A Multi-Objective Optimization Approach for the Integration and Test Order Problem

Wesley Klewerton Guez Assunção[a], Thelma Elita Colanzi[a,b], Silvia Regina Vergilio[a,*], Aurora Pozo[a]

[a] *Computer Science Department, Federal University of Paraná (UFPR), Brazil, CP 19:081, CEP: 81531-970, Curitiba, Brazil*
[b] *Computer Science Department (DIN), State University of Maringá (UEM), Av. Colombo, 5790, Bloco C-56, CEP: 87.020-900, Maringá, Brazil*

Abstract

This text presents the main results from the MOCAITO evaluation, related to the application of the quality indicators and comparison of the multi-objective algorithms.

1. Main Results

Table 1 presents the main results obtained by each algorithm and system in Experiments 2M and 4M. Column 2 presents the cardinality of PF_{true}, formed by the non-dominated solutions obtained considering all algorithms executions. The average number of solutions found by each algorithm per run is also presented, as well as, in parentheses, the number of solutions of the set PF_{known}. In addition, the average of runtime, in seconds, and the standard deviation (in parentheses) are presented for all systems.

*Corresponding author
Email addresses: wesleyk@inf.ufpr.br (Wesley Klewerton Guez Assunção), thelmae@inf.ufpr.br (Thelma Elita Colanzi), silvia@inf.ufpr.br (Silvia Regina Vergilio), aurora@inf.ufpr.br (Aurora Pozo)

A preliminary analysis of this table shows how hard is to solve the integration and test order problem for each system. JBoss, Health Watcher and Toll System are simpler, since they have only one solution, independently of the used objectives. Moreover, all the algorithms find this solution in almost all runs. They are examples of systems in which the objectives are not in conflict, for that reason the three MOEAS achieve a single solution. This does not happen with the other systems with greater number of dependencies (and LOC). For them PF_{true} contains a greater number of solutions in both experiments. It is also noticeable that the greater the number of objectives, the greater this cardinality. The main difference with respect to the PF_{true} cardinality occurs for MyBatis. Maybe in this system the methods have more parameters and return values to be emulated.

We can also observe that SPEA2 requires a higher runtime with greater standard deviation than NSGA-II and PAES and that there is no increase in the execution time for all algorithms in Experiment 4M. A simple analysis shows that all the algorithms can be used to efficiently solve the problem, however the quality indicators described in next section give us a better comparison.

1.1. Coverage

Table 2 presents the results of the C indicator for Experiments 2M and 4M. The solutions of the MOEA that appears in a row have the value of domination on the solutions of the MOEA that appears in the column. For example, in Experiment 2M, the solutions found by PAES for AJHotDraw cover 83% of the solutions found by NSGA-II for the same system. Values greater than 0.5 are significant and indicate more than 50% of dominance.

Table 1: Number of Solutions and Runtime

Experiment 2M							
System	PF_{true}	**NSGA-II**		**SPEA2**		**PAES**	
		# Solutions	**Runtime**	**# Solutions**	**Runtime**	**# Solutions**	**Runtime**
BCEL	29	28.73 (29)	5.37 (0.04)	28.93 (29)	184.51 (21.88)	25.70 (29)	**1.88** (0.07)
JBoss	1	1.00 (1)	19.25 (0.18)	1.00 (1)	2666.66 (585.32)	1.00 (1)	**10.46** (0.08)
JHotDraw	3	1.40 (3)	31.29 (0.18)	1.23 (2)	3213.17 (677.40)	1.83 (2)	**19.06** (0.13)
MyBatis	63	60.60 (63)	79.18 (0.33)	58.60 (57)	132.41 (15.55)	43.00 (54)	**52.30** (0.20)
AJHotDraw	7	4.57 (6)	81.44 (0.41)	4.63 (6)	1375.29 (418.06)	5.87 (7)	**53.53** (0.31)
AJHSQLDB	40	31.53 (35)	67.13 (0.21)	26.40 (36)	101.65 (3.86)	26.23 (40)	**44.57** (0.23)
Health Watcher	1	1.00 (1)	13.00 (0.17)	1.00 (1)	2897.21 (744.19)	1.07 (1)	**6.72** (0.07)
Toll System	1	1.00 (1)	7.10 (0.08)	1.00 (1)	3541.92 (804.71)	1.00 (1)	**2.72** (0.01)
Experiment 4M							
System	PF_{true}	**NSGA-II**		**SPEA2**		**PAES**	
		# Solutions	**Runtime** (s)	**# Solutions**	**Runtime** (s)	**# Solutions**	**Runtime**
BCEL	37	37.43 (37)	**5.91** (0.05)	36.70 (37)	123.07 (18.84)	39.30 (37)	6.58 (1.25)
JBoss	1	1.00 (1)	18.73 (0.20)	1.00 (1)	2455.35 (612.18)	1.13 (1)	**10.69** (0.62)
JHotDraw	11	8.40 (10)	29.85 (0.34)	9.63 (9)	922.99 (373.98)	10.47 (19)	**24.29** (1.50)
MyBatis	789	276.37 (941)	**74.03** (0.87)	248.77 (690)	128.88 (2.65)	243.60 (679)	104.30 (7.91)
AJHotDraw	94	70.03 (79)	75.05 (0.57)	68.87 (78)	195.56 (28.22)	40.73 (84)	**62.07** (2.16)
AJHSQLDB	266	156.63 (360)	**62.34** (0.53)	119.10 (52)	104.29 (0.68)	145.97 (266)	75.62 (5.27)
Health Watcher	1	1.00 (1)	12.72 (0.15)	1.00 (1)	2580.39 (596.29)	1.07 (1)	**8.27** (0.58)
Toll System	1	1.00 (1)	7.33 (0.09)	1.00 (1)	3516.71 (570.76)	1.07 (1)	**4.10** (0.75)

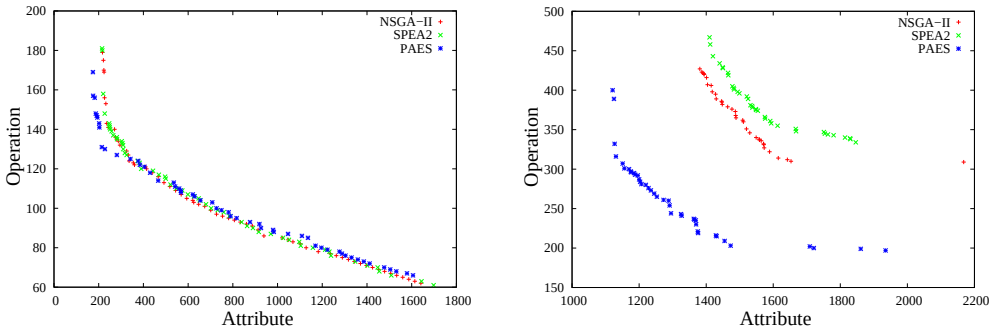
The significant values are highlighted in bold.

Table 2: Indicator C for PF_{know} sets

Experiment 2M								
System	**MOEA**	**NSGA-II**	**SPEA2**	**PAES**	**System**	**NSGA-II**	**SPEA2**	**PAES**
BCEL	**NSGA-II**	-	0.0344828	0.482759	AJHotDraw	-	**1**	0.142857
	SPEA2	0	-	0.448276		0	-	0
	PAES	0	0	-		**0.833333**	**1**	-
JBoss	**NSGA-II**	-	0	0	AJHSQLDB	-	**1**	0
	SPEA2	0	-	0		0	-	0
	PAES	0	0	-		**1**	**1**	-
JHotDraw	**NSGA-II**	-	0	0	Health Watcher	-	0	0
	SPEA2	0	-	0		0	-	0
	PAES	0	0	-		0	0	-
MyBatis	**NSGA-II**	-	**0.666667**	**0.722222**	Toll System	-	0	0
	SPEA2	0.206349	-	**0.62963**		0	-	0
	PAES	0.285714	0.403509	-		0	0	-
Experiment 4M								
System	**MOEA**	**NSGA-II**	**SPEA2**	**PAES**	**System**	**NSGA-II**	**SPEA2**	**PAES**
BCEL	**NSGA-II**	-	0	0.189189	AJHotDraw	-	**0.807692**	**0.952381**
	SPEA2	0.027027	-	0.216216		0.0506329	-	**0.916667**
	PAES	0	0	-		0	0.0128205	-
JBoss	**NSGA-II**	-	0	0	AJHSQLDB	-	0.307692	0
	SPEA2	0	-	0		**0.602778**	-	0
	PAES	0	0	-		**1**	**1**	-
JHotDraw	**NSGA-II**	-	0	**0.947368**	Health Watcher	-	0	0
	SPEA2	0	-	**0.947368**		0	-	0
	PAES	0	0	-		0	0	-
MyBatis	**NSGA-II**	-	0.0231884	**0.976436**	Toll System	-	0	0
	SPEA2	**0.894793**	-	**0.963181**		0	-	0
	PAES	0	0	-		0	0	-

Regarding to Experiment 2M, we observe significant difference in three systems: MyBatis, AJHotDraw and AJHSQLDB. For system MyBatis, the solutions achieved by NSGA-II dominate almost 67% of SPEA2 solutions and 72% of PAES solutions; the SPEA2 solutions dominate almost 63% of PAES solutions; and the PAES solutions do not dominate significantly the solutions of any other MOEA. For system AJHotDraw, the NSGA-II solutions dominate all SPEA2 solutions; PAES solutions dominate 83% of NSGA-II solutions and all SPEA2 solutions (100%); the SPEA2 solutions do not dominate any solutions of any other MOEA. The results for the system AJHSQLDB are similar to the AJHotDraw results, despite of the SPEA2 solutions also dominate all NSGA-II solutions. Figure 1 depicts the solutions on the search space for systems MyBatis and AJHSQLDB. In the case of MyBatis (Fig-

ure 1(a)) all solutions are in the same area, but NSGA-II solutions have better values. For AJHSQLDB (Figure 1(b)) it is clear PAES achieves the best solutions. These pictures corroborate the information about C indicator since they show the distribution of the solutions found by each MOEA.

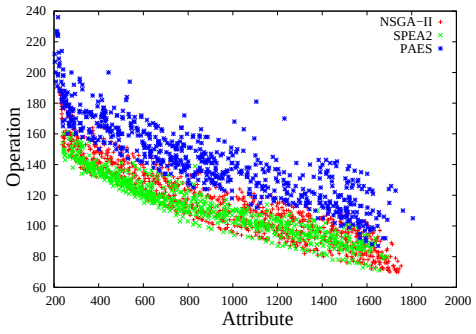


(a) MyBatis (b) AJHSQLDB

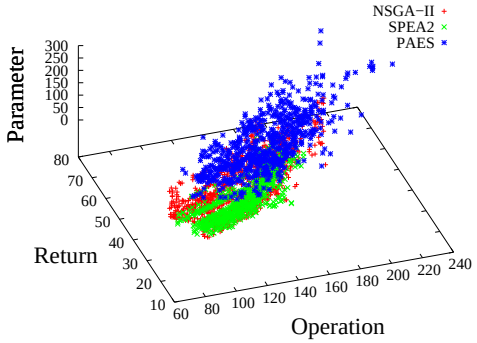
Figure 1: Search Space of Experiment 2M

Regarding to Experiment 4M, in addition to MyBatis, AJHotDraw and AJHSQLDB, we also observe significant difference in the system JHotDraw. In this case, NSGA-II and SPEA2 present similar results with their solutions dominating almost 95% of PAES solutions. Figure 2 presents the solutions on the search space of Experiment 4M for MyBatis and AJHSQLDB where the objectives are represented in two pictures. In the case of MyBatis (Figures 2(a) and 2(b)) the SPEA2 and NSGA-III solutions are close, although SPEA2 solutions are the best. PAES solutions are more spread on the search space. In Figures 2(c) and 2(d) we can observe that again the best solutions were achieved by PAES for AJHSQLDB.

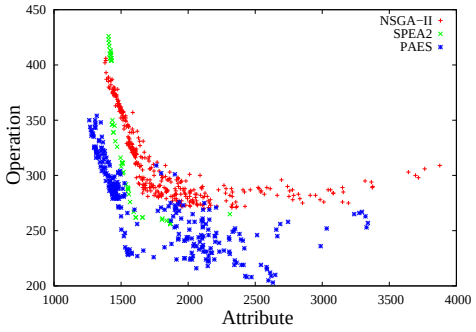
We can summarize the results related to the indicator C considering the number of sets covered by each MOEA. In the case of Experiment 2M, PAES and NSGA-II obtained the best results for the systems with significant differ-



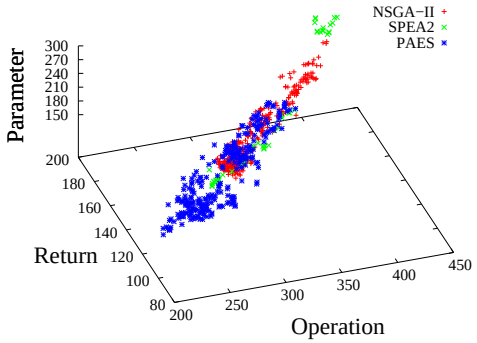
(a) MyBatis (A,O)



(b) MyBatis (O,R,P)



(c) AJHSQLDB (A,O)



(d) AJHSQLDB (O,R,P)

Figure 2: Search Space of Experiment 4M

ence among the MOEAs; each of them dominate four PF_{know} sets. Regarding to Experiment 4M, SPEA2 was the best since it covers six PF_{know} sets, followed by NSGA-II (four PF_{know} sets).

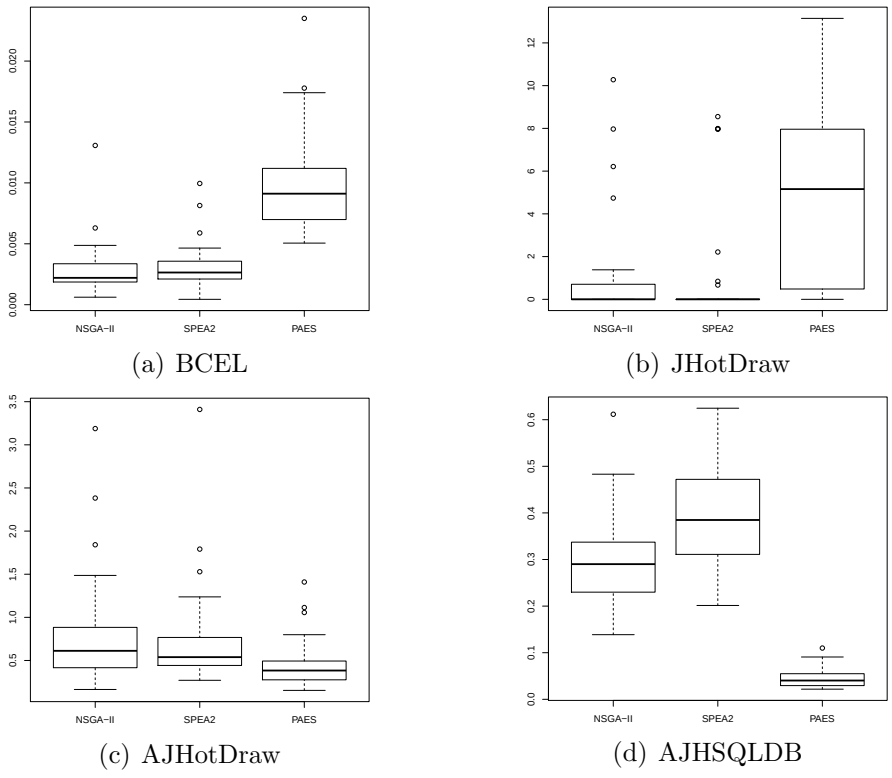
1.2. Generational Distance (GD) and Inverted Generational Distance (IGD)

Table 3 presents the results for GD and IGD indicators. These results are the average and the standard deviation of GD and IGD of the thirty PF_{approx} sets achieved by each MOEA. To verify the MOEAs that present significant difference, the statistical test of Friedmam [1] was used, with confidence level of 95%.

Table 3: Indicators GD and IGD

Experiment 2M							
Indicator	**System**	**NSGA-II**		**SPEA2**		**PAES**	
		Average	**Standard Deviation**	**Average**	**Standard Deviation**	**Average**	**Standard Deviation**
GD	BCEL	**0.002893**	0.002286	**0.003143**	0.001946	0.009942	0.004169
	JBoss	0	0	0	0	0	0
	JHotDraw	**1.133704**	2.599180	**1.206938**	2.793913	4.568271	4.152712
	MyBatis	0.013541	0.006589	0.016307	0.009722	0.014445	0.006950
	AJHotDraw	0.802435	0.667415	0.758403	0.619125	**0.454011**	0.286958
	AJHSQLDB	0.298636	0.099864	0.403558	0.120089	**0.045367**	0.021198
	Health Watcher	0	0	0	0	0.080800	0.312494
	Toll System	0	0	0	0	0	0
IGD	BCEL	**0.002638**	0.002015	**0.003039**	0.002078	0.011752	0.004477
	JBoss	0	0	0	0	0	0
	JHotDraw	**1.213778**	2.976485	**1.322521**	3.141205	5.069102	4.745460
	MyBatis	0.013799	0.006501	0.015677	0.008448	0.017287	0.009075
	AJHotDraw	0.854634	0.966271	0.771509	0.485032	**0.392923**	0.294614
	AJHSQLDB	0.309241	0.189022	0.422809	0.178879	**0.038639**	0.021898
	Health Watcher	0	0	0	0	0.058486	0.227020
	Toll System	0	0	0	0	0	0
Experiment 4M							
Indicator	**System**	**NSGA-II**		**SPEA2**		**PAES**	
		Average	**Standard Deviation**	**Average**	**Standard Deviation**	**Average**	**Standard Deviation**
GD	BCEL	**0.001984**	0.001211	**0.002085**	0.001144	0.011376	0.003297
	JBoss	0	0	0	0	0.069611	0.381275
	JHotDraw	0.303920	0.190225	0.408509	0.304008	0.487535	0.281376
	MyBatis	**0.006973**	0.001603	**0.007766**	0.001727	0.016785	0.002996
	AJHotDraw	**0.034252**	0.013653	0.044081	0.014161	0.052848	0.020588
	AJHSQLDB	0.058134	0.022112	0.086932	0.034211	**0.013284**	0.004251
	Health Watcher	0	0	0	0	0	0
	Toll System	0	0	0	0	0.081650	0.447214
IGD	BCEL	**0.001919**	0.001320	**0.001681**	0.001114	0.018785	0.005098
	JBoss	0	0	0	0	0.063221	0.346276
	JHotDraw	**0.249168**	0.258185	0.380334	0.456610	0.493869	0.330265
	MyBatis	**0.010104**	0.003140	**0.008608**	0.003607	0.024483	0.002652
	AJHotDraw	**0.036723**	0.016616	**0.046578**	0.018384	0.070992	0.026762
	AJHSQLDB	0.054444	0.020275	0.084633	0.035610	**0.016285**	0.004057
	Health Watcher	0	0	0	0	0	0
	Toll System	0	0	0	0	0.074536	0.408248

For Experiment 2M, according to the statistical test, there is significant difference in GD and IGD for four systems: BCEL, JHotDraw, AJHotDraw and AJHSQLDB. Figures 3 and 4 present the boxplots of the systems with difference among the MOEAs, where it is possible to determine which ones are the best. Considering both GD and IGD, NSGA-II and SPEA2 are equivalent and overcame PAES for systems BCEL and JHotDraw the MOEAs, whereas PAES is the best for systems AJHotDraw and AJHSQLDB.

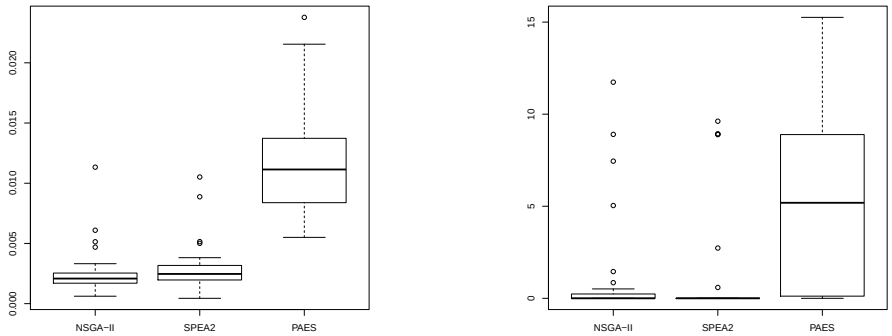


(a) BCEL (b) JHotDraw

(c) AJHotDraw (d) AJHSQLDB

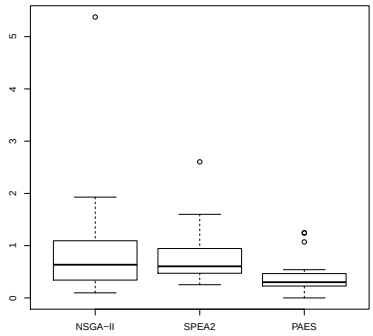
Figure 3: Boxplots for Indicator GD (Experiment 2M)

For Experiment 4M, the statistical test denotes significant difference for the results of the indicator GD for four systems: BCEL, MyBatis, AJHotDraw and AJHSQLDB. For the indicator IGD there is significant difference in

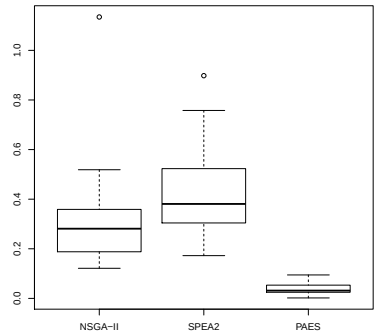


(a) BCEL

(b) JHotDraw



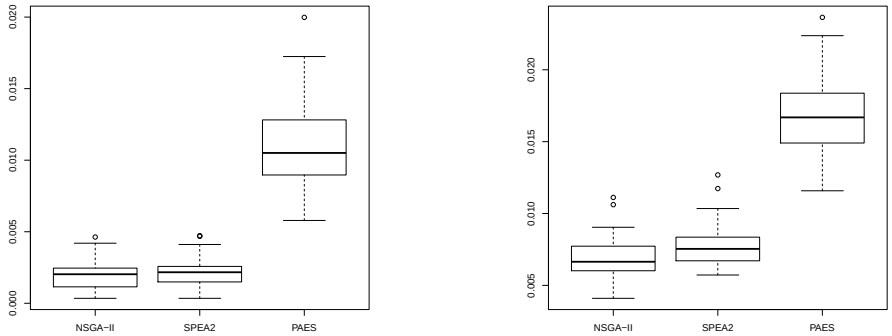
(c) AJHotDraw



(d) AJHSQLDB

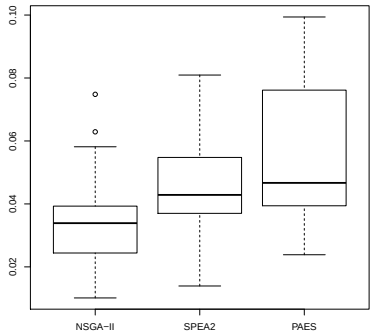
Figure 4: Boxplots for Indicator IGD (Experiment 2M)

five systems: the same four systems with difference in indicator GD (BCEL, MyBatis, AJHotDraw, AJHSQLDB) and JHotDraw. Figures 5 and 6 present the boxplots of the systems with difference among the MOEAs. Regarding to indicator GD, NSGA-II and SPEA2 are equivalent and better than PAES for systems BCEL and MyBatis, NSGA-II is the best for AJHotDraw, and PAES is the best for AJHSQLDB. Regarding to indicator IGD, for the systems BCEL, JHotDraw, MyBatis and AJHotDraw, the MOEAs NSGA-II and SPEA2 are better, and for the system AJHSQLDB, PAES is the best.

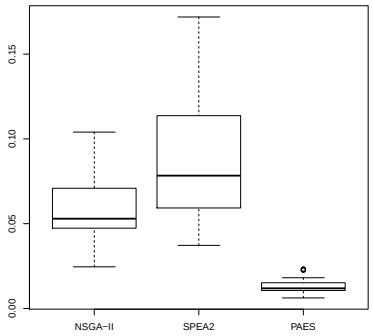


(a) BCEL

(b) MyBatis



(c) AJHotDraw



(d) AJHSQLDB

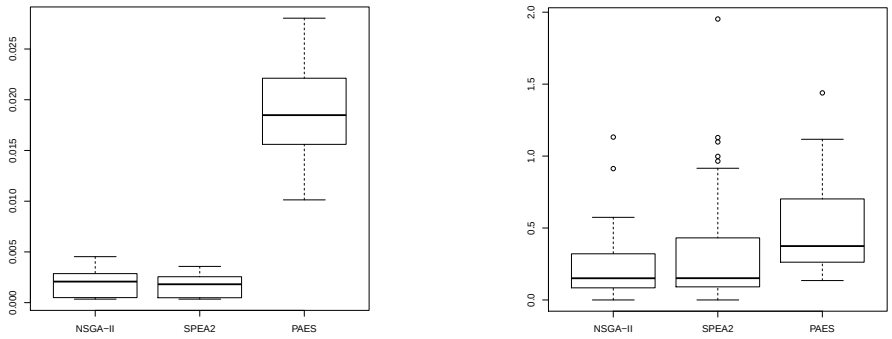
Figure 5: Boxplots for Indicator GD (Experiment 4M)

1.3. Euclidean Distance from the Ideal Solution

Table 4 presents the ED results, regarding to the closest solution to the ideal solution achieved by the MOEAs for each system. The second column presents the cost of the ideal solution, obtained from PF_{true}. The other columns present the distances and the costs of the closest solutions to the ideal solution.

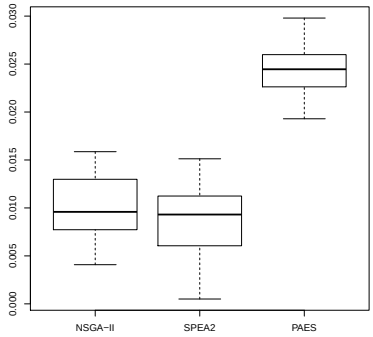
As mentioned before, in both experiments all algorithms achieve only one solution for JBoss, Health Watcher and Toll System. This solution has the best values for each objective, so it represents the ideal solution.

Regarding to Experiment 2M, all the MOEAs achieve the same closest

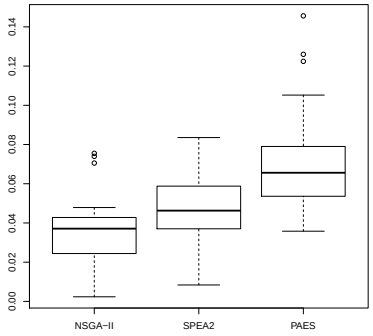


(a) BCEL

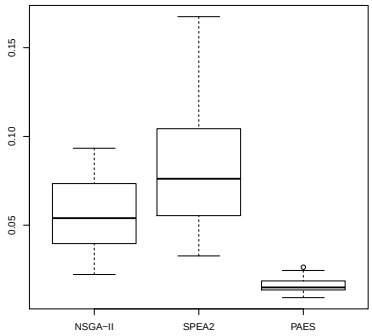
(b) JHotDraw



(c) MyBatis



(d) AJHotDraw



(e) AJHSQLDB

Figure 6: Boxplots for Indicator IGD (Experiment 4M)

solution to ideal solution for JHotDraw. For BCEL, NSGA-II and SPEA2 achieve the best solution. PAES achieves the closest solution for three systems: MyBatis, AJHotDraw and AJHSQLDB.

Table 4: Indicator ED

Experiment 2M							
System	**Ideal Solution Cost**	**NSGA-II**		**SPEA2**		**PAES**	
		Closest ED	**Solution Cost**	**Near ED**	**Solution Cost**	**Closest ED**	**Solution Cost**
BCEL	(45,24)	**21.954498**	(56,43)	**21.954498**	(56,43)	22.203603	(58,42)
JBoss	(10,6)	**0**	(10,6)	**0**	(10,6)	**0**	(10,6)
JHotDraw	(27,9)	**1.414214**	(28,10)	**1.414214**	(28,10)	**1.414214**	(28,10)
MyBatis	(174,61)	102.800778	(236,143)	101.872469	(227,148)	**80.622577**	(214,131)
AJHotDraw	(40,13)	7.810250	(46,18)	12.806248	(48,23)	**5.656854**	(44,17)
AJHSQLDB	(1121,197)	347.131099	(1381,427)	387.191167	(1420,443)	**109.731490**	(1156,301)
Health Watcher	(9,2)	**0**	(9,2)	**0**	(9,2)	**0**	(9,2)
Toll System	(0,0)	**0**	(0,0)	**0**	(0,0)	**0**	(0,0)

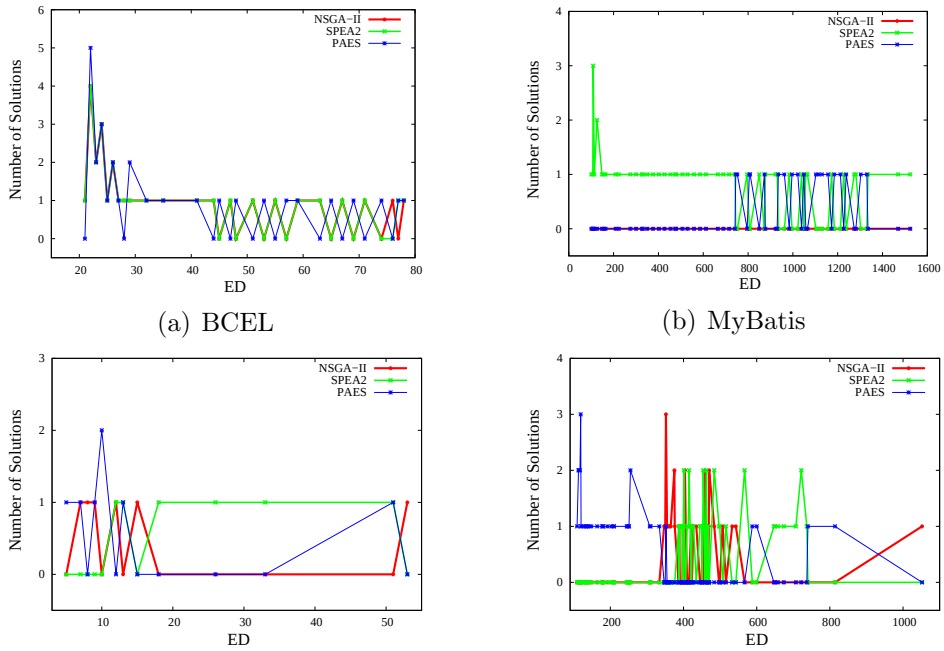
Experiment 4M							
System	**Ideal Solution Cost**	**NSGA-II**		**SPEA2**		**PAES**	
		Closest ED	**Solution Cost**	**Closest ED**	**Solution Cost**	**Near ED**	**Solution Cost**
BCEL	(45,24, 0,96)	**32.186954**	(64,39, 15,111)	**32.186954**	(64,39, 15,111)	**32.186954**	(64,39, 15,111)
JBoss	(10,6,2,9)	**0**	(10,6,2,9)	**0**	(10,6,2,9)	**0**	(10,6,2,9)
JHotDraw	(27,10, 1,12)	**5.000000**	(29,11, 3,16)	**5.000000**	(29,11, 3,16)	7.000000	(30,12, 1,18)
MyBatis	(203,70, 13,47)	153.710767	(248,151, 33,168)	**141.417821**	(250,143, 32,157)	185.237685	(265,172, 49,184)
AJHotDraw	(39,12, 0,18)	**17.058722**	(46,16, 1,33)	18.493242	(46,18, 1,34)	18.841444	(46,19, 1,34)
AJHSQLDB	(1263,203, 91,138)	271.383492	(1391,387, 163,273)	253.288373	(1432,346, 142,250)	**164.872678**	(1314,316, 138,236)
Health Watcher	(9,2,0,1)	**0**	(9,2,0,1)	**0**	(9,2,0,1)	**0**	(9,2,0,1)
Toll System	(0,0,0,0)	**0**	(0,0,0,0)	**0**	(0,0,0,0)	**0**	(0,0,0,0)

Regarding to Experiment 4M, the three algorithms achieve the same closest solution to ideal solution for BCEL. For JHotDraw, NSGA-II and SPEA2 achieve the closest solution. In addition, SPEA2, NSGA-II and PAES achieve the closest solution for one system: MyBatis, AJHotDraw and AJHSQLDB, respectively.

For further analysis on the behavior of MOEAs, using the indicator ED, Figures 7 and 8 present graphs showing the number of solutions for ED. In these pictures, it is possible to verify which MOEA has the greatest concentration of closest solutions to the ideal solution.

Regarding to Experiment 2M, PAES achieves more solutions closest to the ideal solution for BCEL (Figure 7(a)), AJHotDraw (Figure 7(c)) and AJHSQLDB (Figure 7(d)). For MyBatis (Figure 7(b)) SPEA2 finds more closest solutions. So, considering all systems, in addition to achieve the

closest solution for three systems, we can state that PAES can achieve more solutions with lower ED than the other two MOEAs.

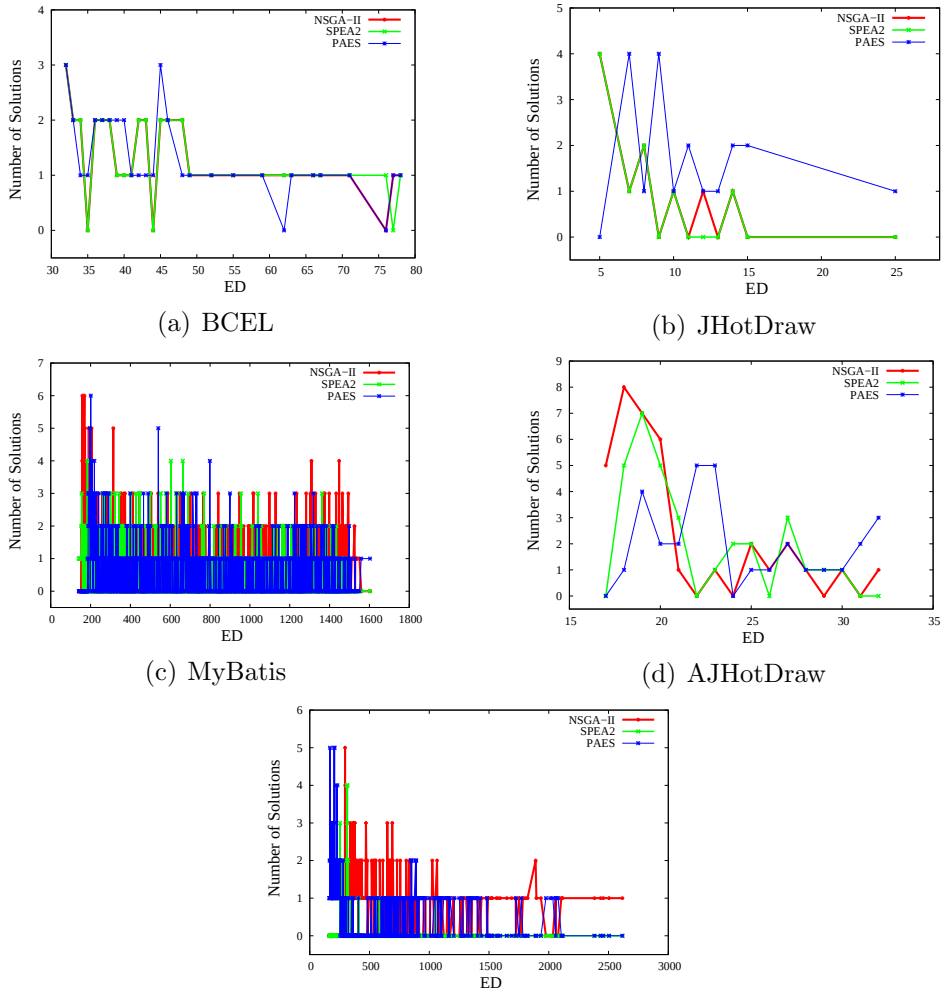


(a) BCEL (b) MyBatis

(c) AJHotDraw (d) AJHSQLDB

Figure 7: Number of Solutions X Indicator ED (Experiment 2M)

Differently from Experiment 2M where PAES finds more solution with lower ED from the ideal solution, in Experiment 4M, NSGA-II achieves more closest solutions to the ideal solution to JHotDraw (Figure 8(b)), AJHotDraw (Figure 8(d)) and MyBatis (Figure 8(c)). PAES achieves the best results (solutions with lower ED) only for AJHSQLDB (Figure 8(e)). From these results we can observe that, in the presence of many objectives (4), NSGA-II can find a greater number of closest solutions to the ideal solution.



(a) BCEL

(b) JHotDraw

(c) MyBatis

(d) AJHotDraw

(e) AJHSQLDB

Figure 8: Number of Solutions X Indicator ED (Experiment 4M)

1.4. Discussion

After presenting the results of the quality indicators, we can analyze these results in a general way. All MOEAs have the same behavior for systems JBoss, Health Watcher and Toll System. It is possible to verify which MOEA has the best behavior for the other systems in Table 5, that presents the

MOEAs with best results for each indicator in each system.

Table 5: Better MOEAs by Quality Indicator

System	Experiment 2M				Experiment 4M			
	C	GD	IGD	ED	C	GD	IGD	ED
BCEL		NSGA-II SPEA2	NSGA-II SPEA2	NSGA-II SPEA2		NSGA-II SPEA2	NSGA-II SPEA2	NSGA-II SPEA2
JHotDraw		NSGA-II SPEA2	NSGA-II SPEA2	NSGA-II	NSGA-II	SPEA2	NSGA-II SPEA2	NSGA-II SPEA2
MyBatis	NSGA-II			SPEA2	SPEA2	NSGA-II SPEA2	NSGA-II SPEA2	SPEA2
AJHotDraw	PAES	PAES	PAES	PAES	NSGA-II	NSGA-II	NSGA-II SPEA2	NSGA-II
AJHSQLDB	PAES	PAES	PAES	PAES	PAES	PAES	PAES	PAES

From the obtained results, presented in Table 5, we can state that:

- Considering the indicators GD and IGD, in general, NSGA-II and SPEA2 have the same performance and they are better than PAES for almost all systems. The single case where PAES is the best for these two indicators is for AJHSQLDB.

- Considering the indicator C, PAES and NSGA-II are better than SPEA2 in Experiment 2M and NSGA-II is the best in Experiment 4M.

- Regarding the indicator ED, all MOEAs achieve solutions with lower ED for some system.

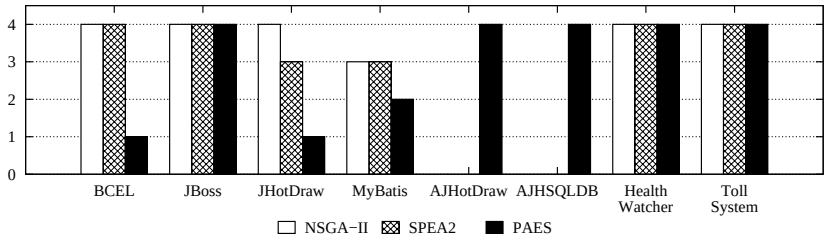
- NSGA-II finds a greater number of closest solutions to the ideal solution in Experiment 4M.

- PAES finds a greater number of closest solutions to the ideal solution in Experiment 2M.

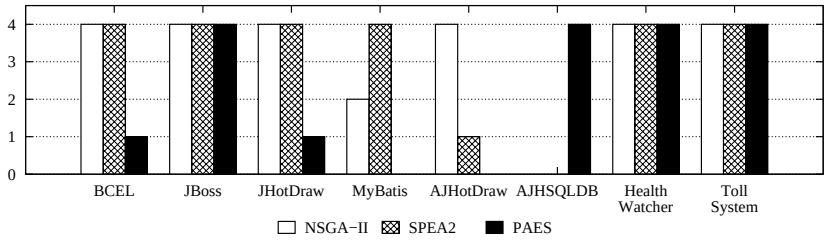
- PAES is the best in all quality indicators for AJHotDraw and AJHSQLDB in Experiment 2M. This algorithm is also the best for AJHSQLDB in Experiment 4M. These two systems has the greatest num-

bers of modules (classes and aspects) and dependencies. So, PAES has better performance to solve the integration and test order problem for more complex systems.

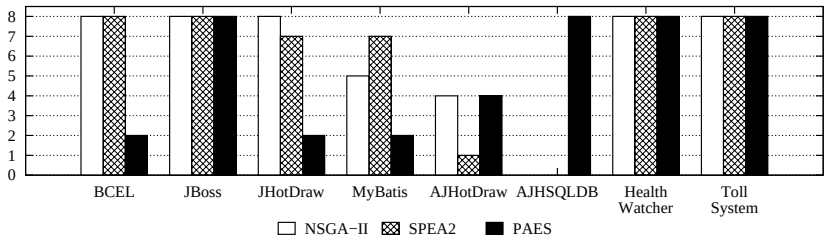
Figure 9 presents the graphs that summarize the performance of each MOEA for each quality indicator. In these graphs each bar represents the number of times that a MOEA overcomes other MOEA during the experiments.



(a) Experiment 2M



(b) Experiment 4M



(c) Experiments 2M and 4M

Figure 9: Number of Best Results X Indicator (C, GD, IGD and ED)

From this information we can highlight the following observations:

- PAES achieves the best results only for AO systems (AJHotDraw and AJHSQLDB) in Experiment 2M (Figure 9(a)).

- In the presence of four objectives (Experiment 4M), PAES does not achieve so good results as NSGA-II and SPEA2 achieve (Figure 9(b)).

- For OO systems, NSGA-II and SPEA2 have the same performance with 2 objectives and SPEA2 was the best with 4 objectives (Figure 9).

The three MOEAs are effective for treating the integration and test order problem using two or four objectives, since they achieve similar results despite using different evolution strategies.

It is possible to affirm that, in general, NSGA-II and SPEA2 have a slight better convergence than PAES, as GD and the cardinality of the PF_{true} indicate. For two objectives, SPEA2 presents good distribution of solutions in the region near the ideal solution, as the ED indicator shows. Often, decision makers prefer solutions near to the ideal solution. So, in this case the SPEA2 should be chosen. The same occurs for NSGA-II with respect to four objectives.

In the context of our empirical evaluation, we can recommend the use of PAES to solve CAITO problem (AO systems) in the presence of two objectives and with complex systems. From all results from the quality indicators, NSGA-II seems to be more appropriate in general cases because: (i) it has good convergence (GD and IGD indicators), (ii) it finds a set of solutions that cover the solutions found by two other MOEAs (C indicator), (iii) it achieve solutions closer to the ideal solution, and (iv) its good results do not change in the presence of four objectives.

Acknowledgments

We would like to thank Edison Klafke Fillus for his contribution in the experiments. This work was supported by CNPq (Productivity grant numbers: 303761/2009-1 and 306608/2009-0 and Universal Project grant numbers: 472510/2010-0 and 481397/2009-4) and by CAPES/Reuni.

References

[1] S. Gárcia, D. Molina, M. Lozano, and F. Herrera. A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 Special Session on Real Parameter Optimization. *Journal of Heuristics*, 15(6):617–644, 2009.