Approximation Algorithms for Dispersion Problems in Euclidean and Metric Spaces

Gautam Kumar Das

January 20, 2026

Abstract

Facility Location Problems (FLPs) have been the subject of extensive research due to their wide range of applications in VLSI design, networking, clustering, and other areas. The covering problem and the dispersion problem are two popular FLPs. The covering problem refers to selecting a subset of objects from a given set such that the union of the selected objects covers all elements. In contrast, the dispersion problem aims to select a subset of objects such that the selected objects are as far apart from one another as possible. In this talk, we focus on the dispersion problem.

For the dispersion problem, we consider the concept of the dispersion partial sum, which generalizes the classical notion of dispersion. Based on this concept, we define several variants of the dispersion problem, namely the 1-dispersion problem, the 2-dispersion problem, and the c-dispersion problem. We study these problems in Euclidean space, specifically the 2-dispersion problem in $\mathbb{R}^1$ and $\mathbb{R}^2$, and the 1-dispersion problem in $\mathbb{R}^2$. The 2-dispersion problem is defined as follows.

Given a set $P = \{p_1, p_2, \ldots, p_n\}$ of n points, the non-negative Euclidean distances between each pair of points $p_i, p_j \in P$, and a positive integer k ($3 \leq k \leq n$), for each point $p \in P$ and a subset $S \subseteq P$, the 2-dispersion cost of the point p with respect to S, denoted by $cost_2(p, S)$, is defined as the sum of the Euclidean distances from p to its closest and second closest points in $S \setminus p$. The 2-dispersion cost of the subset S is defined as $cost_2(S) = \min_{p \in S} cost_2(p, S)$. The objective of the 2-dispersion problem is to find a subset $S \subseteq P$ of cardinality k such that $cost_2(S)$ is maximized.

We present a $(2\sqrt{3}+\epsilon)$-factor approximation algorithm for the 2-dispersion problem in $\mathbb{R}^2$. We also develop a unified framework for dispersion problems in Euclidean space, which yields a $2\sqrt{3}$-factor approximation algorithm for the 2-dispersion problem in $\mathbb{R}^2$ and an optimal algorithm for the same problem in $\mathbb{R}^1$.

Next, we study the 1-dispersion problem, defined as follows. Given a set $P = \{p_1, p_2, \ldots, p_n\}$ of n points, the non-negative distances between each pair of points $p_i, p_j \in P$, and a positive integer k ($2 \leq k \leq n$), for each point $p \in P$ and a subset $S \subseteq P$, the 1-dispersion cost of the point p with respect to S, denoted by $cost_1(p, S)$, is defined as the distance from p to its closest point in $S \setminus p$. The 1-dispersion cost of the subset S is defined as $cost_1(S) = \min_{p \in S} cost_1(p, S)$. The objective of the 1-dispersion problem is to find a subset $S \subseteq P$ of cardinality k such that $cost_1(S)$ is maximized.

We propose a 2-factor approximation algorithm for the 1-dispersion problem in $\mathbb{R}^2$.

We then investigate the dispersion problem in general metric spaces by introducing another variant, called the c-dispersion problem, which is formally defined as follows.

Given a set $P = \{p_1, p_2, \ldots, p_n\}$ of n points, a non-negative distance function $d(p_i, p_j)$ defined for each pair of points $p_i, p_j \in P$, and a positive integer k ($c+1 \leq k \leq n$), for each point $p \in P$ and a subset $S \subseteq P$, the c-dispersion cost of the point p with respect to S, denoted by $cost_c(p, S)$, is defined as the sum of distances from p to its c closest points in $S \setminus p$. The c-dispersion cost of the subset S is defined as $cost_c(S) = \min_{p \in S} cost_c(p, S)$. The objective of the c-dispersion problem is to find a subset $S \subseteq P$ of size k such that $cost_c(S)$ is maximized.

We propose a greedy algorithm for the c-dispersion problem that achieves a $2c$-factor approximation guarantee. We also prove that the c-dispersion problem in a metric space, parameterized by the solution size, is W[1]-hard.

Finally, we consider a variant of the 1-dispersion problem in which the set of locations corresponds to the vertices of a convex polygon. This variant is referred to as the convex 1-dispersion problem and is defined as follows.

Given a set $P = \{p_1, p_2, \ldots, p_n\}$ of n vertices of a convex polygon, with Euclidean distances $d(p, q)$ defined between each pair of vertices $p, q \in P$, the objective of the convex 1-dispersion problem is to find a subset $S \subseteq P$ of size k such that the cost of S, defined as $cost(S) = \min d(p, q) \mid p, q \in S$, is maximized.

We present an $O(n^3)$-time algorithm that computes an optimal solution for the convex 1-dispersion problem when $k = 4$. Additionally, we propose a $\sqrt{3}$ (approximately 1.733)-factor approximation algorithm for the convex 1-dispersion problem for arbitrary values of k.