Logarithmic Approximations for Fair k-Set Selection*

Shi Li[1], Chenyang Xu[2], and Ruilong Zhang[3]

[1]School of Computer Science, Nanjing University, Nanjing, China
[2]Software Engineering Institute, East China Normal University, Shanghai, China
[3]Department of Computer Science, City University of Hong Kong (Dongguan), Dongguan, China
shili@nju.edu.cn, cyxu@sei.ecnu.edu.cn, ruilong.zhang@cityu-dg.edu.cn

Abstract

We study the fair k-set selection problem where we aim to select k sets from a given set system such that the (weighted) occurrence times that each element appears in these k selected sets are balanced, i.e., the maximum (weighted) occurrence times are minimized. By observing that a set system can be formulated into a bipartite graph $G := (L \cup R, E)$, our problem is equivalent to selecting k vertices from R such that the maximum total weight of selected neighbors of vertices in L is minimized. The problem arises in a wide range of applications in various fields, such as machine learning, artificial intelligence, and operations research.

We first prove that the problem is NP-hard even if the maximum degree Δ of the input bipartite graph is 3, and the problem is in P when $\Delta = 2$. We then show that the problem is also in P when the input set system forms a laminar family. Based on intuitive linear programming, we show that a dependent rounding algorithm achieves $O(\frac{\log n}{\log \log n})$-approximation on general bipartite graphs, and an independent rounding algorithm achieves $O(\log \Delta)$-approximation on bipartite graphs with a maximum degree Δ. We demonstrate that our analysis is almost tight by providing a hard instance for this linear programming. Finally, we extend all our algorithms to the weighted case and prove that all approximations are preserved.

*In the Proceedings of IJCAI 2025