NAKHODA

Journal of Algorithms and Computation

journal homepage: http://jac.ut.ac.ir

Fréchet-Like Distances between Two Rooted Trees

Elena Farahbakhsh Touli*[1]

[1]Stockholm University, Department of Mathematics.

ABSTRACT

The purpose of this paper is to extend the definition of Fréchet distance which measures the distance between two curves to a distance (Fréchet-Like distance) which measures the similarity between two rooted trees. In this paper, I prove that the Fréchet-Like distance between two trees is SNP-hard to compute.
Later, I modify the definition of Fréchet-Like distance to measure the distance between two merge trees, and I prove the relation between the interleaving distance and the modified Fréchet-Like distance.

ARTICLE INFO

Article history:
Research Paper
Received 12, April 2020
Received in revised form 15, April 2021
Accepted 6 May 2021
Available online 01, June 2021

Keyword: Merge trees, Fréchet distance, Fréchet-Like distance, modified Fréchet-Like distance, interleaving distance.

AMS subject Classification: 05C78.

1 Introduction

In this paper I am interested in extending the definition of Fréchet distance between curves to a distance between two trees.
Fréchet distance between curves is a distance for measuring the similarity between two curves. For the first time Fréchet distance was defined by Maurice Fréchet [4, 9, 10]. Later, Fréchet distance attracted attention and was worked on by other people [1, 3, 4, 6, 7].
The intuitive definition of Fréchet distance between two curves is as follows: A man and his dog start from the starting points of two curves and a leash connects the dog to the

*Corresponding author: E. F. Touli. Email: elena.touli@math.su.se

man. They can only go forward. The Fréchet distance between the curves is the minimum length of the leash that the man and the dog start from the beginning of the curves and they reach to the end of the curves without separating the leash. In the following I write the mathematical definition of the Fréchet distance between two curves [4].

Definition 1. *[4] Suppose that we have two curves $C_1 : [a, b] \longrightarrow V$ and $C_2 : [a', b'] \longrightarrow V$, such that $a < b$ and $a' < b'$ and V is a vector space. The Fréchet distance between C_1 and C_2 is defined as the infimum distance over all continuous increasing functions $\alpha : [0, 1] \longrightarrow [a, a']$ and $\beta : [0, 1] \longrightarrow [b, b']$ that maximizes the distance between $C_1(\alpha(t))$ and $C_2(\beta(t))$ on $t \in [0, 1]$. In this case, the Fréchet distance is defined as follows:*

$$d_F(C_1, C_2) = \inf_{\alpha,\beta} \max_{t\in[0,1]} \{d(C_1(\alpha(t)), C_2(\beta(t)))\}.$$

Weak Fréchet distance is a special kind of Fréchet distance such that the man and the dog can go backward as well [4]. Both Fréchet distance and weak Fréchet distance can be computed in a polynomial time between two polygonal curves [4], but it is NP-hard to compute the Fréchet distance between two surfaces [12] and till now no one has defined Fréchet distance between trees. Discrete Fréchet distance was discussed by T. Eiter and H. Mannila in 1994 [8]. In 2012, P.K.Agrawal, etc. found an algorithm to find the discrete Fréchet distance between two polygonal curves in sub-quadratic time. [1]

New work. In this paper I will extend the definition of Fréchet distance between curves to define a similar distance between rooted trees.

This is the first time that Fréchet distance is defined between trees. I call it Fréchet-Like distance because of the similarity of this definition to the Fréchet distance between curves. The intuitive definition of the Fréchet-Like distance is as follows:

Two men (A and B) start from the roots of two different trees T_1 and T_2 (A starts from the root of tree T_1 and B starts from the root of tree T_2) and a rope connects A to B. When man A reaches to a node (say x) with the degree of more than 2, he constructs $k - 1$ men which k is the outgoing degree of x. By constructing $k - 1$ men, the rope is divided into k as well. Therefore, each of k man can take one end point of the rope (see Figure 1). Each k men ($k - 1$ constructed men and man A) goes through one edge which connects the node x to its children. Same situation happens for the man B. Each man in tree T_1 is connected by a rope to at least one man in tree T_2 and vice versa. When there is a rope between the man A and B we say that the man B is monitored by A and the man A is monitoring the man B. There are many possibility for the man A

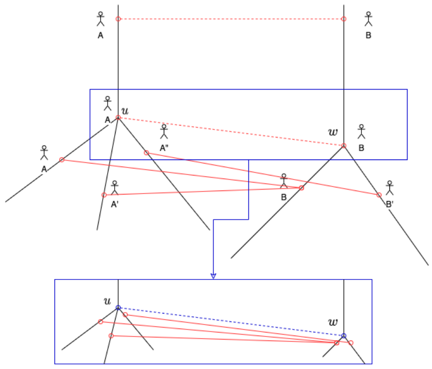

Figure 1: Two men A and B start to walk from the roots of two trees T_1 and T_2. When A reaches to the nodes u with the out degree of 3 and B reaches to w, A constructs A' and A'' and B constructs B'.

such that starts from the root of T_1 and monitoring B and they construct other men when they reach to a node and cover all the nodes in two trees. The matching distance for each possibility is the maximum length of the ropes between the men in tree T_1 and the men in tree T_2 such that are connected by a rope and they all go forward (the geodesic distance between them to the root of the tree increases) and reach to the leaves of the trees (each leaf is connected at least to a leaf). The Fréchet-Like distance is the minimum over all matching distance for each possibility of matching.
Later, I will modify the definition of Fréchet-Like distance to a definition between two merge trees. By considering the merge trees T_1^f and T_2^g, I prove the relation between the modified Fréchet-Like distance and the interleaving distance between two merge trees.

Definition 2. *Merge tree. [14, 15]*
A merge tree is a rooted tree with a function which is defined on each point of the tree. A merge tree T^h is defined by a pair (T, h) such that $h : |T| \longrightarrow \mathbb{R}$ is a monotone function which means that if for $x, y \in |T|$, $x < y^1$, $h(x) < h(y)$.
Intuitively we can define a merge tree (T, h) as follows: consider a tree and a node of the tree as the node u. Hang the tree from the node. I consider the function value $h(u) = 0$ for u that I hang the tree from and for all the other points in the tree, the function of each point of the merge tree T_u^h will be the negative distance between the node u and the point.

The outcome of this paper is as follows: The distance between two trees is discussed in section 2. In section 3, I define the Fréchet-Like distance between trees, both the intuition and mathematical definition of Fréchet-Like distance. In section 4, I prove that it is NP-hard to approximate the Fréchet-Like distance between two rooted trees. Section 5 is considered for modifying the Fréchet-like distance between two merge trees. I also prove the relation between the interleaving distance and the modified Fréchet-like distance between two merge trees in this section. Section 6 is the conclusion.

2 Distance between Trees

Distance between trees is one of the topics that has been discussed in the previous years [2, 5, 13, 14, 15]. The tree edit distance and the tree alignment distance are two well-known distances which were defined between trees [13]. Both the tree edit distance and the tree alignment distance between two trees are MAX SNP-hard to compute. There is a polynomial time algorithm for computing the tree alignment distance between two ordered trees[2] if we bound the degree of each node, however there is no known polynomial algorithm for finding the edit distance between ordered trees with bounded degrees. There is a polynomial algorithm for computing the tree edit distance between trees if we consider trees with bounded depth [13].

Definition 3. *Tree edit distance [13].*

[1] $x < y$ is that x is a descendant of y

[2] Ordered tree is a rooted tree that there is an order between the children of each node [13].

Consider two labeled trees T_1 and T_2. The tree edit distance is the minimum cost of changing one tree to another one by using three editing operations add, remove and rename.

Now, the definition of the tree alignment distance is as follows:

Definition 4. *Tree alignment distance [13].*
Consider two labeled trees T_1 and T_2. The alignment distance between the two trees is obtained as follows: first I add nodes to T_1 and T_2 that the modified trees T_1' and T_2' have the same structures. The related cost would be the the cost of changing the labels that two trees T_1' and T_2' have also same labels. The minimum cost related to the best structural changes is the alignment distance.

Two following notes are satisfied about the tree edit distance and tree alignment distance from [13] and [15] respectively.

Note 1. *[13] Tree alignment distance is always greater than or equal to tree edit distance. For more illumination, look at Figure 2 (b).*

Note 2. *[15] Although there is a polynomial time algorithm for finding tree alignment distance between ordered labeled trees, tree alignment distance cannot capture similarities between trees. Figure 2 (a) illustrates this better.*

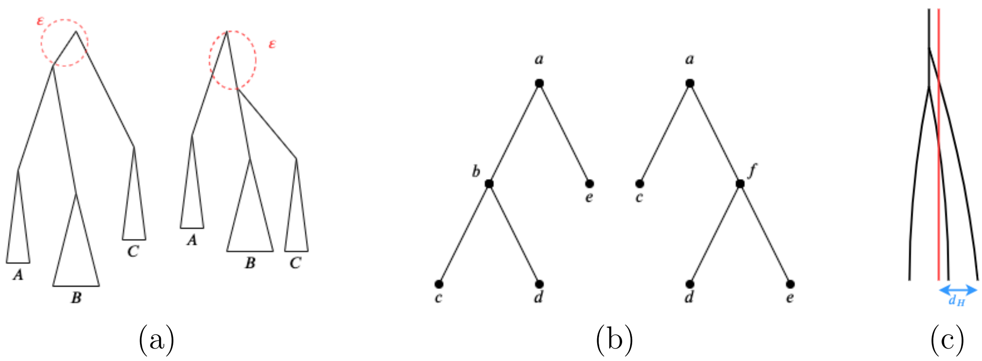

Figure 2: (a) Two trees are very similar to each other, but the alignment distance between them is very large, because tree alignment distance is sensible to the relationship between children and parents. (b) If the costs of relabeling, removing and adding nodes is 1, tree edit distance between two trees is 2, and tree alignment distance between them is 4. (c) Two trees (red color and black color trees) are completely different, however the Hausdorff distance between them is small.

Another distance that we can consider between trees is Hausdorff distance [6]. Hausdorff distance is defined between two sets of points. The Hausdorff distance is defined as follows:

Definition 5. *Hausdorff Distance [6].*
For given sets S_1 and S_2 in a space, for each point s in S_1 we find the closest point to it in S_2 (as s'), and for each point in S_2 we find the closest point to it in the set S_1. The

Hausdorff distance is the maximum over all distances that we find. The mathematical definition of Hausdorff distance is as follows:

$$d_H(S_1, S_2) = \max\left\{\sup_{s\in S_1}\inf_{s'\in S_2} d(s,s'), \sup_{s'\in S_2}\inf_{s\in S_1} d(s,s')\right\}.$$

If we consider the underlying space of trees on $\mathbb{R}^2$, we can define Hausdorff distance between two trees. However, the Hausdorff distance cannot capture dissimilarities between trees. For example in Figure 2 the two trees are very different, however the Hausdorff distance between them is very small.
Another distance that we can consider between trees is interleaving distance[14, 15]. Interleaving distance is defined between merge trees. Interleaving distance between two merge trees T_1^f and T_2^g is defined by two continuous functions α and β and the definition is as follows:

Definition 6. *[2, 14, 15] Interleaving distance between two merge trees T_1^f and T_2^g is defined as follows:*

$$d_I(T_1^f, T_2^g) = inf\{\delta \text{ s.t. there is a pair of } \delta\text{-compatible maps between } T_1^f \text{ and } T_2^g\},$$

where two continuous maps $\alpha_\delta : |T_1^f| \longrightarrow |T_2^g|$ and $\beta_\delta : |T_2^g| \longrightarrow |T_1^f|$ are δ-compatible if and only if the following conditions are satisfied:
(1) For all $u \in |T_1^f|$, $g(\alpha_\delta(u)) = f(u) + \delta$,
(2) For all $v \in |T_2^g|$, $f(\alpha_\delta(v)) = g(v) + \delta$,
(3) For all $u_1, u_2 \in |T_1^f|$ s.t. $f(u_1) = f(u_2)$, $\beta_\delta o \alpha_\delta(u_1) = \beta_\delta o \alpha_\delta(u_2)) = u_1^{2\delta}$,
(4) For all $v_1, v_2 \in |T_2^g|$ s.t. $g(v_1) = g(v_2)$, $\alpha_\delta o \beta_\delta(v_1) = \alpha o \beta_\delta(v_2) = v_1^{2\delta}$.

In [2], P. K. Agrawal, etc., proved that it is NP-hard to compute interleaving distance between two merge trees and it concludes the fact that it it NP-hard to compute the Gromov-Hausdorff distance between trees within a factor of better than 3. Later in 2019 E. Farahbakhsh and Y. Wang [15] defined one ε-good map from T_1^f to T_2^g which is defined as follows:

Definition 7. *[15] A map $\alpha^\delta : |T_1^f| \longrightarrow |T_2^g|$ is called δ-good map if and only if the following conditions are satisfied:*
(C1) α^δ is continuous,
(C2) For every point $u \in |T_1^f|$, $g(\alpha^\delta(u)) = f(u) + \delta$,
(C3) For every pair of points $v_1 = \alpha^\delta(u_1)$ and $v_2 = \alpha^\delta(u_2)$, if $v_1 \geq v_2$, $u_1^{2\delta}$[3] $\geq u_2^{2\delta}$,
(C4) If there is a point $v \in |T_2^g|$ which is not in the image of α^δ, $f(v^F$[4]$) - f(v) \leq 2\delta$.

and by the definition of δ-good map, they proved the following Theorem:

Theorem 1. *[15] $d_I(T_1^f, T_2^g) \leq \delta$ if and only if there is a δ-good map $\alpha^\delta : |T_1^f| \longrightarrow |T_2^g|$.*

[3] $u_1^{2\delta}$ is an ancestor of u_1 in T_1^f such that $f(u_1^{2\delta}) - f(u_1) = 2\delta$
[4] v^F is the nearest ancestor of v such that v^F is in the image of α^δ.

3 Fréchet-Like Distance between two Rooted Trees

In Section **??**, we defined the Fréchet like distance intuitively. In this section, I write the mathematical definition of the Fréchet-Like distance: Given two rooted trees T_1 and T_2 rooted at u and v respectively, the definition of Fréchet-Like distance is as follows:

Definition 8. ***Fréchet-Like Distance***
For two given rooted trees T_1 and T_2, I define Fréchet-Like distance as follows:

$$d_{FL}(T_1^f, T_2^g) := \min_{R \in \mathcal{R}} \sup_{(x,y) \in R} d(x, y)$$

$d(x, y)$ is the Euclidean distance between two points x and y and the correspondence $R \subseteq |T_1^f| \times |T_2^g|$ is defined as follows:

1) $\forall x \in |T_1|$, $\exists y \in |T_2|$ s.t. $(x, y) \in R$
1-i) $\forall y \in |T_2|$, $\exists x \in |T_1|$ s.t. $(x, y) \in R$
2) If $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$ and $x_2 \geq x_1$ and $y_2 \geq y_1$ then
2-i) $\forall x$ s.t. $x_1 \leq x \leq x_2$, $\exists y$ s.t. $y_1 \leq y \leq y_2$ and $(x, y) \in R$ and
2-ii) $\forall y$ s.t. $y_1 \leq y \leq y_2$, $\exists x$ s.t. $x_1 \leq x \leq x_2$ and $(x, y) \in R$.
3) If $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$ then $(x_1 \sim x_2$[5]$, y_1 \sim y_2) \in R$.
4) If $x \in |T_1|$ is a leaf, there should be a leaf $y \in |T_2|$ such that $(x, y) \in R$, unless there is a y' such that $(x, y') \in R$ and $(x^N$[6]$, y') \in R$.
4-i) If $y \in |T_2|$ is a leaf, there should be a leaf $x \in |T_1|$ such that $(x, y) \in R$, unless there is a x' such that $(x', y) \in R$ and $(x', y^N) \in R$.

4 Approximation of the Fréchet-Like Distance is in NP-hard

In this section I prove that computing the Fréchet like distance between two rooted trees is SNP-hard to compute by a reduction from UNRESTRICTED-PARTITION. The way that I prove that it is in SNP-hard is very similar to proving that Gromov-Hausdorff distance between two merge trees is in SNP-complete. [2]

UNRESTRICTED-PARTITION.
Input: a multiset of positive integers $\mathcal{X} = \{a_1, ..., a_n\}$ such that $n = 3k$,
Output: Is there a partition of $\mathcal{X}$ into k multisets $\mathcal{X}_1, ..., \mathcal{X}_m$ such that for each multiset $\mathcal{X}_j$ if we consider by S_j the summation of elements in multiset $\mathcal{X}_j$, $S_j = (\sum_{i=1}^{n} a_i)/m$? [11]

Theorem 2. *The problem UNRESTRICTED-PARTITION is in SNP-complete.*

[5] $x_1 \sim x_2$ is the nearest ancestor of x_1 and x_2
[6] x^N the nearest node which is an ancestor of x

Proof. See ([11]). □

For proving the hardness of the approximation, I construct two special rooted trees such that both of them are merge trees T_1^f and T_2^g as it shown in Figure 3. Two trees are hanging out from their root, it means that we can construct them in just one dimension and the Hausdorff distance between them is 0. Their edges and nodes are separated just to be shown better in the picture. Also, A and B are two large numbers.

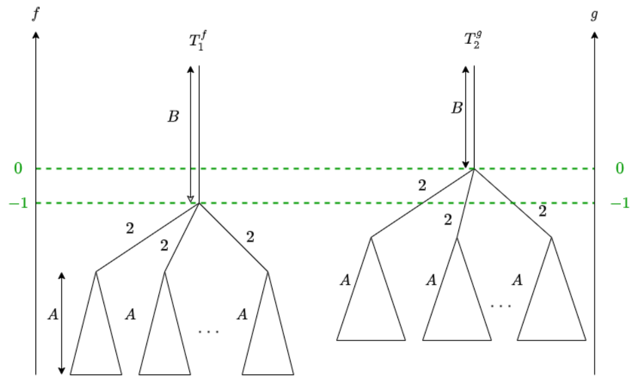

Figure 3: Two trees T_1^f and T_2^g. A and B are two large numbers.

Now, by using Figure 3, I prove the hardness of approximation of Fréchet-Like distance by the following lemmas.

Lemma 1. $d_{FL}(f, g) \leq 1$ *if UNRESTRICTED-PARTITION is a yes instance.*

Proof. If UNRESTRICTED-PARTITION is a yes instance, I can construct a correspondence $R \subseteq |T_1^f| \times |T_2^g|$ such that $\sup_{(x,y)\in R} |f(x) - g(y)| \leq 1$. If UNRESTRICTED-PARTITION is a yes instance, I can partition $\mathcal{X}$ into $X_1, X_2, ..., X_n$ such that $S(X_1) = \frac{S(\mathcal{X})}{k}$, and $X_i = \{a_{i,1}, ..., a_{i,k_i}\}$. Therefore, I map sub-trees rooted at $\{u_{i_1}, ..., u_{i_{k_i}}\}$ to v_i, such that u_{i_j} corresponds to $a_{i,j}$ in the construction of the tree and v_i corresponds to X_i. When I say that I map a point $x \in |T_1^f|$ to a point $y \in |T_2^g|$, we mean that $(x, y) \in R$ If $(u_{i_j}, v_i) \in R$ and $(u_{i_k}, v_i) \in R$, I have that $(u_r, v_i) \in R$. (For more illustration look at Figure 3) Therefore, I could construct a correspondence $R \subseteq |T_1^f| \times |T_2^g|$ such that $\sup_{(x,y)\in R} |f(x) - g(y)| < 1$. □

Lemma 2. *If UNRESTRICTED-PARTITION is a no instance,* $d_{FL}(f, g) \geq 3$.

Proof. If UNRESTRICTED-PARTITION is a no instance, as edges with the length of A are too large, we have to find a correspondence R such that for any pair of points $x_1, x_2 \in T_1^f$ such that $x_1 \| x_2$ [7], there are two different points $y_1, y_2 \in T_2^g$ such that $(x_1, y_1) \in R$, and $(x_2, y_2) \in R$. Therefore, the best correspondence that I can find with the conditions of the Definition 8 is that $x_1 \sim x_2$ map to two different point y_1 and y_2 as shown in Figure 5. Which indicates that the Fréchet distance between T_1^f and T_2^g cannot be smaller than 3. □

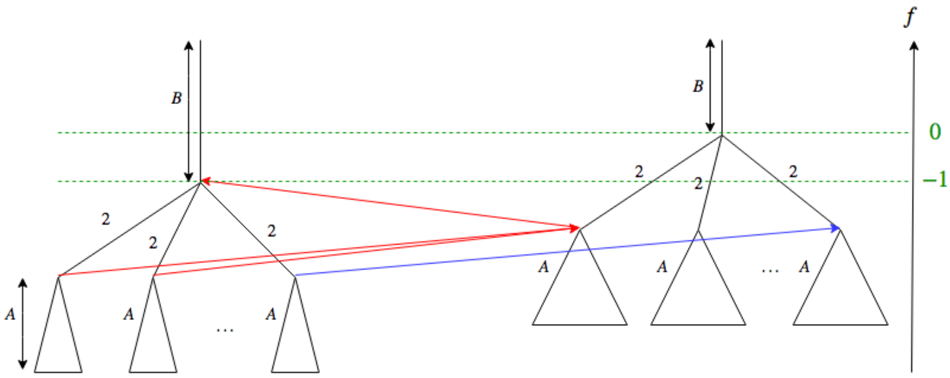

Figure 4: if UNRESTRICTED-PARTITION is a yes instance.

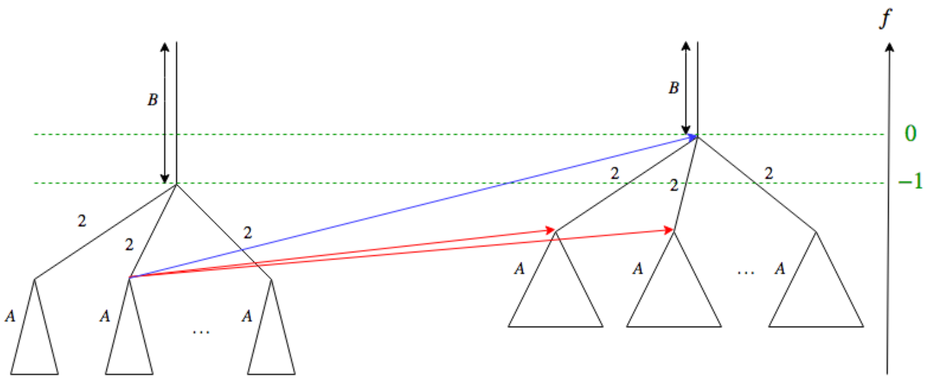

Figure 5: if UNRESTRICTED-PARTITION is a no instance.

From two mentioned lemmas, we can conclude the following result.

Corollary 1. *Computing a $(3-\epsilon)$-approximation of the Fréchet-Like distance between two merge trees T_1^f and T_2^g is NP-complete,*

5 Fréchet-Like distance and the interleaving distance

In this section we define a special kind of Fréchet-Like distance between two merge trees, which we call Fréchet-Like distance between merge trees. Given two merge trees T_1^f and T_2^g rooted at u and v respectively, the definition of Fréchet-Like distance is defined as the minimum distance between the function value of the point that a man is and the function value of the point that the monitored man is located. In the following we write the mathematical definition of modified Fréchet-Like distance between two merge trees:

Definition 9. *Modified Fréchet-Like Distance*
For two given merge trees T_1^f and T_2^g, we define modified Fréchet-Like distance as follows:

$$d_{FL}^M(T_1^f, T_2^g) := \min_{R \in \mathcal{R}} \sup_{(x,y)\in R} |f(x) - g(y)|$$

and the correspondence $R \subseteq |T_1^f| \times |T_2^g|$ is defined similar to the Definition 6 which is as follows:
1) $\forall x \in |T_1|$, $\exists y \in |T_2|$ s.t. $(x, y) \in R$
1-i) $\forall y \in |T_2|$, $\exists x \in |T_1|$ s.t. $(x, y) \in R$
2) If $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$ and $x_2 \geq x_1$ and $y_2 \geq y_1$ then
2-i) $\forall x$ s.t. $x_1 \leq x \leq x_2$, $\exists y$ s.t. $y_1 \leq y \leq y_2$ and $(x, y) \in R$ and
2-ii) $\forall y$ s.t. $y_1 \leq y \leq y_2$, $\exists x$ s.t. $x_1 \leq x \leq x_2$ and $(x, y) \in R$.
3) If $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$ then $(x_1 \sim x_2, y_1 \sim y_2) \in R$.
4) If $x \in |T_1|$ is a leaf, there should be a leaf $y \in |T_2|$ such that $(x, y) \in R$, unless there is a y' such that $(x, y') \in R$ and $(x^N, y') \in R$.
4-i) If $y \in |T_2|$ is a leaf, there should be a leaf $x \in |T_1|$ such that $(x, y) \in R$, unless there is a x' such that $(x', y) \in R$ and $(x', y^N) \in R$.

By the following lemma, we prove the relation between the Fréchet-like distance and the interleaving distance between merge trees.

Lemma 3. *If there exists an ε such that $d_{FL}^M(T_1^f, T_2^g) \leq \varepsilon$, then $d_I(T_1^f, T_2^g) \leq \varepsilon$.*

Proof. For proving this lemma we need to find an ε-good map $\alpha_\varepsilon : |T_1^f| \longrightarrow |T_2^g|$ such that three conditions in the definition of ε-good map are satisfied. First, we consider the ε-good map α^ε as follows:
As the Fréchet-Like distance between T_1^f and T_2^g is not greater than ε, based on the Definition 9 there is a correspondence R such that four conditions of the Definition 9 are

[7] $x_1 \| x_2$ if $x_1 \nleq x_2$ nor $x_2 \nleq x_1$

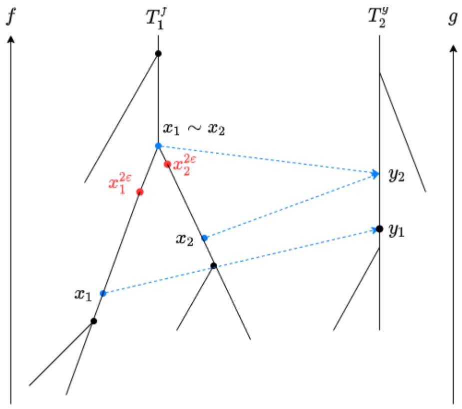

Figure 6: $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$. $y_2 > y_1$ therefore $(x_1 \sim x_2, y_2) \in R$, but $f(x_1 \sim x_2) - g(y_2) > \varepsilon$.

satisfied. Now, for constructing of the ε-good map for any pair of points $(x, y) \in R$ if $g(y) = f(x) + \varepsilon$, we map the point x to y, in another words $\alpha^{\varepsilon}(x) = y$. Otherwise, if $g(y) < f(x) + \varepsilon$, we map x to a point y' such that $y \leq y'$ and $g(y') = f(x) + \varepsilon$, it means that $\alpha_{\varepsilon}(x) = y'$.
Now, we need to prove that α_{ε} is an ε-good map. To do so, we need to prove that four conditions of the Definition 7 for the map α_{ε} are satisfied.
C1. We need to prove that map α_{ε} is continuous. To do so, we use the similar method as is written in [15].
C2. Based on the construction of the map α^{ε} for any pair of points $(x, y) \in R$ we map x to a point which is ε distance higher than x. As for all the point x in $|T_1^f|$ there is at least one y such that $(x, y) \in R$, we can conclude that for all the point x in $|T_1^f|$, $g(\alpha^{\varepsilon}(x)) = f(x) + \varepsilon$, which satisfies the condition (C2) of the Definition 7.
C3. If two pairs of points $(x_1, y_1) \in R$ and $(x_2, y_2) \in R$, and $y_1 \leq y_2$, we know that $g(y_1) \leq g(y_2)$. Therefore, based on the construction of the map α^{ε}, we have that $f(x_1) \leq f(x_2)$. Two cases can happen:
Case1: $x_1 \leq x_2$, which in this case we have that $x_1^{2\epsilon} \leq x_2^{2\epsilon}$.
Case2: $x_1 \| x_2$, in this case if by contradiction $x_1^{2\epsilon} \not\leq x_2^{2\epsilon}$, therefore we have that $x_1^{2\epsilon} \| x_2^{2\epsilon}$ as $f(x_1) \leq f(x_2)$. Based on the definition of Fréchet-Like distance the highest y such that $(x_2, y) \in R$ is y_2. Therefore by using the condition 3 of the Fréchet-Like distance the highest point $y \in T_2^g$ that $(x_1 \sim x_2, y) \in R$ is y_2 and $f(x_1 \sim x_2) - g(y_2) > \varepsilon$. It is a contradiction with the fact that the Fréchet-Like distance between T_1^f and T_2^g is less than or equal to ε. For more illustration take a look at Figure 6.
C4. If there is a point $y \in T_1^f$ such that there is no $x \in T_2^g$ map to y under the map α_{ε}, as we already proved in C1 that the map is continuous, the point should be a branch connects a leaf (For example y^L) to the tree, and none of the point $y' \leq y$ are in the image of the map α_{ε}. Now, I just need to prove that $g(y^F) - g(y) \leq 2\varepsilon$. By contradiction if $g(y^F) - g(y) > 2\varepsilon$ and x is the point that $(x, y^F) \in R$ based on the definition of Fréchet-

Like distance condition 4, x is a leaf. Therefore, $(x, y^L) \in R$ and it is a contradiction by the fact that $d^M_{FL}(T_1^f, T_2^g) \leq \varepsilon$. □

6 Concluding Remarks

In this paper, I extended the definition of Fréchet distance between two curves to the Fréchet-Like distance between two rooted trees. In section 2, I discussed some distances that have been defined between two trees. I defined a new definition for computing the similarity between two trees in Section 3. I called the new distance, Fréchet-Like distance because of the similarity of the definition to Fréchet distance between curves. The hardness of approximation was discussed later in Section 4. Here, we also proved that although there is a polynomial time algorithm for computing the Fréchet distance between polygonal curves [4], it is NP-hard to approximate Fréchet-Like distance between two trees even if we consider merge trees. The relation between Fréchet-Like distance between two merge trees and the interleaving distance was discussed in section 5.

References

[1] Agarwal, P. K., Avraham, R. B., Kaplan, H., and Sharir, M. Computing the discrete Fréchet distance in subquadratic time. *SIAM J. Comput. 43* (2013), 429–449.

[2] Agarwal, P. K., Fox, K., Nath, A., Sidiropoulos, A., and Wang, Y. Computing the Gromov-Hausdorff distance for metric trees. In *ISAAC* (2015).

[3] Akitaya, H. A., Buchin, M., Ryvkin, L., and Urhausen, J. The k-Fréchet distance. *CoRR abs/1903.02353* (2019).

[4] Alt, H., and Godau, M. Computing the frchet distance between two polygonal curves. *Int. J. Comput. Geometry Appl. 5* (1995), 75–91.

[5] Bille, P. A survey on tree edit distance and related problems. *Theor. Comput. Sci. 337*, 1-3 (June 2005), 217–239.

[6] Buchin, K., Buchin, M., and Wenk, C. Computing the Fréchet distance between simple polygons in polynomial time. *Proceedings of the Annual Symposium on Computational Geometry 2006* (01 2006), 80–87.

[7] Buchin, M., Driemel, A., and Speckmann, B. Computing the Fréchet distance with shortcuts is NP-hard. In *Symposium on Computational Geometry* (2013).

[8] Eiter, T., and Mannila, H. Computing discrete Fréchet distance. Tech. Rep. CD-TR 94/64, CD-Laboratory for Expert Systems, TU Vienna, Austria, May 1994.

[9] Ewing, G. *Calculus of Variations with Applications*. Dover Books on Mathematics. Dover Publications, 1985.

[10] Fréchet, M. M. Sur quelques points du calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo (1884-1940) 22*, 1 (Dec 1906), 1–72.

[11] Garey, M. R., and Johnson, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences)*, first edition ed. W. H. Freeman, 1979.

[12] Godau, M. *On the complexity of measuring the similarity between geometric objects in higher dimensions*. PhD thesis, Freie Universitt Berlin, 1999.

[13] Jiang, T., Wang, L., and Zhang, K. Alignment of trees — an alternative to tree edit. In *Combinatorial Pattern Matching* (Berlin, Heidelberg, 1994), M. Crochemore and D. Gusfield, Eds., Springer Berlin Heidelberg, pp. 75–86.

[14] Morozov, D., Beketayev, K., and Weber, G. H. Interleaving distance between merge trees. In *Workshop on Topological Methods in Data Analysis and Visualization: Theory, Algorithms and Applications* (2013).

[15] Touli, E. F., and Wang, Y. FPT-algorithms for computing Gromov-Hausdorff and interleaving distances between trees. In *ESA* (2018).