TABLE I: Comparison of ComSat and MonoMod for the CF-EVRP over a set of generated problem instances. Instances are sorted by the parameters N-V-J (nodes, vehicles, jobs), value of edge reduction, and time horizon. For each resulting class, five instances are evaluated and the number of feasible and unfeasible ones is reported, together with the average solving time (in seconds) for that specific class. The average generation time (in seconds) is also reported. The symbol "-" means that no instance for that category was either feasible or feasible, depending on where the symbol appears.

	N-V-J	15-3-5														
T	Edge Red.	0					25					50				
		Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)
20	ComSat	4/5	4.96	1/5	9.2	0.36	3/5	6.58	2/5	8.9	0.33	1/5	2.35	3/5	3.46	0.08
	MonoMod	4/5	11.05	1/5	1.16	22.85	3/5	8.63	2/5	4.59	22.94	1/5	10.98	4/5	4.24	22.38
25	ComSat	4/5	5.83	1/5	17.55	0.33	3/5	6.91	2/5	14.14	0.35	3/5	3.99	2/5	3.71	0.12
	MonoMod	4/5	20.79	1/5	2.18	41.98	3/5	24.02	2/5	33.95	40.88	3/5	39.42	2/5	21.7	43.38
30	ComSat	4/5	6.06	1/5	21.69	0.36	3/5	7.25	2/5	16.49	0.34	3/5	3.41	2/5	4.25	0.1
	MonoMod	4/5	913.15	0/5	-	64.74	3/5	66.87	1/5	845.16	70.68	3/5	65.85	1/5	5.06	65.5
40	ComSat	4/5	6.23	1/5	22.97	0.32	3/5	8.35	2/5	19.96	0.36	3/5	4.05	2/5	4.02	0.09
	MonoMod	4/5	162.33	0/5	-	94.36	3/5	178.07	0/5	-	98.15	3/5	240.75	1/5	901.82	102.1
	N-V-J	25-4-7														
T	Edge Red.	0					25					50				
		Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)
20	ComSat	2/5	582.1	3/5	35.13	1.24	0/5	-	5/5	38.6	1.09	0/5	-	5/5	17.92	0.84
	MonoMod	2/5	42.09	3/5	26.91	75.04	0/5	-	5/5	23.93	67.66	0/5	-	5/5	24.38	70.36
25	ComSat	5/5	44.61	0/5	-	1.2	4/5	63.11	1/5	75.48	1.17	1/5	20.99	4/5	227.08	0.92
	MonoMod	5/5	182.52	0/5	-	125.76	4/5	163.15	1/5	333.66	109.85	1/5	110.78	4/5	113.66	99.82
30	ComSat	5/5	41.39	0/5	-	1.23	4/5	177.0	1/5	84.86	1.13	1/5	34.72	4/5	303.96	0.87
	MonoMod	5/5	1160.64	0/5	-	195.9	4/5	995.07	0/5	-	169.96	1/5	479.7	1/5	1197.42	191.34
40	ComSat	5/5	66.63	0/5	-	1.11	4/5	139.65	1/5	101.52	1.14	1/5	30.29	4/5	351.01	0.85
	MonoMod	4/5	2312.39	0/5	-	285.38	4/5	3147.12	0/5	-	275.71	1/5	318.09	0/5	-	293.11
	N-V-J	35-6-8														
T	Edge Red.	0					25					50				
		Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)	Feas	Av.(sec)	Unfeas	Av.(sec)	Gen.(sec)
20	ComSat	1/5	16.15	4/5	22.26	2.46	0/5	-	5/5	26.99	2.13	0/5	-	5/5	32.31	1.77
	MonoMod	1/5	80.24	4/5	97.01	167.36	0/5	-	5/5	134.58	132.39	0/5	-	5/5	82.83	131.8
25	ComSat	4/5	177.73	1/5	22.28	2.32	4/5	81.2	1/5	30.37	2.32	4/5	216.74	1/5	33.99	1.56
	MonoMod	4/5	896.66	1/5	48.93	220.63	4/5	636.35	1/5	32.75	203.98	4/5	644.79	1/5	31.24	201.95
30	ComSat	4/5	113.54	1/5	24.32	2.27	4/5	188.78	1/5	27.8	2.15	4/5	103.26	1/5	34.29	1.9
	MonoMod	4/5	2268.8	1/5	138.78	425.56	4/5	937.94	1/5	77.62	376.42	4/5	890.02	1/5	125.29	347.38
40	ComSat	4/5	216.67	1/5	21.65	2.29	4/5	116.29	1/5	29.81	2.45	4/5	167.9	1/5	32.84	1.6
	MonoMod	4/5	2689.35	0/5	-	586.41	4/5	1167.67	1/5	913.9	597.01	4/5	1261.64	1/5	210.42	492.4

V. Computational Analysis

In order to compare MonoMod and ComSat we generated a set of problem instances. The parameters we used are the number of nodes, vehicles, and jobs (grouped in an index called N-J-V), as well as the time horizon and the value called 'edge reduction', which indicates the connectivity of the graph (the higher the value, the fewer edges). For each combination of these parameters, five different problems are randomly generated.

For the analysis we used Z3 4.8.9. The time limit for MonoMod is set to 10800 seconds (three hours); the model generation time is measured separately, since it is implementation-dependent and can be dealt with using more efficient formulations, as discussed in our previous work [23]. As for ComSat, we only computed *ten* paths for each pair of customers. We also set an upper bound of *ten* to the number of iterations between the *Routing Problem* and the *Assignment Problem*. Also, the generation time refers to the time taken to generate the paths between each pair of customers. All the experiments were performed on an *Intel Core i7 6700K, 4.0 GHZ, 32GB RAM* running *Ubuntu-18.04 LTS*.

Though Z3 allows for optimization of the objective function, the size of the problems evaluated with MonoMod is such that no optimum is expected to be found in any reasonable time. Therefore Z3 is set to find feasible, sub-optimal solutions [1].

Table I summarizes the results of the computational analysis. The generation time for ComSat is actually negligible, whereas for MonoMod it increases with N-J-V (nodes, vehicles, and jobs), and time horizon and it decreases with the *edge reduction*, presumably because fewer edges means fewer constraints to declare. By comparing the number of solved instances in each category, whether they turned out to be *sat* or *unsat*, it is possible to notice that there is a number of instances that were determined *unknown* by ComSat but were declared *unsat* by MonoMod. On the other hand, both methods usually agree on the feasibility of the *sat* instances, except for some cases with high values of N-J-V and time horizon where MonoMod run out of time. The reason for the *unknown* responses lies in the termination criterion we set up for the algorithm. By running early experiments we noticed that the algorithm was rather slow in evaluating unsatisfiable problem instances; therefore we decided to limit the number of iterations between the *Routing Problem* and the *Assignment Problem*. On average, ComSat is between 10 and 100 times faster than the at solving instances that are intrinsically *sat*.