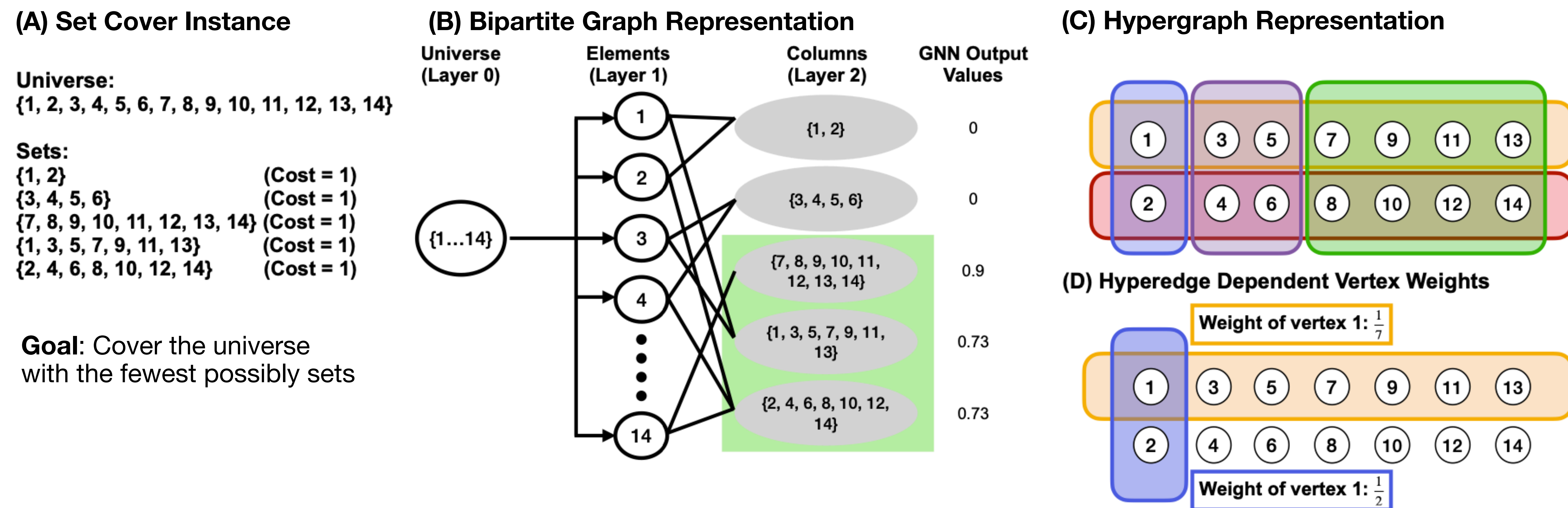


Problem Statements:

Can we use graph machine learning to speed up the set cover problem?
 Can we do so without loss in solution quality?
 Can we do so in a solver agnostic manner?

Datasets:

We generate 4 instance types, each spanning a range of characteristics. Here, m is the number of rows, n the number of columns, and d is the density of the instance.



Instance	m	n	d	Cost
Type 1	100-400	100-1000	0.22-0.29	Uniform [100-200]
Type 2	100-300	100-500	0.16-0.28	Equal
Type 3	200-350	300-350	0.13-0.18	Poisson ($\lambda = 20$)
Type 4	200-250	1000-3000	0.04-0.05	Poisson ($\lambda = 20$)
Type 5 (OR Library)	100-500	1000-10000	0.02-0.2	Uniform [1-100]

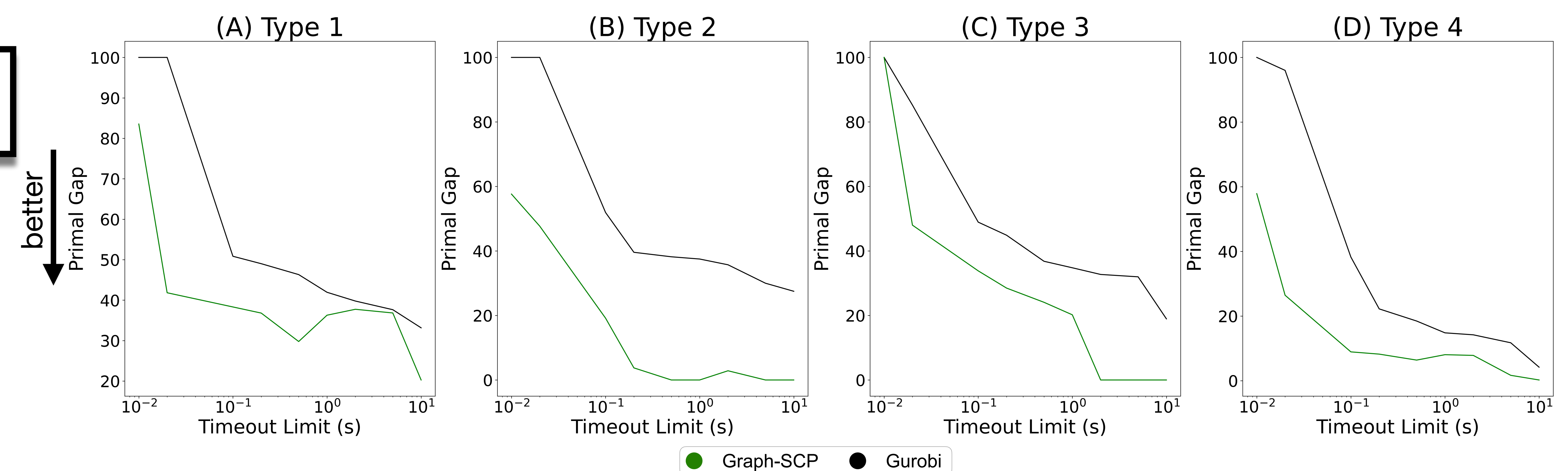
Results:

Comparison of Gurobi and SCIP with and without Graph-SCP. Graph-SCP runs significantly faster while maintaining solution quality. Results are averaged across 30 instances for each instance type.

Instance Type	Speedup Factor Gurobi	Mean Runtime Gurobi (s)	Speedup Factor SCIP	Mean Runtime SCIP (s)
Type 1	8.19 ± 3.82	5.61 ± 0.53	11.61 ± 3.02	21.87 ± 3.07
Type 2	70.94 ± 16.11	17.2 ± 4.87	491.53 ± 134.74	168.15 ± 48.07
Type 3	12.26 ± 4.2	35.57 ± 4.11	18.48 ± 5.05	313.5 ± 46.26
Type 4	3.65 ± 0.24	40.74 ± 7.02	11.96 ± 5.28	751.16 ± 176.33

Results:

Graph-SCP aids the traditional solvers in finding better incumbent solutions faster.



How?

- Graph-SCP uses a bipartite and hypergraph representation of a set cover instance
- It computes features for each variable from each representation, like degree, PageRank and an edge-weight dependent variation of PageRank for the hyper graph
- A GNN is trained offline to predict the subset of variables that lie in the solution in a supervised manner with an unsupervised component reducing the LP relaxation of the input set cover instance
- The top k^{th} percent of variables are selected as the input to a traditional solver. If the solution found is higher than a user-defined objective criteria, the percentile is reduced to pass in a larger subset of variables

