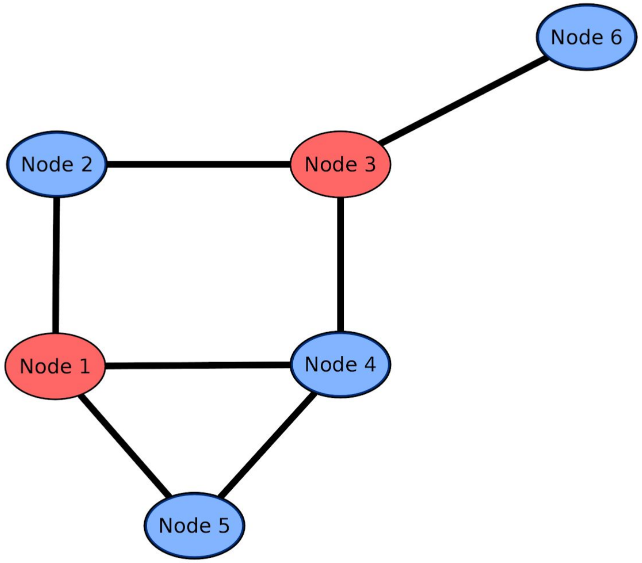
Minimum Dominating Set: Empirical Comparisons of Graph Theoretical Algorithms vs Mathematical Programming Formulations

Stephen K. Grady[1], Austin R. Wyer[2], Hesam Shams[3], Faisal N. Abu-Khzam[4], Charles A. Phillips[2] and Michael A. Langston[1,2]

1. Genome Science and Technology, University of Tennessee
2. Electrical Engineering and Computer Science, University of Tennessee
3. Industrial Engineering, University of Tennessee
4. Department of Computer Science and Mathematics, Lebanese American University

Minimum Dominating Set (MDS)

- Given a graph G=(V,E) a dominating set (DS) is a set D $\subseteq$ V s.t. for all vertices v $\subseteq$ V, v is either in D or adjacent to at least one vertex in D.
- An MDS is a DS of minimum cardinality.
- The size of an MDS is known as the dominating number.



Nodes 1 and 3 in red form a minimum dominating set.

Determining Critical, Redundant and Intermittent Vertices

- Must solve MDS N times.
- To determine critical: for each vertex v in MDS, exclude from being in solution and resolve. If MDS size increases, v is critical.
- To determine redundant: for each vertex v not in MDS, force in solution and resolve. I MDS size increases, v is redundant.
- Intermittent are those vertices that are neither.

Motivation

- Determining critical, intermittent, and redundant vertices found to be important for studying biological networks.
- Critical vertices are in every solution of MDS.
- Intermittent vertices are in one or more, but not all solutions of MDS.
- Redundant vertices are in no solution of MDS.

Direct Search-Tree Algorithm

- Preprocess all degree one vertices.
- Backtracking algorithm of Fomin et al.
- Two reduction rules:
  - Closed-neighborhood subset rule
  - Unique neighbor rule.

- Vertex selection
  - Vertex with most uncovered neighbors.

ILP Formulation of MDS

- For each vertex $v_i \in V$ create a binary variable x_i
- For each closed neighborhood $N[v_i]$ create the constraint: $\sum_{j \in N[v]} x_j \geq 1$

$$\min \sum_{i \in V} x_i$$

$$\sum_{j \in N[v]} x_j \geq 1 \ \forall i \in V$$

$$X_i \in \{0,1\} \ \forall i \in V$$

Direct Graph Method vs. ILP

An Experimental study

Testbed

- Erdos-Renyi random graphs with 100 vertices and varying density.
- We use a variety of graphs from real-world datasets.
- Useful to test on different graph topologies.
- Transcriptomic network, protein-protein interaction network, social networks, food-web network, road network and epidemiology network.

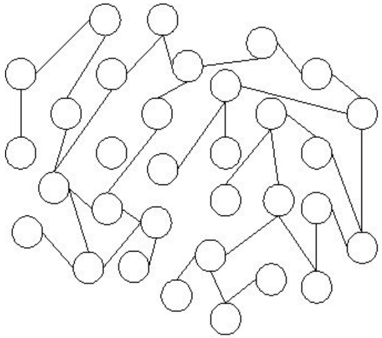
Testing Environment

- All experiments were performed on a dell laptop with an intel core 15-52000U @ 2.20GHz x4 with 8GiB of memory.
- ILP formulation implemented using GUROBI 7.5.2

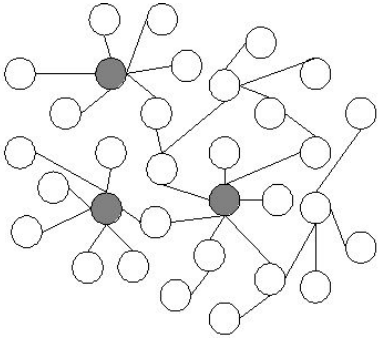
DS-Min vs. ILP: Random Graphs

Instances Graph Algorithm Outperforms

- Graphs where reduction rules can take advantage.
  - Power-law degree distributions.

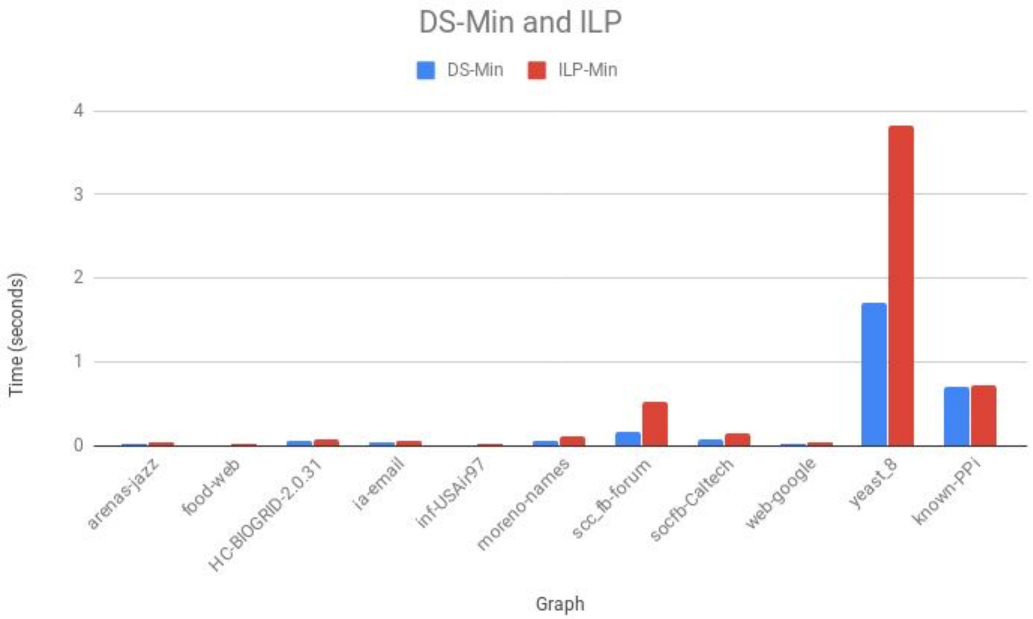


(a) Random network



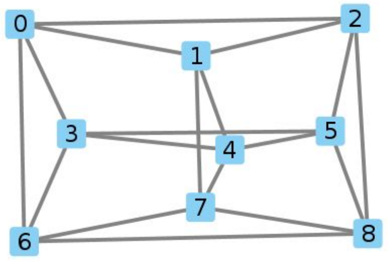
(b) Scale-free network

DS-Min vs. ILP: Power-Law Graphs

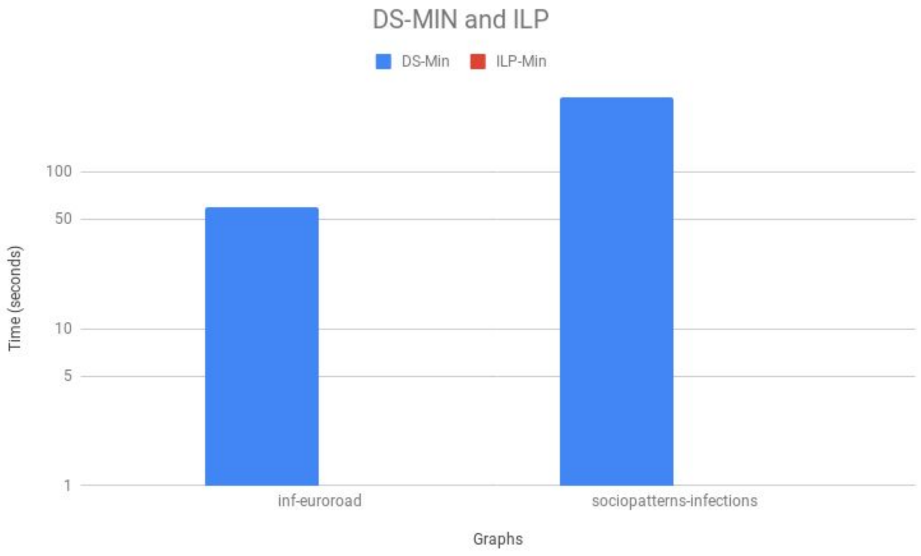


Instances ILP Outperforms

- Sparse highly regular graphs
  - Road networks, certain epidemiological networks

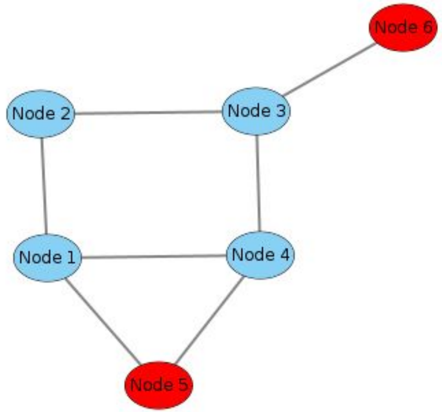


DS-Min vs. ILP: Highly regular Graphs

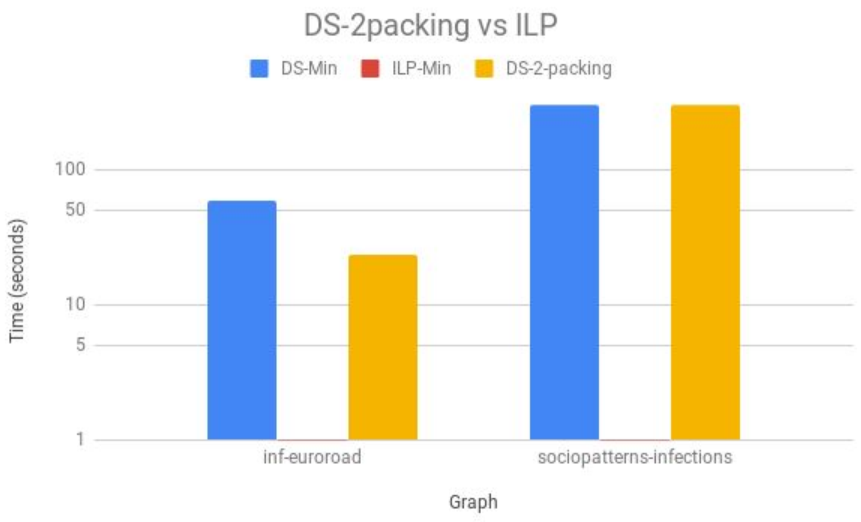


Lower Bounds: 2-packing

- A set of vertices P, s.t. the intersection of the closed neighborhoods of any two vertices u,v ∈ P is empty.
- Can be thought of an independent set on 2-hop graph.
- 2-packing (2*p*) is a lower bounds to MDS.

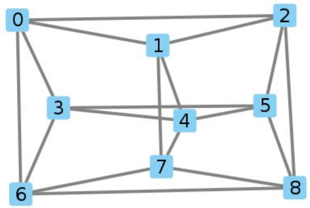


Lower Bounds: 2-packing

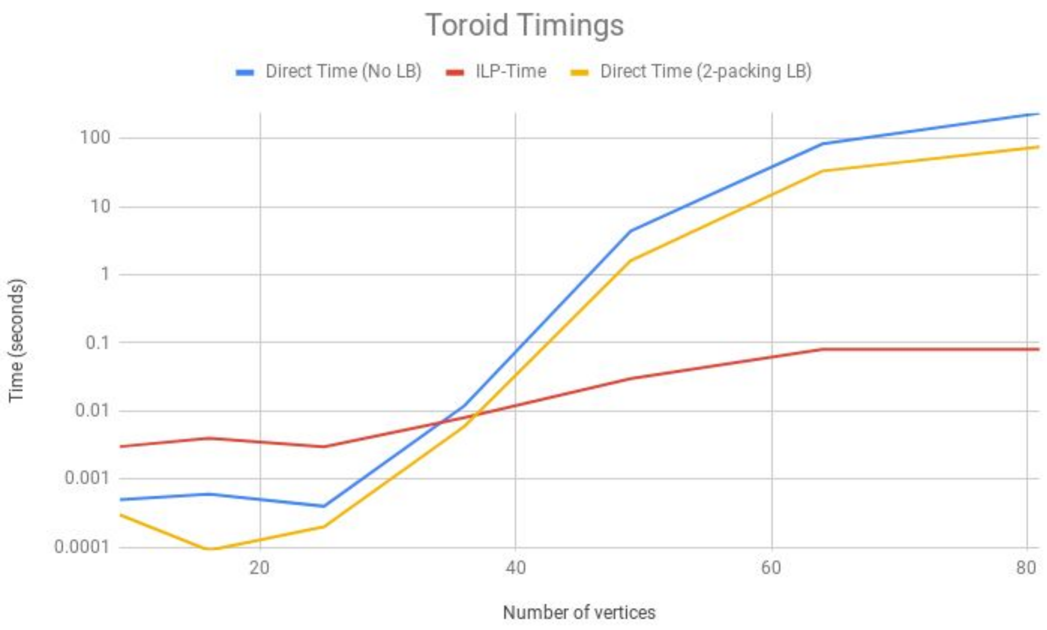


Toroidal Graphs

- ILP performs significantly better on highly regular graphs.
- Regular graph with degree 4.
- Constructed as intentionally hard case for graph theoretic algorithm.
  - Reduction rules do not apply at start.



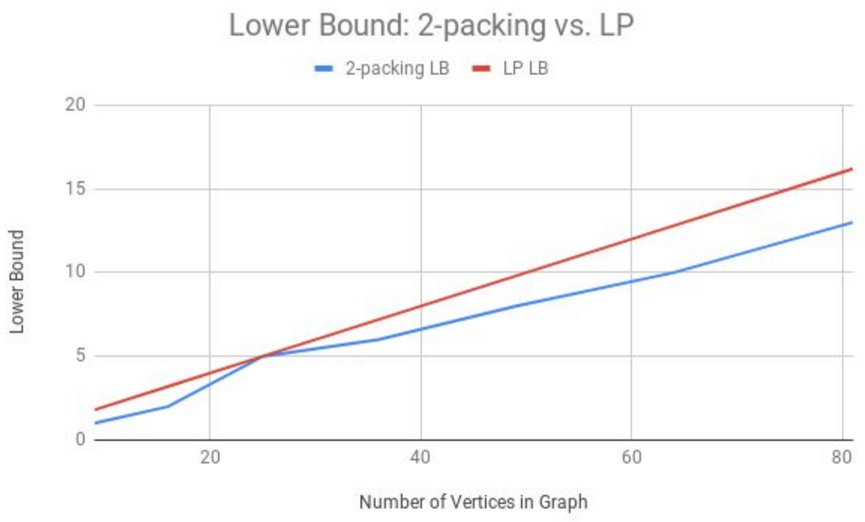
Toroidal Graphs: Timings



Lower Bounds: LP Relaxation

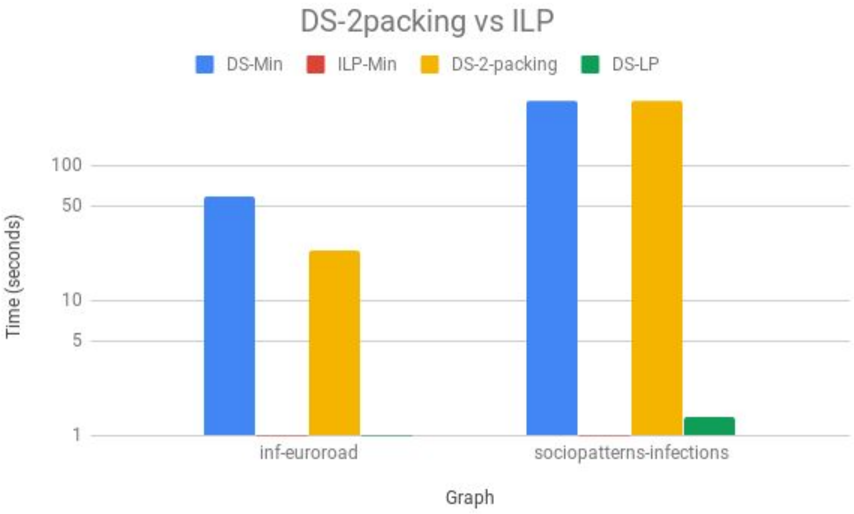
- The LP relaxation of MDS is also a lower bounds for MDS.
- Unfortunately, $2p \leq$ LP-MDS $\leq$ MDS.
- The LP relaxation is generally a tighter bound.
- Added advantage; if LP relaxation is integral, optimum solution is found.

Lower Bounds: LP vs 2-packing



Toroidal graphs: LP Relaxation

Lower Bounds: LP Relaxation

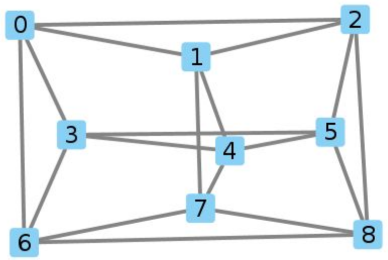


The Story so Far

- Direct method generally better for power-law graphs.
  - Can take advantage of reduction rules.
- ILP generally better for regular graphs.
- Studying “extreme” instances can offer insight into algorithmic behavior.
- Studying the techniques used by one implementation can aid in the design of the other.

Future Work: Symmetry

- ILP is able to exploit symmetry.
- Allows branching on orbits. (a set of vertices that can be swapped preserving isomorphism)



Future Work: Choosing Lower Bounds

- Exploiting knowledge of graph topology to choose lower bound during search.
- When is 2-packing sufficient?
- When is LP relaxation worth the time cost?
- Can we employ machine learning methods to determine this choice?

Future Work: When to Check for Connected Components

- MDS deconstructs a graph when branching.
- Graph can become disconnected during search.
- Each connected component can be solved independently.

Open Question

- When to use kernalization via MDS's fixed parameter tractable (FPT) dual dominated set (also known as nonblocker)?
- Analogous to vertex cover for clique.

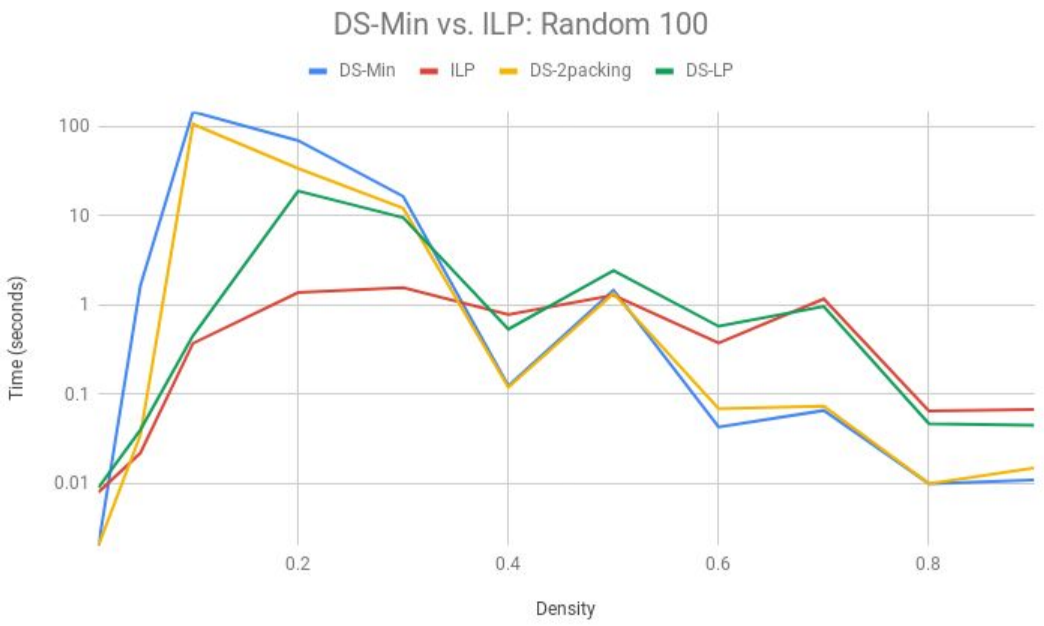
Questions?

I would like to thank:

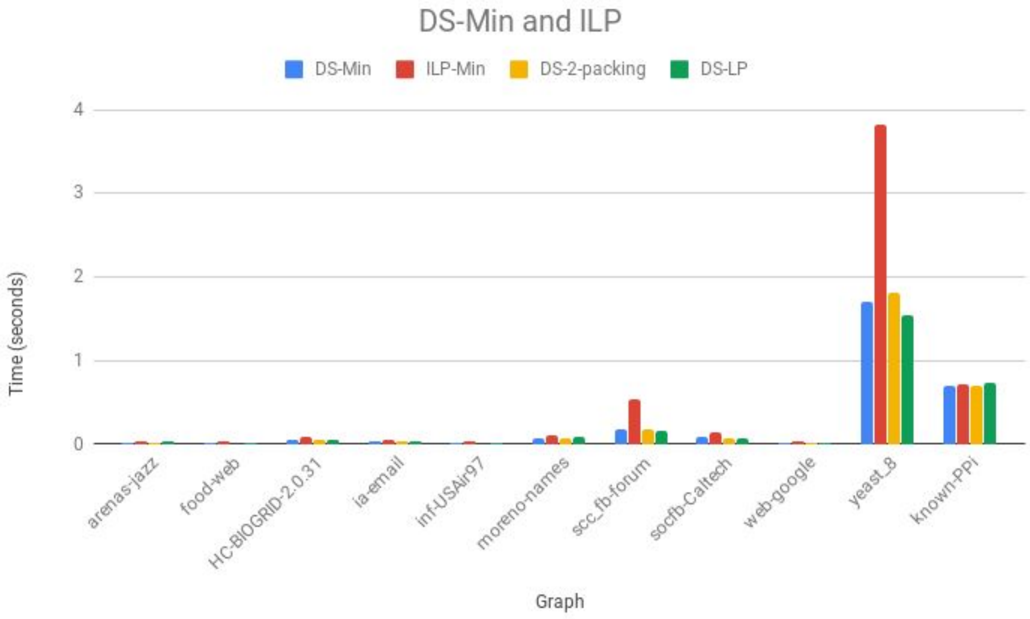
- My advisor, my coauthors and colleagues.
- AFRL, EPA, and NIH
- Program for Genome science and Technology.

References

1. Nacher, J. C., and T. Akutsu. “Analysis of Critical and Redundant Nodes in Controlling Directed and Undirected Complex Networks Using Dominating Sets.” *Journal of Complex Networks*, vol. 2, no. 4, 2014, pp. 394–412., doi:10.1093/comnet/cnu029.
2. Fomin, Fedor V., et al. “Measure and Conquer: Domination – A Case Study.” *Automata, Languages and Programming Lecture Notes in Computer Science*, 2005, pp. 191–203., doi:10.1007/11523468_16.
3. Rubalcaba, R., and M. Walsh. “Minimum Fractional Dominating Functions and Maximum Fractional Packing Functions.” *Discrete Mathematics*, vol. 309, no. 10, 2009, pp. 3280–3291., doi:10.1016/j.disc.2008.09.049.
4. CC BY-SA 3.0, from Wikipedia https://en.wikipedia.org/wiki/Scale-free_network



DS-Min vs. ILP: Random 100
DS-Min
ILP
DS-2packing
DS-LP
100
10
1
0.1
0.01
Time (seconds)
0.2
0.4
0.6
0.8
Density


DS-Min and ILP
DS-Min
ILP-Min
DS-2-packing
DS-LP
4
3
2
1
0
Time (seconds)
arenas-jazz
food-web
HC-BIOGRID-2.0.31
ia-email
inf-USAir97
moreno-names
scc_fb-forum
socfb-Caltech
web-google
yeast_8
known-PPi
Graph