Computing the discrete Fréchet distance upper bound of imprecise input is NP-hard

Chenglin Fan, Binhai Zhu
Department of Computer Science
Montana State University, Bozeman, MT 59717, USA
chenglin.fan@msu.montana.edu, bhz@cs.montana.edu

October 6, 2015

1 Introduction

The Fréchet distance is a natural measure of similarity between two curves [4]. The Fréchet distance between two curves is often referred to as the "dog-leash distance". Alt and Godau [4] presented an algorithm to compute the Fréchet distance between two polygonal curves of n and m vertices in $O(nm\log(nm))$ time. There has been a lot of applications using the Fréchet distance to do pattern/curve matching.

A slightly simpler version of the Fréchet distance is the *discrete Fréchet distance*, where only the vertices of polygonal curves are considered. It takes $O(mn)$ time to compute the discrete Fréchet distance using a standard dynamic programming technique [6].

The discrete Fréchet distance with shortcuts was also studied by Avraham *et al.* recently [5]. A novel technique, based on distance selection, was designed to compute the discrete Fréchet distance with shortcuts efficiently.

The computational geometry with imprecise objects has drawn much interest to researchers since a few years ago. There are two models: one is the *continuous* model, where a precise point is selected from an erroneous region (say a disk, or rectangle) [7]; the other is the *discrete* or *color-spanning* model, where a precise point is selected from several discrete objects with the same color and all colors must be selected [1]. A lot of algorithms have been designed to handle imprecise geometric problems on both models.

Ahn *et al.* studied the problem of computing the discrete Fréchet distance between two imprecise point sequences, and gave an efficient algorithm for computing the lower bound (of the distance) and efficient approximation algorithms for the corresponding upper bound (under a realistic assumption) [2,3]. It is unknown whether computing the discrete Fréchet distance upper bound for imprecise input is solvable in polynomial time or not, so Ahn *et al.* left that as an open problem [2,3]. In this paper, we proved that the problem is in fact NP-hard. We also consider the same problem under the discrete Fréchet distance with shortcuts and give efficient polynomial-time solutions.

Main results: We study the problem of computing the upper bound of the discrete Fréchet distance for imprecise input, and prove that the problem is NP-hard. This solves an open problem posted in 2010 by Ahn *et al.* If shortcuts are allowed, we show that the upper bound of the discrete Fréchet distance with shortcuts for imprecise input can be computed in polynomial time and we present several efficient algorithms[1].

2 Preliminaries

Definition For a region q_i, a precise point a_i is called a realization of q_i if $a_i \in q_i$; For a region sequences $Q = (q_1, q_2, ..., q_n)$, the precise point sequence $A = (a_1, a_2, ..., a_n)$ is called a *realization* of Q if we have $a_i \in q_i$ for all $1 \le i \le n$.

For the discrete Fréchet distance of imprecise input, we use the same notions such that the *realization* of an imprecise input sequence as in [2]. To be consistent with these notations, we also use $F(A, B)$ to denote the discrete Fréchet distance between A and B (i.e., $F(A, B) = d_{dF}(A, B)$).

Definition For two region sequences $Q = (q_1, q_2, ..., q_n)$ and $H = (h_1, h_2, ..., h_m)$, $A = (a_1, a_2, ..., a_n)$ (resp. $B = (b_1, b_2, ..., b_m)$) is a possible realization of H (resp. Q) . The Fréchet distance upper bound $F^{\max}(Q, H) = \max\{F(A, B)\}$, where A (resp. B) is a possible realization of Q (resp. H).

[1]The full version of this paper can be found at http://arXiv:1509.02576

We show in the next section that computing $F^{\max}(Q, H)$ is NP-hard, which was an open problem posted by Ahn *et al.* in [2].

3 Overview of the Proof

In this section, we give an overview to prove $F^{\max}(Q, H) > \epsilon$ is NP-hard. In fact, this holds even when H is a precise vertex sequence, and Q is an imprecise vertex sequence (where each vertex is modeled as a rectangle, not axis aligned). As the proof is quite complex, we separate it in several steps.

Firstly, we prove that another induced subgraph connectivity problem of colored sets is NP-hard, which is useful for the proof of deciding $F^{\max}(Q, H) > \epsilon$. We define the induced subgraph connectivity problem of colored sets (ISCPCS) as follows: let G be the graph with n vertices and each vertex is colored by one of the m colors in the plane, a fixed source vertex s, a fixed destination vertex t, and some directed edges between the vertices (where no two edges cross), choose an induced subgraph G_s consisting of exactly one vertex of each color such that in G_s there is no path from s to t. We prove that the ISCPCS problem is NP-hard by a reduction from 3SAT.

Secondly, we construct an equivalent free space diagram (grid graph) corresponding to the graph G in ISCPCS. In the special free space diagram, some cells can be passed, and some other cells can not be passed, and the rest of cells are colored, for the cells colored by same color, at least one of them can be passed. We show that $F^{max}(Q, H) > \epsilon$ if and only if there exist a choice that choose exactly one passable cell of each color such that there is no monotone path from lower-left cell to the upper-right cell in the equivalent free space grid graph.

To complete the proof that deciding $F^{\max}(Q, H) > \epsilon$ is NP-hard, we need to construct a precise vertex sequence H and an imprecise vertex sequence Q (where each imprecise vertex is modeled as a rectangle) such that the free space grid graph constructed above can be geometrically realized.

Our NP-hardness proof is quite complicated, the construction has a combinatorial and a geometric part. In the combinatorial part, we interpret the imprecise discrete distance in terms of finding monotone paths through a colored grid graph; In the geometric part, we show that the relevant colored free space diagram grids can be realized geometrically.

It would be interesting to consider the continuous Fréchet distance of imprecise input in the future.

References

[1] M. Abellanas, F. Hurtado, C. Icking, R. Klein, E. Langetepe, L. Ma, B. Palop, and V. Sacristan. Smallest color-spanning objects. *Proc. 9th European Sympos. Algorithms (ESA'01)*, pp. 278–289, 2001.

[2] H-K. Ahn, C. Knauer, M. Scherfenberg, L. Schlipf, and A. Vigneron. Computing the discrete Fréchet distance with imprecise input. In *Proc. ISAAC'10*, pp. 422-433, 2010.

[3] H-K. Ahn, C. Knauer, M. Scherfenberg, L. Schlipf, and A. Vigneron. Computing the discrete Fréchet distance with imprecise input. Int. J. Comput. Geometry Appl., 22(1):27-44, 2012.

[4] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *Int. J. Comput. Geometry Appl.*, 5:75-91, 1995.

[5] R. Avraham, O. Filtser, H. Kaplan, M. Katz, and M. Sharir. The discrete Fréchet distance with shortcuts via approximate distance counting and selection. In Proc. SoCG'14, pages 377, 2014.

[6] T. Eiter and H. Mannila. Computing discrete fréchet distance. Technical report, Technische Universitat Wien, 1994.

[7] M. Löffler and M.J. van Kreveld. Largest and smallest tours and convex hulls for imprecise points. In *Proc. SWAT'06*, LNCS 4059, pp. 375-387, 2006.