Approximating (k, ℓ)-center Clustering for Curves

Michael Horton
University of Sydney and New York University

May 1, 2018

Abstract

The Euclidean k-center problem is a classical problem that has been extensively studied in many areas of computer science. Given a set $\mathcal{G}$ of n points in Euclidean space, the problem is to determine a set $\mathcal{C}$ of k centers (not necessarily part of $\mathcal{G}$) such that the maximum distance between a point in $\mathcal{G}$ and its nearest neighbor in $\mathcal{C}$ is minimized. In this paper we study the corresponding (k, ℓ)-center problem for polygonal curves under the Fréchet distance, that is, given a set $\mathcal{G}$ of n polygonal curves in $\mathbb{R}^d$, each of complexity m, determine a set $\mathcal{C}$ of k polygonal curves in $\mathbb{R}^d$, each of complexity ℓ, such that the maximum Fréchet distance of a curve in $\mathcal{G}$ to its closest curve in $\mathcal{C}$ is minimized.

We show that there is no polynomial-time $(\frac{3}{2} - \varepsilon)$-approximation algorithm for any $\varepsilon > 0$ unless $\mathsf{P} = \mathsf{NP}$. This bound even holds for one-dimensional curves under the continuous Fréchet distance, and is further improved to $(3\sin(\frac{\pi}{3}) - \varepsilon)$ if the curves may lie in the plane and if the discrete Fréchet distance is used. These hardness results hold even when $k = 1$, i.e., for the minimum-enclosing ball version of the problem. At the same time, we prove that a careful adaptation of Gonzalez' algorithm in combination with a curve simplification yields a 3-approximation in any dimension, provided that an optimal simplification can be computed exactly. We give a comprehensive analysis of this approximation scheme in spaces of arbitrary dimension and for both the discrete and continuous Fréchet distances. Overall, our results significantly extend and improve the known approximation bounds for the (k, ℓ)-center clustering problem.

Joint work with Kevin Buchin, Anne Driemel, Joachim Gudmundsson, Irina Kostitsyna, and Maarten Löffler.