Active search is an increasingly important learning problem in which we use a limited budget of label queries to discover as many members of a certain class as possible. Numerous real-world applications may be approached in this manner, including fraud detection, product recommendation, and drug discovery. Active search has model learning and exploration/exploitation features similar to those encountered in active learning and bandit problems, but algorithms for those problems do not fit active search well.

Previous work on the active search problem showed that the optimal algorithm requires a lookahead evaluation of expected utility that is exponential in the number of selections to be made and proposed a truncated lookahead heuristic. In this talk, we propose a myopic method for active search on graphs. We suggest selecting points by maximizing a score considering the potential impact of selecting a node, meant to emulate lookahead while avoiding exponential search. We test the proposed algorithm empirically on real-world graphs and show that it outperforms popular approaches for active learning and bandit problems as well as truncated lookahead of a few steps.

This is joint work with Jeff Schneider and Roman Garnett.