Dispersion on Trees

An M.Sc. Proposal by Nadav Krasnopolsky

Abstract

In the dispersion problem we want to select k nodes of a given graph as to maximize the minimum distance between any two chosen nodes. This can be seen as a generalization of the independent set problem, where the goal is to select nodes so that the minimum distance is larger than 1. We design an optimal $O(n)$ time algorithm for the dispersion problem on trees consisting of n nodes, thus improving the previous $O(n \log n)$ time solution.

Then we consider the weighted version. For the decision version, where the goal is to find a set of nodes of total weight at least k such that any two nodes are at distance at least λ, we present tight $\Theta(n \log n)$ upper and lower bounds. For the optimization version, where we wish to maximize λ, we present an $O(n \log^2 n)$ upper bound improving the previous $O(n \log^4 n)$ time solution.

1 Introduction

Facility location is a family of problems, where the goal is to place a number of facilities as to minimize the total cost while preserving given constraints. In its basic version, called the metric k-center problem, we wish to designate up to k nodes of a given weighted graph to be facilities, as to minimize the maximum distance of a node to its closes facility. Even this simplest version is already NP-hard [Vaz03]. However, a simple greedy 2-approximation algorithm is known [Gon85]. The situation for different objective functions or more restrictions on the placement is of course more complex, but nevertheless one can often design an efficient approximation algorithm, see [STA97] and references therein.

Given that facility location problems are usually if not always NP-hard, it makes sense to consider them on restricted graphs families. The simplest yet still non-trivial such a family are trees on n nodes. There are multiple possible versions of the k-center problem on trees: the facilities can be either only the nodes of the tree or any points on an edge, and we might either minimize the distance of any node or any point on an edge to its closest facility. All these versions have been extensively studied, culminating in an $O(n)$ time and space algorithm given by Frederickson [Fre91b] that solves all but one of them. This was an improvement on the previous $O(n \log n)$ time algorithm by Frederickson and Johnson [FJ], which in turn was an improvement on the $O(n \log^2 n)$ time algorithm of Megiddo et al. [MTZC].

The weighted version of these problems has also been studied. In this version, every node has an associated weight and the distance is multiplied by the corresponding weight. The currently fastest solution is the $O(n \log^2 n)$ time algorithm of Megiddo and Tamir [MT]. Among the facility location problems on trees there is the max-min tree k-partitioning [PS] and the min-max tree k-partitioning [BSP], both solved by Frederickson in $O(n)$ time [Fre91a] using a clever approach based on the parametric search that he has later extended to solve the k-center problem.

In this paper we focus on the max-min tree k-dispersion problem that falls within the class of facility location problems on trees. In this problem, the goal is to select k nodes so as to maximize the minimum distance between any two chosen nodes. In other words, we wish to select k nodes that are as spread-apart as possible. This generalizes the maximum independent set problem by binary searching for the largest value of k for which the minimum distance is at least 2. The best previously known algorithm for the max-min tree k-dispersion was given by Bhattacharya and Houle [BH91] and takes $O(n \log n)$ time, improving on an earlier $O(kn + n \log n)$ solution for the one dimensional case given by Wang and Kuo [WK88] . Using the framework of Frederickson, we construct an optimal $O(n)$ time algorithm for this problem. In our solution we also develop a slightly modified version of this framework, which might be simpler to understand.

For a subset of nodes $P \subseteq V$, let $d(u, v)$ denote the distance between nodes u and v and let $f(P) = min_{u,v \in P}\{d(u, v)\}$. The problems we solve are formally defined as:

1. **The Dispersion Optimization Problem:** Given a tree T with non-negative edge lengths, and a number p, find a subset $P \subseteq V$ of size p s.t. $f(P)$ is maximized.

2. **The Dispersion Decision Problem:** Given a tree T with non-negative edge lengths, a number p, and a number λ, find a subset $P \subseteq V$ of size p s.t. $f(P) \geqslant \lambda$.

We then move to the weighted version, where the tree nodes are weighted and the goal is to select P nodes with total weight at last k. There, the best previously known solution requires $O(n \log^4 n)$ time [BH91]. We present a significantly faster $O(n \log^2 n)$ time algorithm. We also show that the decision version, where the goal is to check if there exists a set of of nodes total weight k such that any two nodes are at distance at least b, requires $\Omega(n \log n)$ time in the algebraic decision tree model. We also present a matching $O(n \log n)$ time algorithm, which is then used to construct our $O(n \log^2 n)$ time solution for the weighted max-min tree k-dispersion using the standard parametric search paradigm of Fredrickson.

References

[BH91] B.K. Bhattacharya and M.E. Houle. Generalized maximum independent sets for trees in subquadratic time. In *ISAAC*, 1991.

[BSP] R.I. Becker, S.R. Schach, and Y. Perl. A shifting algorithm for min-max tree partitioning. *J. ACM*, 29(1):56.

[FJ] G.N. Frederickson and D.B. Johnson. Finding k-th paths and p-centers by generating and searching good data structures. *J. Algorithms*, 4(1):61.

[Fre91a] G.N. Frederickson. Optimal algorithms for tree partitioning. In *SODA*, 1991.

[Fre91b] G.N. Frederickson. Parametric search and locating supply centers in trees. In *WADS*, 1991.

[Gon85] T.F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 1985.

[MT] N. Megiddo and A. Tamir. New results on the complexity of p-center problems. *SIAM J. Computing*, 12(3):751.

[MTZC] N. Megiddo, A. Tamir, E. Zemel, and R. Chandrasekaran. An $o(nlog^2n)$ algorithm for the k-th longest path in a tree with applications to location problems. *SIAM J. Computing*, 10(2):328.

[PS] Y. Perl and S.R. Schach. Max-min tree partitioning. *J. ACM*, 28(1):5.

[STA97] D.B. Shmoys, É. Tardos, and K. Aardal. Approximation algorithms for facility location problems. In *STOC*, 1997.

[Vaz03] V.V. Vazirani. *Approximation Algorithms.* Springer, 2003.

[WK88] D.W. Wang and Y. Kuo. A study on two geometric location problems. *Information Processing Letters*, 1988.