Constant Approximation of Fréchet Distance in Strongly Subquadratic Time*

Siu-Wing Cheng† Haoqiang Huang† Shuo Zhang‡

Fréchet distance is a popular and natural distance metric to measure the similarity between two curves. It is useful in many real-world applications. Consider a polygonal curve τ in $\mathbb{R}^d$, with vertices $(v_1, v_2, \ldots, v_n)$. A *parameterization* of τ is a continuous function $\rho : [0,1] \to \tau$ such that $\rho(0) = v_1$, $\rho(1) = v_n$, and for all $t, t' \in [0,1]$, $\rho[t]$ is not behind $\rho[t']$ from v_1 to v_n along τ if $t \leq t'$. Let ϱ be a parameterization for another polygonal curve σ of m vertices. The pair (ρ, ϱ) define a *matching* between τ and σ such that $\rho(t)$ is matched to $\varrho(t)$ for all $t \in [0,1]$. The distance realized by $\mathcal{M}$ is $d_{\mathcal{M}}(\tau,\sigma) = \max_{t\in[0,1]} d(\rho(t), \varrho(t))$, where $d(\rho(t), \varrho(t))$ is the Euclidean distance between $\rho(t)$ and $\varrho(t)$. The Fréchet distance of σ and τ is $d_F(\sigma,\tau) = \inf_{\mathcal{M}} d_{\mathcal{M}}(\tau,\sigma)$. If we restrict $\mathcal{M}$ to match every vertex of σ to a vertex of τ, and vice versa, the resulting $\inf_{\mathcal{M}} d_{\mathcal{M}}(\tau,\sigma)$ is the *discrete Fréchet distance*, which we denote by $\tilde{d}_F(\tau,\sigma)$.

The state-of-the-art for approximating $d_F(\tau,\sigma)$ is an $O(\alpha)$-approximation algorithm that runs in $O((n + mn/\alpha)\log^2 n)$ time due to Van der Horst and Ophelders. For $\tilde{d}_F(\tau,\sigma)$, the best known result is an $O(n\log n + n^2/\alpha^2)$-time $O(\alpha)$-approximation algorithm assuming that $m = n$ and $\alpha \in [1, \sqrt{n/\log n}]$ due to Chan and Rahmati. All these known approximation schemes can only achieve approximation ratios of n^c in strongly subquadratic time. In spite of the extensive effort devoted in designing fast algorithms for computing $d_F(\tau,\sigma)$ and $\tilde{d}_F(\tau,\sigma)$, the existence of constant approaximation algorithms in stongly subquadratic time, i.e., $O\big((nm)^{1-c}\big)$ for some constant $c \in (0,1)$, was repeatedly raised as an open problem in the literature.

We answer this open problem affirmatively. Our main result is a *randomized* $(7+\varepsilon)$-approximation decision procedure that runs in $O(nm^{0.99})$ time. Specifically, given two polygonal curves τ, σ and a fixed value δ, for any $\varepsilon \in (0,1)$, if our decision procedure returns yes, then $d_F(\tau,\sigma) \leq (7+\varepsilon)\delta$; if it returns no, then $d_F(\tau,\sigma) > \delta$ with probability at least $1 - n^{-7}$. The technical ideas can be adapted to get a *randomized* strongly subquadratic $(7+\varepsilon)$-approximation decision procedure for the discrete Fréchet distance as well. We then follow the conventional approaches to turn our decision procedures into $(7+\varepsilon)$-approximation algorithms for computing $d_F(\tau,\sigma)$ and $\tilde{d}_F(\tau,\sigma)$, respectively. These approaches only introduce extra polylog factors in the running times. Our algorithms are the first strongly subquadratic algorithms that can achieve constant approximation ratios, which improves the previous ratio of n^c significantly.

Our results is based on a novel framework that can accelerate the classic dynamic programmings for computing $d_F(\tau,\sigma)$ and $\tilde{d}_F(\tau,\sigma)$. It involves a clever use of curve simplification and sampling techniques. While the ratio of $(7+\varepsilon)$ is a significant improvement, there is potential that several components within the framework can be further optimized to get better approximation ratio and running time, which is an interesting future work.

---

*This paper will appear at STOC 2025.

†Department of Computer Science and Engineering, HKUST, Hong Kong. Email: scheng@cse.ust.hk, haoqiang.huang@connect.ust.hk

‡Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China. Email: zhangshuo1422@ruc.edu.cn