CHARLES UNIVERSITY

CSI

IÚUK

Mgr. Martin Koutecký, Ph.D.
Charles University
Computer Science Institute
Malostranské nám. 25, 3rd floor, room #326
118 00 Praha 1, Czech Republic
Telephone: +420 774 853 316
Email: koutecky@iuuk.mff.cuni.cz

Review of Bachelor Thesis of Aryan Kumar

The thesis at hand provides an experimental evaluation of two possible solution approaches to a problem from computational social choice, called "bribery in society graphs" (R-BSG). The two approaches are via **1)** integer linear programming (ILP), or **2)** Presburger Arithmetic (PrA). Thus, in some sense, the thesis is a comparison of solvers for those two problems on a common computational problem. The ILP approach (first described by a paper of mine from IJCAI 2018) uses "encoding tricks" to implement logical disjunctions which are known to create difficulty for solvers; intuitively, they break the close relationships between the original ILP and its continuous relaxation. Thus, one would expect that a tool for which disjunctions are "natural" or "first order", i.e., logical solvers, would perform better. This is the premise of this thesis – validating whether the logical solvers could beat ILP solvers or not.

In order to get to the answer, the author of the thesis had to understand the R-BSG problem. I will not go into details of its description, but I wish to note that he had to absorb a non-trivial amount of background in computational social choice. Moreover, to understand the ILP encoding presented in the IJCAI '18 paper, and to come up with the PrA encoding, he had to also gain a lot of background in those two methods. Consequently, a "side contribution" of the thesis is a clear and detailed exposition of how to encode R-BSG as an ILP and as a PrA formula. The first encoding is present in a paper, but despite best efforts by its authors (myself included), the presentation was not entirely clear to a non-expert, and I believe that the author of the thesis has made this more accessible. The second encoding was simply not present in the literature at all.

Onto the results: the author first begins by comparing two approaches to finding an optimum value using PrA. The "logical" approach is to embed in the formula the statement that the found solution x is **a)** a solution, and **b)** not worse than any other solution, i.e., it is optimal. The alternative is to encode that x has cost at most B, and then perform binary search over B. The results overwhelmingly favor the binary search approach. The main message of the remainder of the results is that the ILP approach consistently beats the PrA approach, and that Gurobi (unsurprisingly) is the best among ILP solvers.

My notes / questions for the author are as follows:

1. It seems like the results would benefit from running more experiments. For example, there seems to be a lot of noise from few instances in all the plots - for instance in Figure 4.9, the lone dark square at diffusion steps = 8 and $\tau = 17$ seems surprising; it is hard to imagine that this combination of parameters is particularly easy, so it seems to me that the only explanation is that the number of instances corresponding to each cell is small.

Question: what *is* the number of instances corresponding to the various squares (I know it won't be the same for all; perhaps, what is the lowest, or median)?

2. Would it be feasible to get more data? Or is that computationally too hard?

3. You mention in the Conclusion testing solvers equipped for optimization - did you have some specific solvers in mind?

The thesis is written well and the results are presented using good structure and helpful plots.

Overall, the thesis provides an experimental counterpart to the theoretical results on the applicability of ILP and PrA, and it gives answers to natural questions so far not addressed in the literature. I suggest the grade of "1" without hesitation.

Sincerely,
Martin Koutecký