WARWICK
THE UNIVERSITY OF WARWICK

LeNSE: Learning To Navigate Subgraph Embeddings For Large-Scale Combinatorial Optimisation

David Ireland – WMG, University of Warwick

Prof. Giovanni Montana – Department of Statistics, University of Warwick

Combinatorial Optimisation

- Involves the search of maxima/minima over an objective function with a (large) discrete domain;
- The number of possible solutions 'explode' as the size of the problem increases – typically Combinatorial Optimisation problems are NP-hard;
- Most problems are naturally expressed over graphs;
- Examples of problems: Influence Maximisation, Max-Cut, Vertex Covering

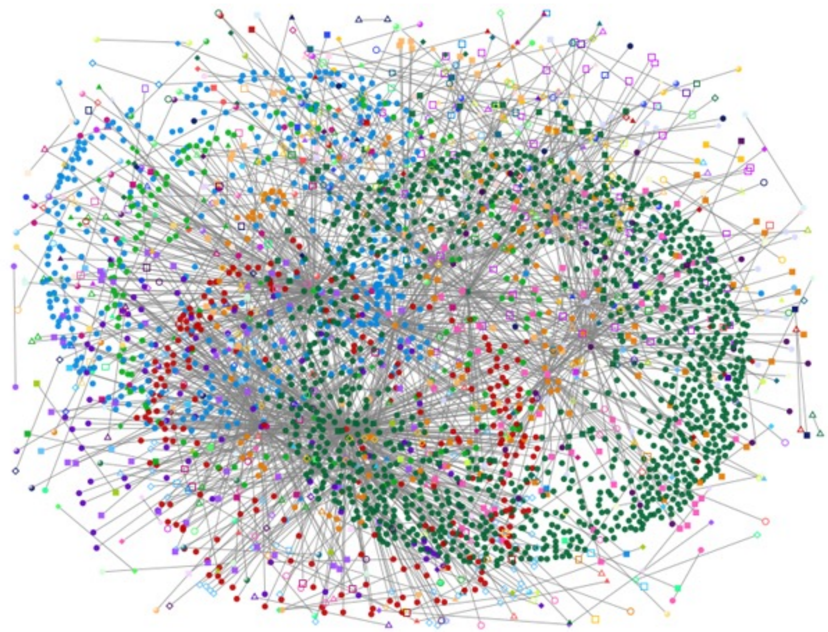


Photo Credit: https://threatpost.com/researchers-graph-social-networks-spot-spammers-061711/75346/

Learning to Navigate Subgraph Embeddings

- We make the observation that, for budget constrained problems, subgraphs of the original problem contain all the information needed to find a solution;

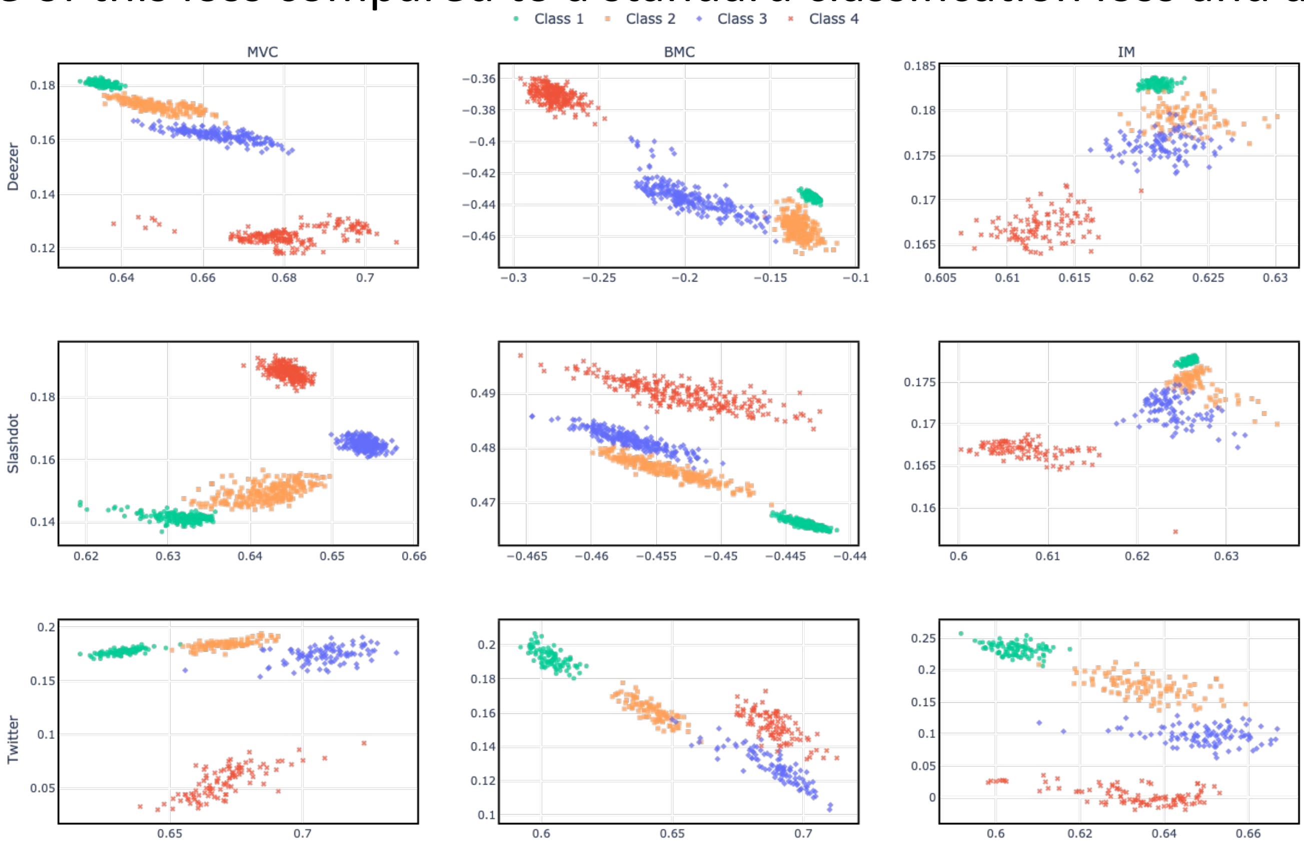
- Using this, we propose to learn to find an optimal subgraph, and then use an existing heuristic to find the solution in a much more efficient manner;

- To find an optimal subgraph, we learn to navigate subgraph embeddings;

- Problem Formulation: Given a graph $G = (V, E)$, objective function $f(\cdot)$, heuristic $\mathcal{H}(\cdot)$ and optimal solution X^*, we want to find a subgraph $S = (V_S, E_S)$ where $V_s \subset V, E_S \subset E$, such that $|V_S| \ll |V|$ and $f\left(\mathcal{H}(S)\right)/f(X^*) = 1$.

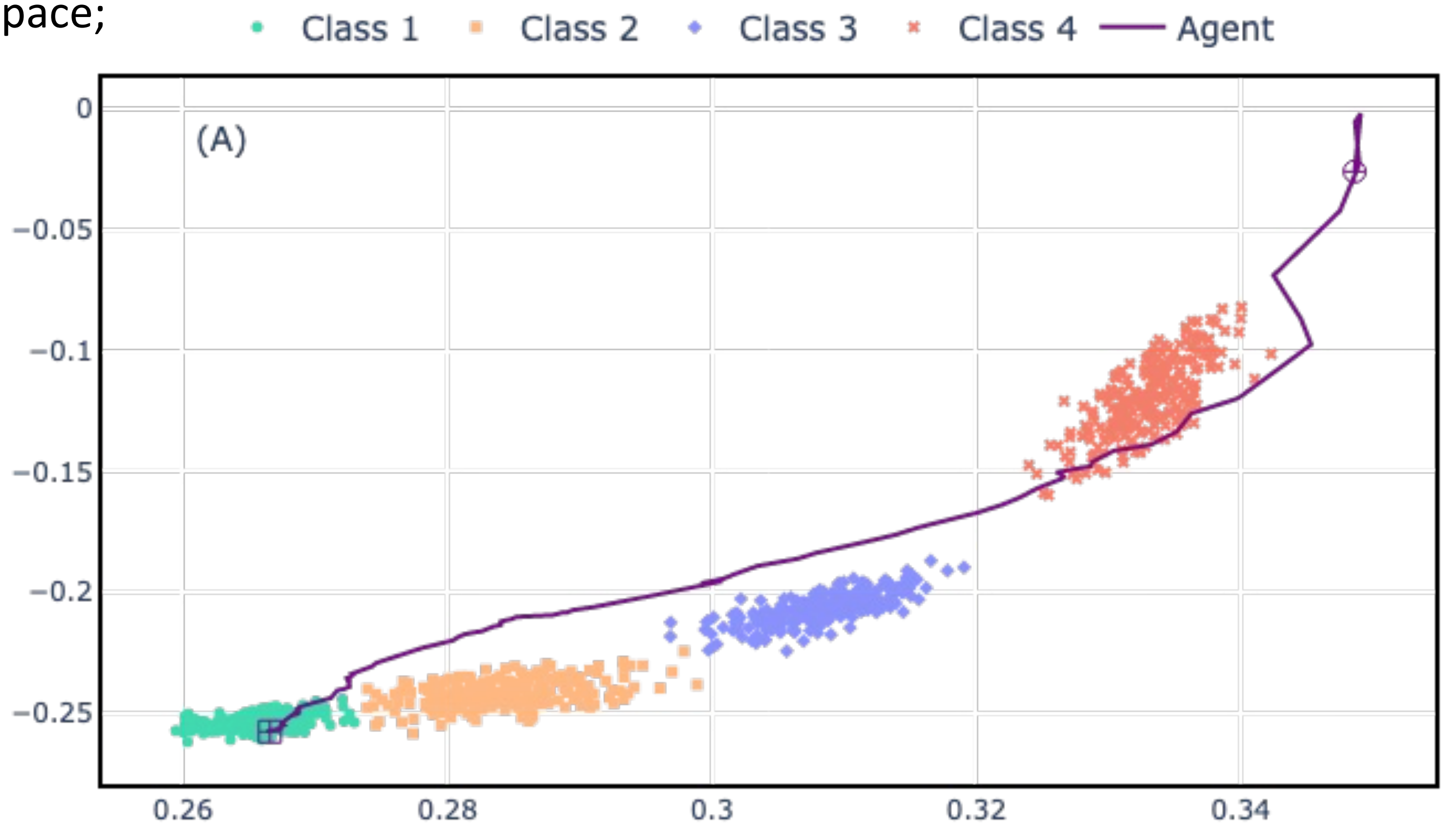
Learning a discriminative subgraph embedding

- To learn a discriminative subgraph embedding, we first generate a dataset of subgraphs;
- We take a small training graph $G_T \subset G$ where we can readily obtain solutions and sample N subgraphs, ranking each using the quality of the solution returned by the heuristic;
- An encoder is then trained by minimising the InfoNCE loss [1] – we perform an ablation justifying the choice of this loss compared to a standard classification loss and an ordinal based loss;



Subgraph Navigation

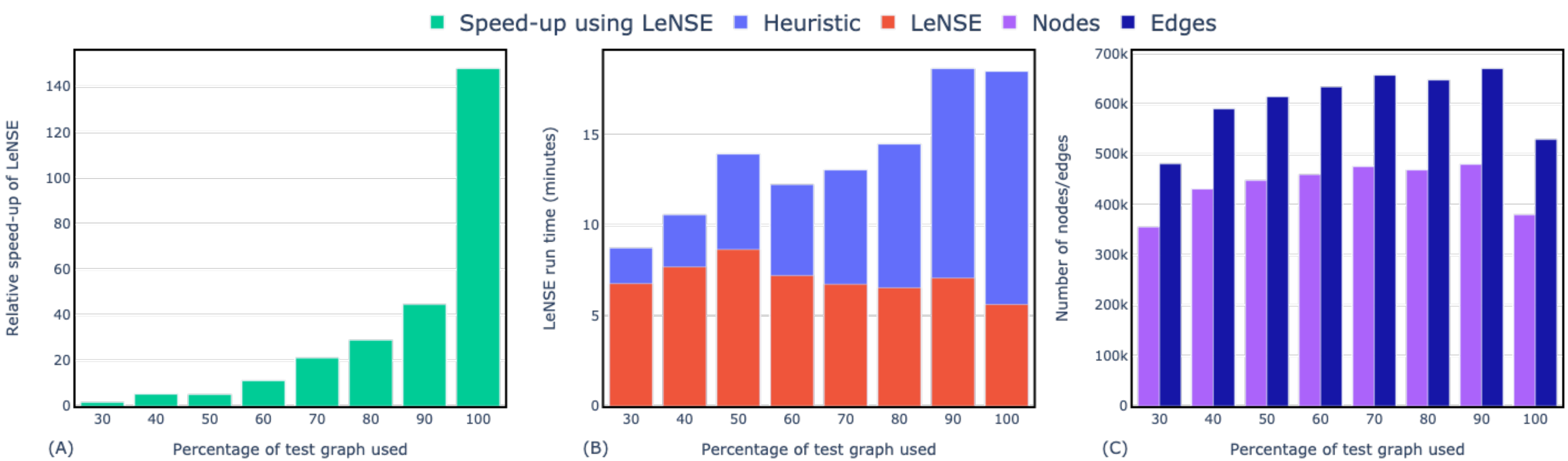
- Once the encoder is trained, we then learn how to navigate the embedding space;
- We formulate this as a sequential decision making problem and apply the standard Markov Decision Process framework;
- Starting from a random subgraph, an agent must learn to take subgraph modification operations (actions) that change subgraph, eventually leading to an optimal subgraph;
- The agent is rewarded for getting closer to the region of high scoring subgraphs in the pre-trained embedding space;



Results

THE UNIVERSITY OF WARWICK

- We found that LeNSE is capable of efficiently pruning the graph without any significant degradation in performance;
- Compared to the competing methods, LeNSE was able to prune significant amounts of the graph whilst also maintaining close to optimal performance;
- In some instances LeNSE was able to prune more than 90% of the nodes and edges;



Conclusion and Future Work

- We have introduced a framework that is able to scale up existing CO solvers by efficiently pruning the problem through subgraph navigation;
- When tested on several real world graphs we found that we can, in the best case, prune more than 90% of the graph without a significant performance degradation;
- Further, we also demonstrate that on a large, real world graph the speed up of using the LeNSE framework provides a 140x speed-up over using the heuristic without LeNSE;
- Future directions include looking to generalise this framework to non-budget constrained problems, where the size of the solution is not known *a-priori*.

References

- [1] Representation learning with contrastive predictive coding – Oord et al. 2018