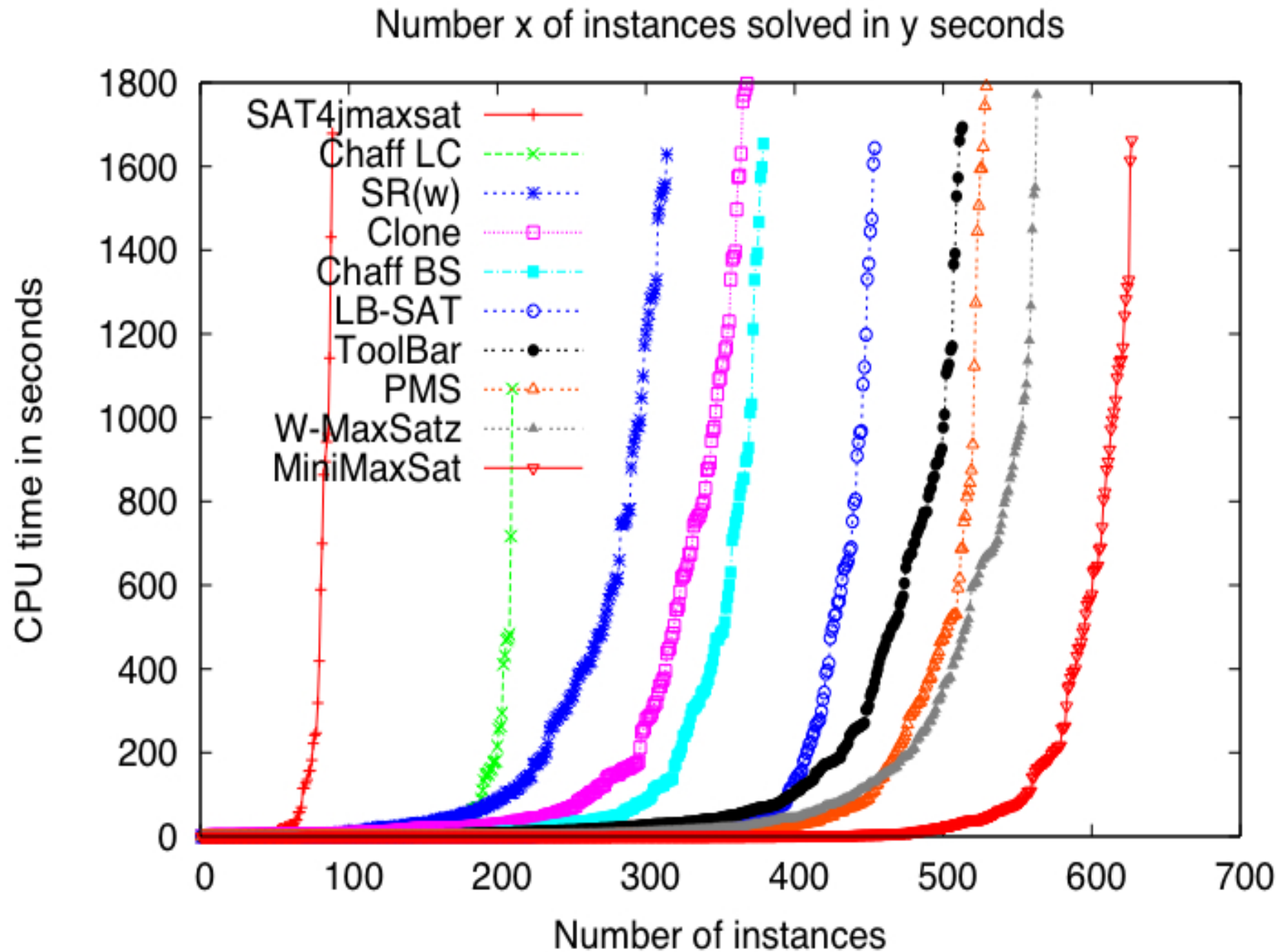
Partial Max-SAT Category

Set Name	NI	Chaff_BS	Chaff_LC	Clone	LB-SAT	MiniMaxSat	PMS	SAT4jmaxsat	SR(w)	ToolBar	W-MaxSatz
RANDOM/PMAX2SAT/	90	-	-	8.20 (1)	305.93 (59)	221.57(83)	220.27 (44)	-	-	149.86(89)	40.42(90)
RANDOM/PMAX3SAT/	60	40.25(24)	22.44(22)	251.62(19)	52.42 (59)	156.47(58)	80.83 (59)	4.57 (20)	327 (16)	172.31(47)	59.01(60)
MAXCLIQUE/RANDOM/	96	146.24(54)	-	189.65(79)	9.89 (96)	2.39 (96)	68.19 (96)	-	225.38 (55)	11.39(96)	49.34(80)
MAXCLIQUE/STRUCTURED/	62	282.83(19)	54.44(9)	308.72(16)	128.34 (32)	85.26(36)	171.13 (27)	13.16(1)	19.35 (9)	202.68(33)	153.30(22)
MAXONE/3SAT/	80	402.14(23)	11.67(41)	420.67(54)	62.18 (76)	1.30 (80)	4.23 (80)	1013.93(5)	273.87 (70)	102.34(80)	199.16(77)
MAXONE/STRUCTURED/	60	52.98(57)	81.21(2)	258.19(32)	2.29 (2)	31.04(60)	176.71 (37)	412.66(3)	443.59 (22)	221.31(44)	385.89(54)
PSEUDO/garden/	7	1.34 (5)	0.78 (5)	2.59 (5)	0.47 (5)	7.13 (5)	0.55 (5)	1.42 (3)	2.55 (5)	1.82 (4)	2.16 (4)
PSEUDO/logic-synthesis/	17	39.42(2)	32.16(4)	-	865.73 (3)	216.28(2)	2.55 (1)	-	-	-	-
PSEUDO/primes-dimacs-cnf/	148	72.92(99)	41.25(46)	89.72(99)	82.68 (35)	88.15(107)	124.09 (88)	82.11(45)	67.03 (77)	68.71(60)	129.97(85)
PSEUDO/routing/	15	180.33(15)	0.22 (14)	19.08(5)	-	93.88(14)	25.98 (5)	-	-	-	143.94(5)
WCSP/MAXCSP/DENSE_LOOSE/	20	324.93(14)	143.86(6)	831.09(1)	1.16 (20)	0.65 (20)	2.03 (20)	-	588.37 (1)	336.71(15)	7.19 (20)
WCSP/MAXCSP/DENSE_TIGHT/	20	65.83(20)	106.81(18)	25.90(20)	2.87 (20)	0.68 (20)	2.25 (20)	-	199.93 (18)	461.84(20)	10.53(20)
WCSP/MAXCSP/SPARSE_LOOSE/	20	19.16(20)	41.80(19)	122.28(13)	1.86 (20)	0.35 (20)	1.42 (20)	222.86(10)	264.08 (16)	4.18 (10)	25.55(20)
WCSP/MAXCSP/SPARSE_TIGHT/	20	28.87(20)	16.23(19)	29.58(20)	7.14 (20)	0.85 (20)	2.19 (20)	-	219.99 (19)	20.36(10)	26.04(20)
WCSP/WQUEENS/	7	13.94(7)	18.94(5)	80.49(4)	5.26 (7)	0.52 (6)	12.95 (7)	11.17(2)	45.17 (6)	12.73(5)	85.10(6)

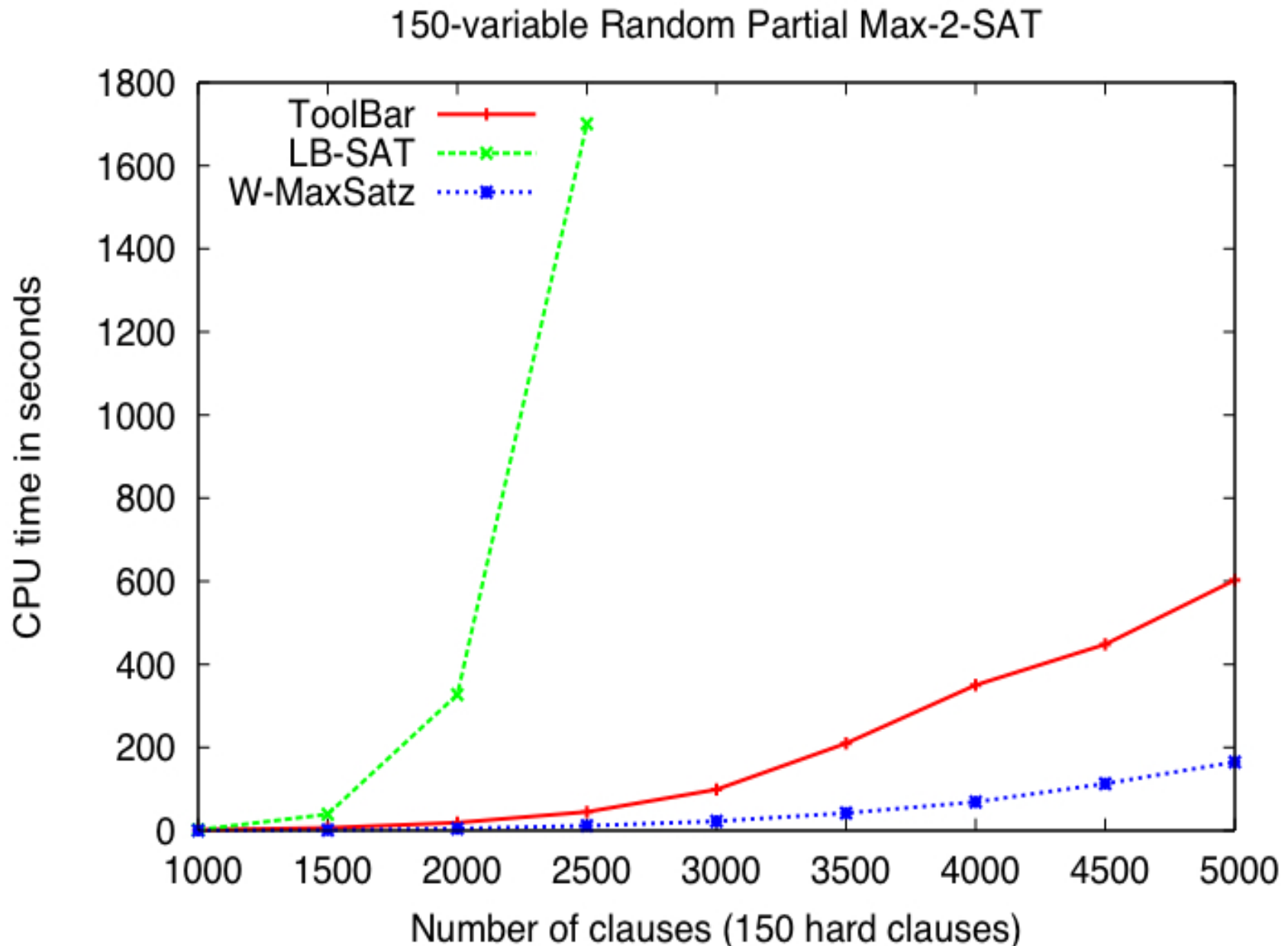
- For each solver and for each set of instances, we display the mean time of the solved instances and the number of solved instances (in brackets). Time in seconds. Timeout: 30 minutes.

- The best performing solver: it solves the maximum number of instances in minimum time
- Solvers solving the same number of instances as the best performing solver

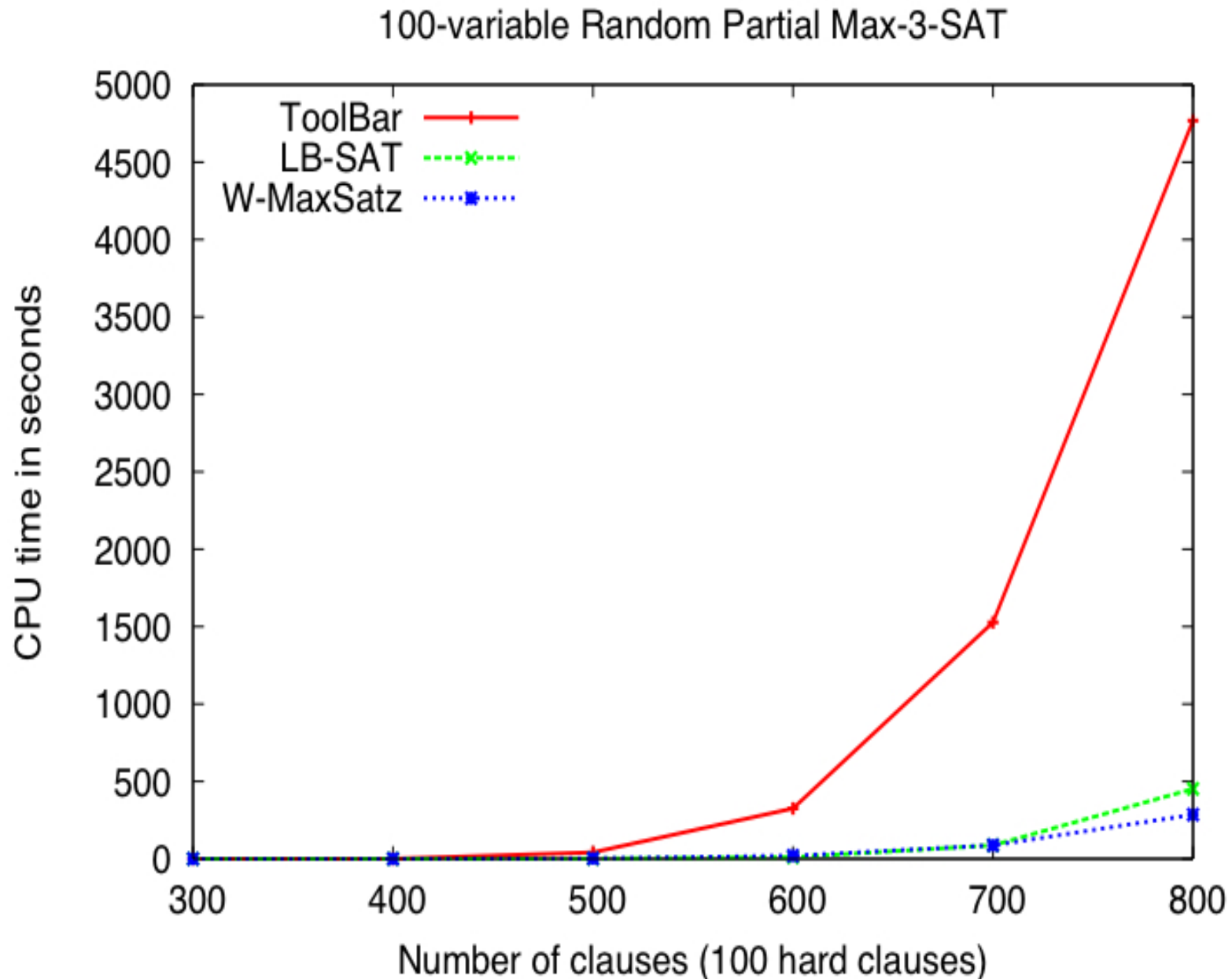
Partial Max-SAT Category



Partial Max-SAT Category



Partial Max-SAT Category



Weighted Partial Max-SAT Category

Solvers

- Clone
- MiniMaxSat
- Sat4jmaxsat
- SR(w)
- Toolbar
- W-MaxSatz

Weighted Partial Max-SAT Category

Benchmarks

- **Weighted Partial Max-SAT Instances**
 - ▶ 746 instances with a timeout of 30 minutes
 - ▶ Random Weighted Partial Max-2-SAT and
Weighted Partial Max-3-SAT without timeout

Weighted Partial Max-SAT Category

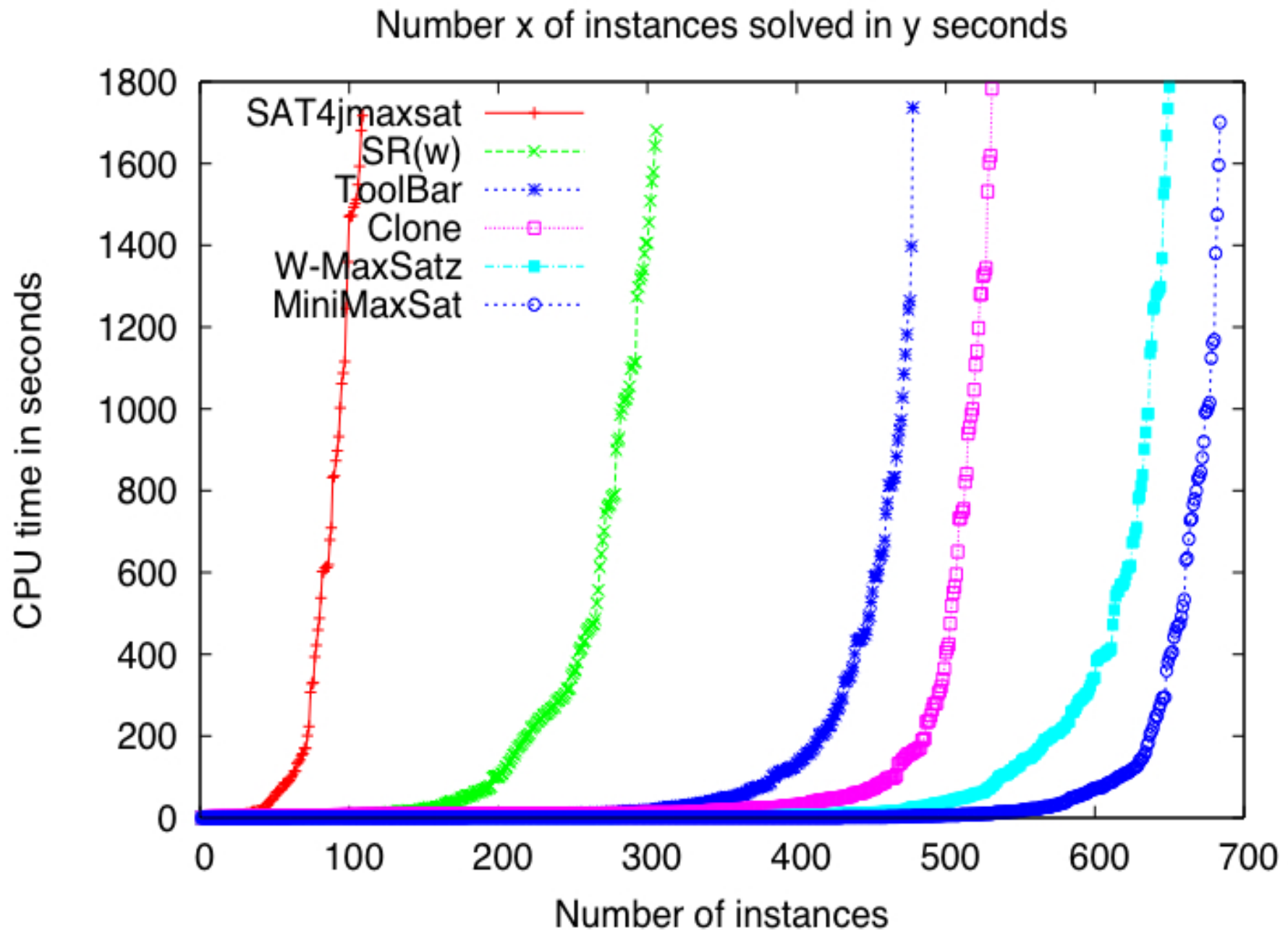
Set_Name	NI	Clone	MiniMaxSat	SAT4jmaxsat	SR(w)	ToolBar	W-MaxSatz
RANDOM/WPMAx2SAT/	90	-	246.28(81)	-	-	213.24(88)	196.30(88)
RANDOM/WPMAx3SAT/	60	136.28(21)	186.63(58)	6.41 (20)	275.44 (17)	188.75(47)	91.81(60)
AUCTIONS/AUC_PATHS/	88	50.78(88)	31.55(88)	-	163.45 (77)	48.68(88)	243.98(70)
AUCTIONS/AUC_REGIONS/	84	30.51(84)	1.61 (84)	-	130.20 (82)	6.45 (84)	6.70 (84)
AUCTIONS/AUC_SCHEDULING/	84	228.16(74)	46.22(84)	-	231.83 (55)	74.11(82)	103.85(82)
PSEUDO/factor/	186	9.85 (186)	1.17 (186)	598.29(55)	-	246.39(12)	0.43 (186)
PSEUDO/miplib/	16	132.42(5)	41.66(5)	6.74 (3)	244.85 (6)	2.92 (4)	1.50 (4)
QCP/	25	-	25.01(20)	377.01(14)	652.50 (5)	191.07(12)	37.54(11)
WCSP/PLANNING/	71	261.15(62)	9.97 (71)	73.22(16)	365.48 (52)	22.82(52)	101.49(59)
WCSP/SPOT5/DIR/	21	9.32 (6)	3.83 (3)	0.56 (1)	2.92 (6)	128.04(5)	17.35(2)
WCSP/SPOT5/LOG/	21	7.27 (5)	9.18 (4)	0.54 (1)	14.91 (6)	111.41(4)	640.86(4)

- For each solver and for each set of instances, we display the mean time of the solved instances and the number of solved instances (in brackets). Time in seconds. Timeout: 30 minutes.

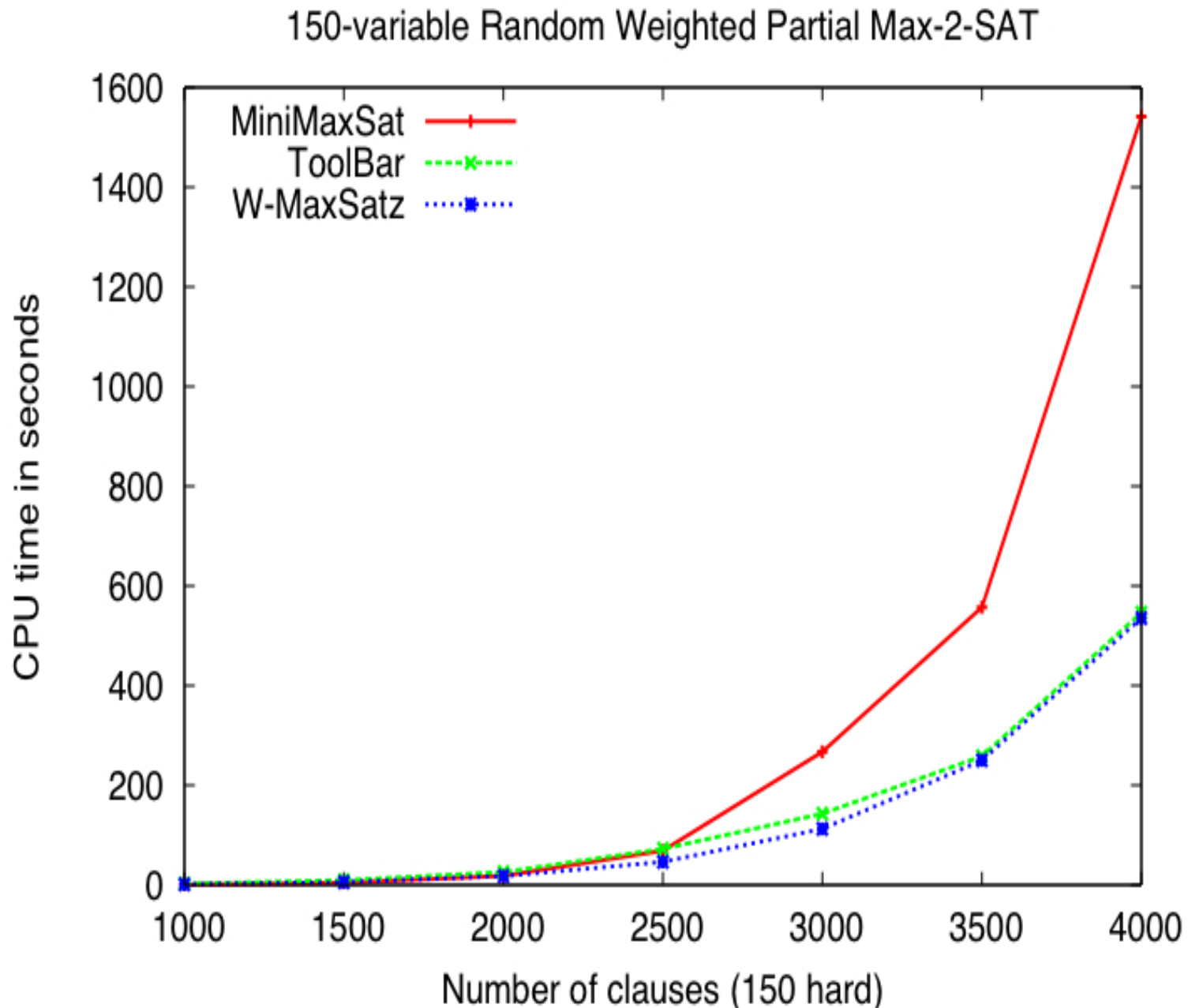
- The best performing solver: it solves the maximum number of instances in minimum time

- Solvers solving the same number of instances as the best performing solver

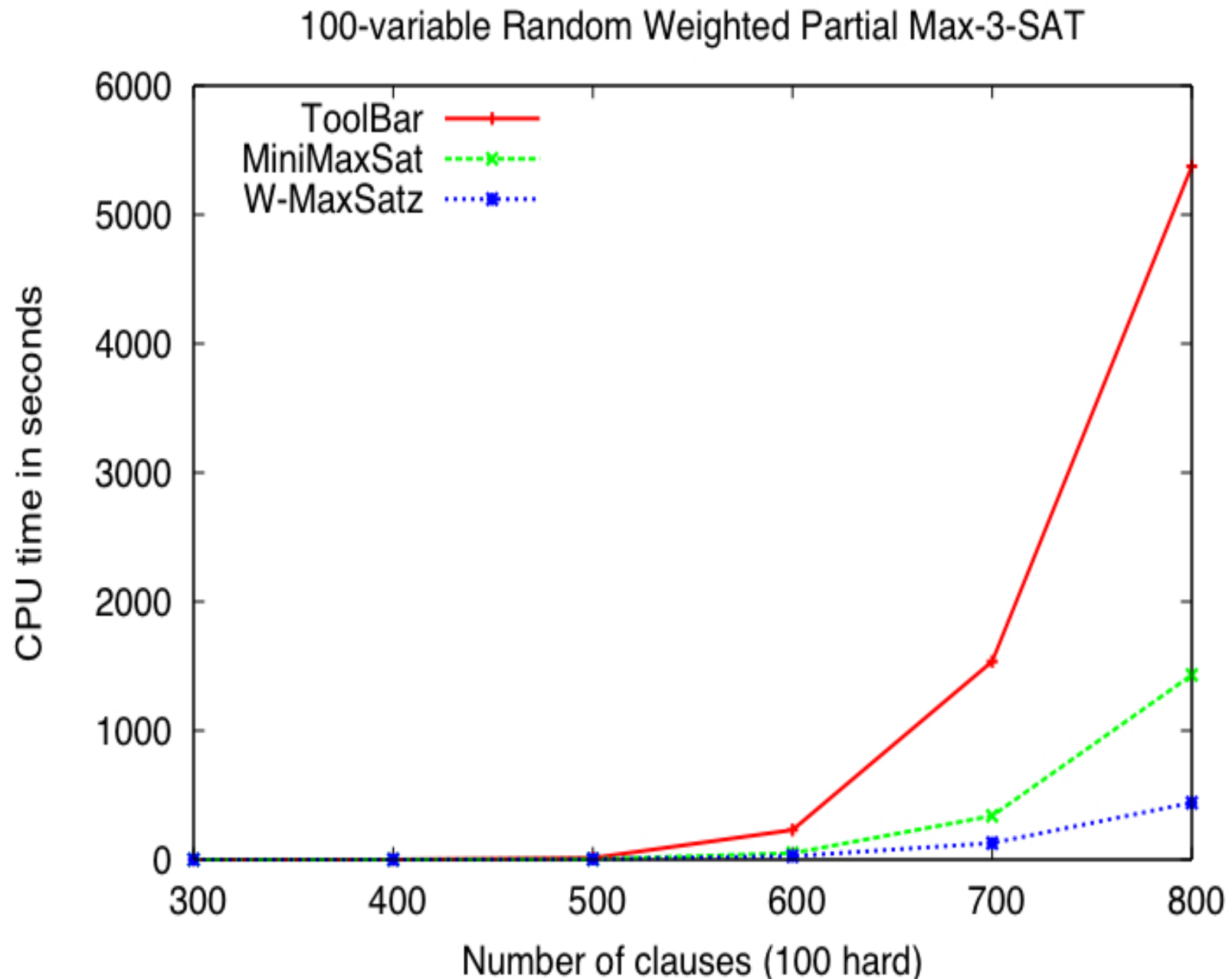
Weighted Partial Max-SAT Category



Weighted Partial Max-SAT Category



Weighted Partial Max-SAT Category



Thanks to the people that contributed solvers and benchmarks, and to the Universitat de Lleida for allowing to use its cluster

We encourage you to participate in the Max-SAT Evaluation 2008

