Can state-of-the-art MaxSAT-based preprocessing and solving be effectively used for cluster editing?

Imko Mariinissen

Supervisor: Dr. Emir Demirović

Problem Overview Cluster Editing Unweighted Partial MaxSAT Given two sets, hard clauses and soft clauses, what is an Given a Graph, what is the minimum number of edge assignment of literals that satisfies all the hard clauses modifications to turn it into a disjoint union of cliques? and minimizes the number of falsified soft clauses? Easier to Cluster Cluster MaxSAT MaxSAT Solve Editina Solver Reconstructor Editing Encoder Preprocessor Solution Instance MaxSAT Solution Graph Instance Figure 1: General Pipeline

Pipeline Specifics

Modelling [1]

- Transitive encoding (prevent conflict triples)
- Binary Encoding (assign vertices to cluster)
- Domain-specific Knowledge

Solvers

- MaxHS (Implicit Hitting Set) [2]
- EvalMaxSAT (Core-Guided) [3]
- Both performed well in MaxSAT Evaluation 2020

Preprocessor

- MaxPre [4]
- SAT Techniques
- (Weighted) MaxSAT Techniques

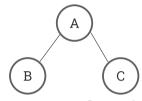
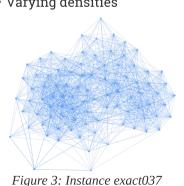
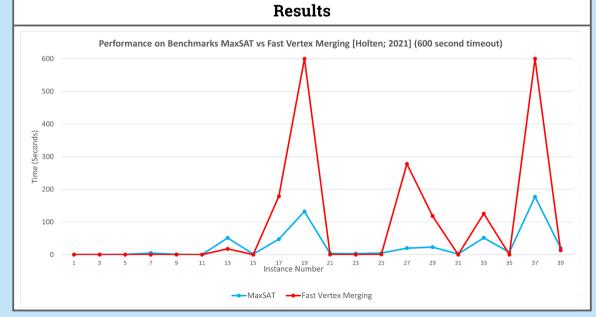


Figure 2: Conflict Triple

Benchmarks

- · Between 10 90 Vertices
- Between 11 1456 Edges
- Between 3-703 Solution Size
- Varying Structures
- Varying densities





Conclusions

- · Non-Obvious Interactions between modelling, preprocessing and solving
- · Applicability depends on input instance
- · Modelling is not very scalable (up to 500.000 clauses for 100 vertices)
- · MaxHS + Preprocessed Transitive Encoding appears to be the most stable
- MaxSAT approaches are possible but might not be guided in the same way as theoretical approaches

[1] J. Berg and M. Järvisalo, Cost-optimal constrained correlation clustering via weighted partial maximum satisfiability. [2] F. Bacchus, MaxHS in the 2020 MaxSat Evaluation. [3] F. Avellaneda, A short description of the solver EvalMaxSAT. [4] T. Korhonen, J. Berg, P. Saikko and M. Järvisalo, MaxPre: An Extended MaxSAT Preprocessor