УДК 512.5

O. Berezsky, M. Zarichnyi

RÉCHET DISTANCE BETWEEN WEIGHTED ROOTED TREES

O. Berezsky, M. Zarichnyi. *Fréchet distance between weighted rooted trees*, Mat. Stud. **48** (2017), 165–170.

The aim of this note is to extend the notion of Fréchet distance over the set of weighted rooted trees. The weighted trees naturally appear as skeletons of planar domains. The defined distance allows for defining a distance between (weighted) threes, which is merely a symmetric, i.e., does not necessarily satisfy the triangle inequality.

1. Introduction. Comparison of images is widely used in different areas of artificial intelligence, e.g., in pattern recognition, robototechnics, computer vision. Algorithms of signal shapes comparison are also used in natural language processing as well as CAD-systems. The shape of planar objects is often described by their contours, which, in general, are represented by curves. Therefore, if is important to evaluate the dissimilarity between curves. One of classical tools for that is the so-called Fréchet distance introduced in [10] (see also [11]).

The Fréchet distance and its generalizations are investigated in numerous papers (see, e.g., [3, 12, 16, 19]). In particular, in the important paper [3], an algorithm for computing the Fréchet distance between polygonal curves can be computed in $O(mn\log(mn))$ time.

In [5], the Freḉhet distance is considered for the curves which are sequences of cured pieces and it is shown that, under some mild conditions, this distance can be computed in $O(mn))$ time, where m, n is the number of pieces in these curves.

In the paper of Schlesinger, Vodolazskiy and Yakovenko [13], an algorithm is discussed that determines whether the Fréchet distance between two closed polynomial curves is less than a given number. The algorithm takes $O(mn)$ time, where m, n are the numbers of vertices.

A. Mosig and Clausen [15] introduced approximate algorithms for computing the Fréchet distance between two polynomial curves. They and other authors [16] also defined a discrete version of the Fréchet distance.

The computation of Fréchet distances between polygonal curves whose vertices are imprecise (i.e., are given by regions in which they lie) is made in [17]).

The notion of Fréchet distance was also extended over more complicated geometric objects, e.g., surfaces (see [2]).

The Fréchet distances can be naturally defined for oriented and closed curves. In this paper, we consider this notion for the so-called weighted trees [4]. These trees naturally appear as skeletons middle sets of planar domains; the weight of a point is then the distance from this point to the boundary of the domain. Remark that the skeletons are widely used

2010 *Mathematics Subject Classification:* 37E25, 54E35.
Keywords: group; semigroup; maximal linked family; superextension; automorphism group.
doi:10.15330/ms.48.2.165-170

in computer vision. Our approach is based on combining both the Fréchet distance between oriented curves and the coupling between the sets of end vertices; the latter is related to some metrics such as the Hausdorff metric on sets and the Kantorovich metric (see, e.g., [20]) on measures. We obtain a metric in the case of rooted trees; an attempt to extend our considerations over the unrooted case leads only to a function which is a symmetric (in general, it does not satisfy the triangle inequality). The Hausdorff metric and the Fréchet metric are used by the authors for the quantitative estimations of the image segmentation algorithms (see [5]). The latter was successfully applied to analysis of biomedical images [6].

Remark that the introduced metric on the set of weighted rooted trees induces a symmetric on the set of non-rooted trees.

2. Preliminaries. We assume that all the maps under consideration are continuous.

2.1. Metric functions. Recall that a metric on a a set X is a function $d\colon X \times X \to \mathbb{R}$ satisfying the properties: (1) $d(x, y) \geq 0$, $d(x, y) = 0$ if and only if $x = y$; (2) $d(x, y) = d(y, x)$, and (3) $d(x, y) \leq d(x, z) + d(z, y)$.

If we drop property (3) in this definition, we obtain the notion of symmetric on X (see, e.g., [8]).

2.2. Fréchet metric. In the sequel we will need the notion of Fréchet metric between oriented curves in a metric space (X, d) (see, e.g., [10]). Given oriented curves $\gamma_i\colon [0, 1] \to X$, $i = 1, 2$, define

$$d_F(\gamma_1, \gamma_2) = \inf\{d(\gamma_1(\alpha(t)), \gamma_2(\alpha(t))) \mid \alpha\colon [0, 1] \to [0, 1] \text{ is an increasing homeomorphism}\}.$$

The function d_F is known to be a metric (the Fréchet metric).

2.3. Trees. A tree is a connected graph without cycles. A rooted tree is a pair (T, t_0), where T is a tree and $t_0 \in T$.

A *weighted rooted tree* is a triple (T, t_0, w), where (T, t_0) is a rooted tree and $w\colon T \to \mathbb{R}_+$ is a function (weight).

By $e(T)$ we denote the set of all endpoints of T. For any $a \in e(T)$, by $L_a = L_a^T$ we denote the (unique) geodesic connecting t_0 and e.

Given two sets, A and B, we say that a subset $C \subset A \times B$ is a *coupling* between A and B if the following holds:

1. for any $a \in A$, there exists $b \in B$ such that $(a, b) \in C$;
2. for any $b \in B$, there exists $a \in A$ such that $(a, b) \in C$.

If C' is a coupling between A_1 and A_2 and C'' is a coupling between A_2 and A_3, then the subset

$$C = C' \circ C'' = \{(a_1, a_3) \in A_1 \times A_3 \mid (a_1, a_2) \in C',\ (a_2, a_3) \in C'' \text{ for some } a_2 \in A_2\}$$

is a coupling between A_1 and A_3.

An oriented curve in a metric space X is a map $\gamma\colon [0, 1] \to X$. We identify γ and $\gamma\alpha$, for every orientation-preserved homeomorphism $\alpha\colon [0, 1] \to [0, 1]$.

Suppose that (X, d) is a metric space. The *Fréchet distance* between curves $\gamma_1, \gamma_2\colon [0, 1] \to X$ is

$$d_F(\gamma_1, \gamma_2) = \inf\{\sup\{d(\gamma_1(\alpha(t)), \gamma_2(t)) \mid t \in [0, 1]\} \mid$$

$$\alpha\colon [0,1] \to [0,1] \text{ is an orientation-preserved homeomorphism}\}.$$

Given a curve L in a metric space X and a function $w\colon L \to \mathbb{R}_+$, let $\check{L} = \{(x, w(x)) \mid x \in L\}$. Then $\check{L}$ is a curve in the space $\mathbb{R}^3$.

2.4. Skeletons. For the sake of simplicity, we consider here only subsets in the Euclidean plane $\mathbb{R}^2$. By d we denote the Euclidean metric. Given a nonempty subset $A \subset \mathbb{R}^2$ and $x \in \mathbb{R}^2$, we let $d(x, A) = \inf\{d(x, a) \mid a \in A\}$.

Let D be a bounded domain in $\mathbb{R}^2$. We denote by ∂D its boundary. For any $x \in D$, denote by $p(x)$ its metric projection,

$$p(x) = \{y \in \partial D \mid d(x, y) = d(x, \partial D)\}.$$

The *skeleton* (or *central set*) of D is the set

$$\operatorname{sk} D = \{x \in D \mid \#p(x) \geq 2\},$$

where $\#A$ stands for the cardinality of A.

See, e.g., [9] for metric and topological properties of skeletons of open sets in Euclidean spaces.

There are domains whose skeletons demonstrate irregular behavior (see, e.g., [7]). However, this is not the case for domains that appear in applications. In particular, it is known that the skeleton of any domain with a smooth enough boundary is a deformation retract of it. Moreover, the skeletons of polyhedral simple connected domains are trees.

The function $w\colon \operatorname{sk} D \to \mathbb{R}_+$ defined by the formula $w(x) = d(x, \partial D)$, is a weight on $\operatorname{sk} D$.

3. Results. We consider the trees embedded into the plane $\mathbb{R}^2$. Clearly, our considerations can be easily extended over any euclidean space.

For any two weighted rooted trees (T, t_0, w) and (S, s_0, u), we let

$$D_F((T, t_0, w), (S, s_0, u)) = \min\{\max\{d_F(\check{L}^T_a, \check{L}^S_b) \mid (a, b) \in C\} \mid C \text{ is a coupling between } e(T) \text{ and } e(S)\}.$$

Theorem 1. *The function D_F is a metric on the set of weighted rooted trees embedded in $\mathbb{R}^2$.*

Proof. Clearly, if $D_F((T, t_0, w), (S, s_0, u)) = 0$, then $(T, t_0, w) = (S, s_0, u)$.

It is also obvious that the function D_F is symmetric.

We are going to prove the triangle inequality for D_F. Let $(T, t_0, w), (S, s_0, u), (R, r_0, v)$ be weighted rooted trees. There exist a coupling C between $e(T)$ and $e(S)$ (resp. C' between $e(S)$ and $e(R)$) such that

$$D_F((T, t_0, w), (S, s_0, u)) = \max\{d_F(\check{L}^T_a, \check{L}^S_b) \mid (a, b) \in C\}$$

(resp.

$$D_F((S, s_0, u), (R, r_0, v)) = \max\{d_F(\check{L}^S_b, \check{L}^R_c) \mid (b, c) \in C'\}).$$

Define $C'' = C \circ C'$. Then

$$D_F((T, t_0, w), (R, r_0, v)) \leq \max\{d_F(\check{L}^T_a, \check{L}^R_c) \mid (a, c) \in C''\}.$$

There exists $(a',c')\in C''$ such that $D_F((T,t_0,w),(R,r_0,v))=d_F(\check L^T_{a'},\check L^R_{c'})$.

Then there exists $b\in e(S)$ such that $(a',b)\in C$, $(b,c')\in C'$. We conclude that

$$\begin{gathered}D_F((T,t_0,w),(R,r_0,v))=d_F(\check L^T_{a'},\check L^R_{c'})\le d_F(\check L^T_{a'},\check L^S_{b})+d_F(\check L^S_{b},\check L^T_{a'})\le\\ \le\max\{d_F(\check L^T_{a},\check L^S_{b'})\mid (a,b')\in C\}+\max\{d_F(\check L^S_{b'},\check L^R_{c})\mid (b',c)\in C'\}\le\\ \le D_F((T,t_0,w),(S,s_0,u)).\end{gathered}$$

□

The definition of the distance between weighted rooted trees allows us to define the distance between rooted trees. One should simply put $w\equiv 1$. Note that, as in the case of curves, the distance between rooted trees is estimated by the Hausdorff distance from below.

Remark that, given a tree T, one can regard arbitrary its point as a root. This allows for modifying the Fréchet distance as follows

$$\tilde D_F((T,w),(S,u))=\inf\{D_F((T,t_0,w),(S,s_0,u))\mid t_0\in T,\ s_0\in S\}.$$

Theorem 2. *The function $\tilde D_F((T,w),(S,u))$ is a symmetric on the set of weighted trees.*

Proof. Clearly, $\tilde D_F((T,w),(S,u))\ge 0$ for all weighted trees (T,w) and (S,u). Suppose now that $\tilde D_F((T,w),(S,u))=0$. Given $r>0$ and $a\in e(T)$, there exists $b(r)\in e(S)$ such that r-neighborhood of $\check L^T_a$ (with respect to the Fréchet metric) contains $\check L^S_{b(r)}$. Compactness of S implies that $\check L^T_a=\check L^S_b$, for some $b\in e(S)$. Therefore, $T\subset S$ and $u|T=w$. One can similarly prove that $S\subset T$ and $w|S=u$, whence the equality follows.

The symmetry of $\tilde D_F$ (i.e., property (2) from the definition of symmetric) is obvious. □

The following example demonstrates that, in general, the function $\tilde D_F$ is not a metric, i.e. it does not satisfy the triangle inequality.

Example 1. We suppose that the weights on the trees under considerations are identically 1 and, for the sake of simplicity, we drop the notation for the weight.

Let

$$T=\{(x,y)\in\mathbb R^2\mid y=r|x|-1,\ |y|\le 1\},\ S=\{(0,y)\mid |y|\le 1\},\ R=\{(x,-y)\mid (x,y)\in T\}$$

(we suppose here that $r>0$ is small enough).

Then, clearly, if we regard $(0,-1)$ as the root in T and S, we obtain

$$\tilde D_F(S,T)\le D_F((S,(0,-1)),(T,(0,-1)))\le r.$$

One can similarly show that also $\tilde D_F(S,R)\le r$. However, simple geometric considerations allow us for demonstrating that, for any choice of roots t_0 and r_0 in T and R respectively and any matching $C\subset e(T)\times e(R)$, there exists $(t,r)\in C$ such that the Fréchet distance between $[t_0,t]$ and $[r_0,r]$ is at least 1. Therefore $\tilde D(T,S)\ge 1$ and, if $r<1/2$, we obtain

$$\tilde D_F(T,R)\ge 1>2r\ge \tilde D_F(S,T)+\tilde D_F(S,R),$$

which contradicts the triangle inequality.

A polygonal tree in $\mathbb R^2$ is a tree which is a finite union of linear segments.

The following is an easy consequence of properties of the distance d_F.

Proposition 1. *The set of all polygonal trees is dense (with respect to the metric D_F) in the space of all trees.*

Remark 1. From the definition it easily follows that the distance D_F coincides with the Fréchet distance d_F on the set of oriented curves. Here, every oriented curve $\gamma\colon [0,1] \to X$ can be considered as a rooted tree with the root $\gamma(0)$.

Remark 2. Like in the case of curves, the topology induced by the metric D_F does not coincide with the topology induced by the Hausdorff metric d_H. This easily follows from the above remark.

Let $\mathcal{T}_n$ denote the set of trees T with $\#e(T) \leq n$. Again, in order to simplify the exposition we assume that the weight is identically 1 and drop the notation for the weight.

Proposition 2. *The set $\mathcal{T}_n$ is closed in the set of all trees, for every $n \geq 1$.*

Proof. Let $(T, t_0) \notin \mathcal{T}_n$. Then $\#e(T) > n$ and there is $r > 0$ such that the following holds:

1. for every $t \in e(T)$, $d(t, t_0) > r$,
2. for every $t, t' \in e(T)$, $t \neq t'$, $d(t, t') > r$.

Let (S, s_0) be a tree such that $D_F((T, t_0), (S, s_0)) < r/2$. Then There exists a coupling $C \subset e(T) \times e(S)$ such that

$$D_F((T, t_0), (S, s_0)) = \max\{d_F(L_a^T, L_b^S) \mid (a, b) \in C\}.$$

For any $a \in e(T)$ there exists $b = b(a) \in e(S)$ such that $(a, b(a)) \in C$. Because of choice of r, $b(a) \neq b(a')$ whenever $a \neq a'$. Therefore, $\#e(S) \geq \#e(T) > n$. We conclude that the complement of the set $\mathcal{T}_n$ is open in the set of all trees. □

4. Concluding remarks. Note that there is a non-monotonic version of the Fréchet distance. Using the procedure as described at the beginning of Section one can define a version of the Fréchet distance.

In connection to Theorem 2 an open problem is to find a suitable metric on the set of (non-rooted) trees. This seems to be important for comparing skeletons of (planar) domains as they usually do not possess naturally defined roots.

REFERENCES

1. P.K. Agarwal, R.B. Avraham, H. Kaplan, M. Sharir, *Computing the discrete Fréchet distance in subquadratic time*, SIAM J. Comput., **43** (2014), №2, 429–449.
2. H. Alt, M. Buchin, *Can we compute the similarity between surfaces?* Discrete Comput. Geom., **43** (2010), №1, 78–99.
3. H. Alt, M. Godau, *Computing the Fréchet distance between two polygonal curves*, Int. J. of Computational Geometry and Applications, **5** (1995), 75–91.
4. O. Berezsky, *Frechet metric for trees*, Proceedings of the 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, (2016), 213–217.
5. O. Berezsky, K. Berezska, *Quantified estimation of image segmentation quality based on metrics*, Upravlyayuschie sistemyi i mashinyi, **6** (2015), 59–65. (in Russian)

6. O. Berezsky, O. Pitsun, *Automated processing of cytological and histological images*, Proceedings of the XIIth International Conference “Perspective Technologies and Methods in MEMS Design”, MEMSTECH’2016, Lviv-Polyana, (2016), 51–53.
7. C.J. Bishop, H. Hakobyan, *A central set of dimension 2*, Proc. Amer. Math. Soc., **136** (2008), №7, 2453–2461.
8. E. Deza, M. Deza, Encyclopedia of distances, Springer, 2014.
9. D.H. Fremlin, *Skeletons and central sets*, Proc. London Math. Soc., **74** (1997), №3, 701–720.
10. M. Fréchet, *Sur quelques points du calcul fonctionnel*, Rendiconti del Circolo Mathematico di Palermo, **22** (1906), 1–74.
11. G.M. Ewing, Calculus of variations with applications, Dover Publ., 1985.
12. G. Rote, *Computing the Fréchet distance between piecewise smooth curves*, Computational Geometry, **37** (2007), 162–174.
13. M.I. Schlesinger, E.V. Vodolazskiy, V.M. Yakovenko, *Fréchet similarity of closed polygonal curves*, International Journal of Computational Geometry & Applications **26** (2016), №1, 53–66.
14. L.J. Guibas, J.E. Hershberger, J.S.B. Mitchell, J.S. Snoeyink, *Approximating polygons and subdivisions with minimum link paths*, Internat. J. Comput. Geom. Appl., **3** (1993), №4, 383–415.
15. A. Mosig, M. Clausen, *Approximately matching polygonal curves with respect to the Fréchet distance*, Computational Geometry, **30** (2005), 113–127.
16. T. Eiter, H. Mannila, Computing discrete Frechet distance, Technical Report CDTR 94/64, Christian Doppler Laboratory for Expert Systems, TU Vienna, Austria, 1994.
17. H.-K. Ahn, Ch. Knauer, M. Scherfenberg, L. Schlipf, A. Vigneron, *Computing the discrete Frechet distance with imprecise input*, International Journal of Computational Geometry & Applications, **22** (2012), №1, 27–44.
18. K. Buchin, M. Buchin, C. Wenk, *Computing the Fréchet distance between simple polygons*, Computational Geometry, **41** (2008), 2–20.
19. Atlas F. Cook I.V., A. Driemel, J. Sherette, C. Wenk, *omputing the Frechet distance between folded polygons*, Computational Geometry, **50** (2015), 1–16.
20. C. Villani, Topics in Optimal Transportation, Amer. Math. Soc., Providence, Rhode Island, 2000.

Ternopil National Economic University
Ternopil, Ukraine
ob@tneu.edu.ua

Department of Mechanics and Mathematics
Lviv National University, Lviv, Ukraine
zarichnyi@yahoo.com

Received 23.12.2017