Proximity problems for high-dimensional data

Ioannis Psarros

Thesis Advisor: Ioannis Z. Emiris
National and Kapodistrian University of Athens

June 10, 2019

Summary

1. Random projections with false positives
2. Weak dimension reduction for doubling subsets of ℓ_1
3. High dimensional approximate r-nets
4. Approximate nearest neighbors for polygonal curves
5. Vapnik–Chervonenkis dimension for polygonal curves

1 Random projections with false positives

2 Weak dimension reduction for doubling subsets of ℓ_1

3 High dimensional approximate r-nets

4 Approximate nearest neighbors for polygonal curves

5 Vapnik–Chervonenkis dimension for polygonal curves

Problem definition

Definition (c-Approximate Nearest Neighbor (c-ANN) problem)

Given a finite set $P \subset M$, a distance function $\mathrm{d}(\cdot,\cdot)$, and an approximation factor $c > 1$, preprocess P into a data structure which supports the following type of queries:

$$\forall q \in M, \text{ find } p^* \text{ such that } \forall p \in P: \ \mathrm{d}(q,p^*) \leq c \cdot \mathrm{d}(q,p).$$

Problem definition

Definition ((c,r)-ANN Problem)

Given a finite set $P \subset M$, a distance function $\mathrm{d}(\cdot,\cdot)$, an approximation factor $c > 1$, and a range parameter r, preprocess P into a data structure which supports the following type of queries:

- if $\exists p^* \in P$ s.t. $\mathrm{d}(p^*, q) \leq r$, then it returns any point $p' \in M$ s.t. $\mathrm{d}(p', q) \leq c \cdot r$,
- if $\forall p \in P$, $\mathrm{d}(p, q) > c \cdot r$, then report "Fail".

The data structure is allowed to return either a point at distance $\leq c \cdot r$ or "Fail".

It is known that one can solve logarithmically many instances of the (c,r)-ANN problem to solve the c-ANN problem.

Problem definition

Two regimes for normed spaces.

“Low dimensional”

Time/space complexity: $exp(d)$, but “good” dependence on n.

Example: $(1+\epsilon)$-ANN in space $\tilde{O}(dn)$ and query time $O\left(\frac{1}{\epsilon}\right)^d$.

“High dimensional”

Time/space complexity: $poly(d)$, but “worse” dependence on n.

Example: $(1+\epsilon)$-ANN in space $\tilde{O}(dn^{1+\rho})$ and query time $\tilde{O}(dn^{\rho})$, where $\rho = \rho(\epsilon) < 1$.

Random Projections

Johnson-Lindenstrauss lemma

Let $X \subset \mathbb{R}^d$ and $|X| = n$. There exists a distribution over linear maps $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ with $d' = O\left(\epsilon^{-2} \log n\right)$ s.t., for any $p, q \in X$:

$$\|f(p) - f(q)\|_2 \in (1 \pm \epsilon)\|p - q\|_2,$$

with high probability.

Aiming for linear-space data structures.

"Low dimensional solution" + JL

- Space: $O(d'n)$.
- Query time: $\left(\frac{1}{\epsilon}\right)^{\Theta(\epsilon^{-2} \log n)} = \omega(n)$.

Embeddings with slack

c-kANNs problem: approximate the set of k nearest neighbors.

Definition (Anagnostopoulos, Emiris, P '15)

Let (Y, d_Y), (Z, d_Z) be metric spaces and $X \subseteq Y$. A distribution over mappings $f : Y \to Z$ is a *locality preserving embedding* with

- distortion $D \geq 1$,
- probability of correctness $P \in [0, 1]$,
- and locality/slack parameter k,

if, $\forall c \geq 1$ and $\forall q \in Y$, with probability $\geq \mathcal{P}$, when $S_q = \{f(p_1), \ldots, f(p_k)\}$ is a solution to c-kANNs for $f(q)$, then $\exists f(x) \in S_q$ such that x is a $(D \cdot c)$-ANN of q in X.

Random projections with slack

Theorem (Anagnostopoulos, Emiris, P '15)

Consider $X \subset \mathbb{R}^d$*, query* $q \in \mathbb{R}^d$ *and approximation error* $\epsilon > 0$*. Sample linear map* $f : \mathbb{R}^d \to \mathbb{R}^{d'}$ *from a JL distribution, with* $d' = \Theta\left(\epsilon^{-2} \log\left(\frac{n}{k}\right)\right)$*. Then, w.c.p. the following hold:*

- *if* p^* *is the NN of* q*, then* $\|f(p^*) - f(q)\|_2 \in (1 \pm \epsilon)\|p^* - q\|_2$*,*
- $|\{p \in X \setminus \{p^*\} : \|f(p) - f(q)\|_2 \notin (1 \pm \epsilon)\|p - q\|_2\}| \leq k$

Low dimension + JL with slack

- Query time: $\left(\frac{1}{\epsilon}\right)^{\Theta(\epsilon^{-2}\log(n/k))} + k = dn^{1-\Theta(\epsilon^2/\log(1/\epsilon))}$.
- Space: $O(dn)$.

Locality Sensitive Hashing

Random partition: near points to the same part.

Definition

Let reals $r_1 < r_2$ and $p_1 > p_2 > 0$. We call a family F of hash functions (p_1, p_2, r_1, r_2)-sensitive for a metric space $\mathcal{M}$ if, for any $x, y \in \mathcal{M}$, and h distributed randomly in F, it holds:

- $\mathrm{d}_{\mathcal{M}}(x, y) \leq r_1 \implies Pr[h(x) = h(y)] \geq p_1$,
- $\mathrm{d}_{\mathcal{M}}(x, y) \geq r_2 \implies Pr[h(x) = h(y)] \leq p_2$.

Known LSH function for several metrics.

Locality Sensitive Hashing

Transform to binary LSH function

Suppose that there exists a (p_1, p_2, r_1, r_2)-sensitive family of functions F for a metric space $\mathcal{M}$. Then, there exists a

$$\left(\frac{1+p_1}{2}, \frac{1+p_2}{2}, r_1, r_2\right)\text{-sensitive}$$

family of functions F' for $\mathcal{M}$, which maps points to $\{0, 1\}$.

Proof.

For each non-empty bucket assign a random bit. □

LSH + Random projections with slack

Theorem (P, Avarikioti, Samaras, Emiris '17)

Consider $X \subset \mathbb{R}^d$*, query* $q \in \mathbb{R}^d$ *and radius* $r > 0$*, approximation error* $\epsilon > 0$*. There exists a distribution over mappings* $f : \mathbb{R}^d \to \{0,1\}^{d'}$ *such that if* $d' = \Theta(\epsilon^{-2} \log(\frac{n}{k}))$*, then, w.c.p.*

- $\|p - q\|_2 \leq r$ *implies* $\|f(p) - f(q)\|_1 \leq r'$,
- $|\{p \in X : \|p - q\|_2 \geq (1+\epsilon)r$ *and* $\|f(p) - f(q)\|_1 \leq r'\}| \leq k$,

Low dimension Hamming + LSH projection with slack

- Query time: $2^{\Theta(\epsilon^{-2}\log(n/k))} + k = dn^{1-\Theta(\epsilon^2)}$.
- Space: $O(dn)$.

For any LSHable metric, we obtain linear space and sublinear query.

Summary

Near-linear space regime. Results whp.

	Space	Query
Entropy-based LSH [Panigrahy '06]	$\tilde{O}(dn)$	$dn^{O((1+\epsilon)^{-1})}$
Entropy-based LSH [Andoni '08]	$\tilde{O}(dn)$	$dn^{O((1+\epsilon)^{-2})}$
JL with slack	$\tilde{O}(dn)$	$dn^{1-\Theta(\epsilon^2/\log(1/\epsilon))}$
LSH tradeoffs [Andoni et al. '17]	$\tilde{O}(dn)$	$O(dn^{(2(1+\epsilon)^2-1)/(1+\epsilon)^4})$
LSH-projection with slack	$\tilde{O}(dn)$	$dn^{1-\Theta(\epsilon^2)}$

1 Random projections with false positives

2 Weak dimension reduction for doubling subsets of ℓ_1

3 High dimensional approximate r-nets

4 Approximate nearest neighbors for polygonal curves

5 Vapnik–Chervonenkis dimension for polygonal curves

Scenario

Definition (Doubling constant)

Consider any metric space (X, d_X). The *doubling constant* of X, denoted λ_X, is the smallest integer λ_X such that for any $p \in X$ and $r > 0$, the ball $B(p, r)$ can be covered by λ_X balls of radius $r/2$ centered at points in X.

The *doubling dimension* of (X, d_X) is defined to be equal to $\log \lambda_X$.

Dimension reduction hardness [Naor et al '04]

There are arbitrarily large n-point subsets $P \subseteq \ell_1$ which are doubling with constant 6, such that every embedding[1] with distortion D of P into $\ell_1^{d'}$ requires $d' = n^{\Omega(1/D^2)}$.

[1] bi-Lipschitz

Scenario

Definition (Near-neighbor preserving embedding)

Let (Y, d_Y), (Z, d_Z) be metric spaces and $X \subseteq Y$. A distribution over mappings $f : Y \to Z$ is a *near-neighbor preserving embedding* with range $r > 0$, distortion $D \geq 1$ and probability of correctness $\mathcal{P} \in [0, 1]$ if, $\forall \alpha \geq 1$ and $\forall q \in Y$, if $x \in X$ is such that $d_Y(x, q) \leq r$, then with probability at least $\mathcal{P}$,

- $d_Z(f(x), f(q)) \leq D \cdot r$,
- $\forall p \in X : \ d_Y(p, q) \geq D \cdot \alpha \cdot r \implies d(f(p), f(q)) \geq \alpha \cdot r$.

The $(D\alpha, r)$-ANN problem in Y reduces to the $(\alpha/D, Dr)$-ANN problem in Z.

Results

Theorem ([Emiris, Margonis, P '19])

For n points in ℓ_1:

1. *for every $\epsilon \in (0, 1/2]$ and $c \geq 1$, there is a near neighbor-preserving embedding $h : \ell_1^d \to \ell_1^{d'}$ that can be computed in time $\tilde{O}(dn^{1+1/\Omega(c)})$, with distortion $1{+}6\epsilon$ and probability of correctness $\Omega(\epsilon)$, where*

$$d' = (\log \lambda_P \cdot \log(c/\epsilon))^{\Theta(1/\epsilon)} / \zeta(\epsilon),$$

2. *For every $\epsilon \in (0, 1/2]$, there is a near neighbor-preserving embedding $h' : \ell_1^d \to \ell_1^{d'}$ that can be computed in time $O(dkn)$, with distortion $1{+}6\epsilon$ and probability of correctness $\Omega(\epsilon)$, where*

$$d' = (\log \lambda_P \cdot \log(d/\epsilon))^{\Theta(1/\epsilon)} / \zeta(\epsilon).$$

Results

	Target dimension	Time
[I06] [2]	$d' = (\log n)^{\Theta(1/\epsilon)}/\zeta(\epsilon)$	$O(dd'n)$
[IN07] [1] [3]	$d' = \log(1/\epsilon)\log\lambda_P/\epsilon^2$	$O(dd'n)$
[EMP19]	$d' = (\log\lambda_P \cdot \log(\mathbf{c}/\epsilon))^{\Theta(1/\epsilon)}/\zeta(\epsilon)$	$\tilde{O}(dn^{1+1/\Omega(c)})$
[EMP19]	$d' = (\log\lambda_P \cdot \log(\mathbf{d}/\epsilon))^{\Theta(1/\epsilon)}/\zeta(\epsilon)$	$O(dd'n)$

[2] Nearest-Neighbor preserving

[3] ℓ_2

1 Random projections with false positives

2 Weak dimension reduction for doubling subsets of ℓ_1

3 High dimensional approximate r-nets

4 Approximate nearest neighbors for polygonal curves

5 Vapnik–Chervonenkis dimension for polygonal curves

Problem definition

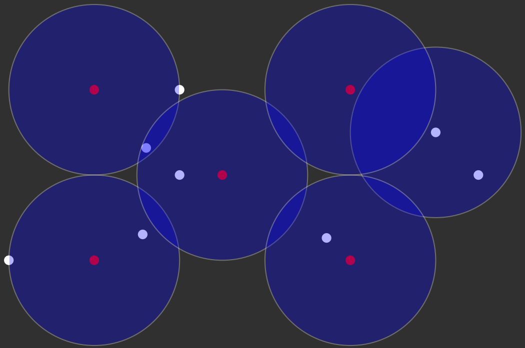
Definition

Given a pointset $X \subseteq \mathbb{R}^d$, a parameter $r > 0$, an r-net of X is a subset $N \subseteq X$ s.t. the following properties hold:

- (packing) For every $p \neq q \in N$, we have that $\|p - q\|_2 > r$.
- (covering) For every $p \in X$, there exists $q \in N$ s.t. $\|p - q\|_2 \leq r$.

Equivalently, an r-net is a maximal r-packing subset of X, or a minimal r-covering subset of X.

Problem definition



Problem definition

Definition (Approximate r-nets)

Given a pointset $X \subseteq \mathbb{R}^d$, a parameter $r > 0$ and an approximation parameter $\epsilon > 0$, a $(1+\epsilon)$-approximate r-net of X is a subset $N \subseteq X$ s.t. the following properties hold:

1. (packing) For every $p \neq q \in N$, we have that $\|p-q\|_2 \geq r$.
2. (covering) For every $p \in X$, there exists $q \in N$ $\|p-q\|_2 \leq (1+\epsilon)r$.

The result

Computing r-nets is a fundamental primitive in Computational Geometry.

Recent improvements in high dimensional "offline" problems:

- LSH: Approximate closest pair in time $\tilde{O}(dn^{2-\Theta(\epsilon)})$.
- [Valiant '12]: Approximate closest pair in time $\tilde{O}(dn^{2-\Theta(\sqrt{\epsilon})})$.

-Can we extend this improvement for the problem of computing r-nets?

The result

Theorem ([Avarikioti, Emiris, Kavouras, P '17])

There is a randomized algorithm which outputs w.h.p. an $(1+\epsilon)$-approximate r-net for n points in ℓ_2^d in time $\tilde{O}(dn^{2-\Theta(\sqrt{\epsilon})})$.

One application:

Theorem ([Avarikioti, Emiris, Kavouras, P '17])

There is a randomized algorithm which outputs w.h.p. an $(1+\epsilon)$-approximation to the kth nearest neighbor distance for n points in ℓ_2^d in time $\tilde{O}(dn^{2-\Theta(\sqrt{\epsilon})})$.

1 Random projections with false positives

2 Weak dimension reduction for doubling subsets of ℓ_1

3 High dimensional approximate r-nets

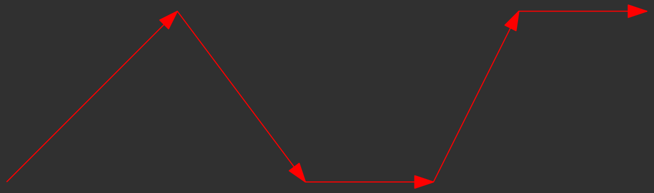
4 Approximate nearest neighbors for polygonal curves

5 Vapnik–Chervonenkis dimension for polygonal curves

Polygonal curves

What is a polygonal curve?

A sequence of vertices $v_1, \ldots, v_m$ in $\mathbb{R}^d$, with edges $\overline{v_1 v_2}, \overline{v_2 v_3}, \ldots \overline{v_{m-1} v_m}$.



Why curves?

Trajectories, data from mobiles, GPS sensors, video analysis etc.

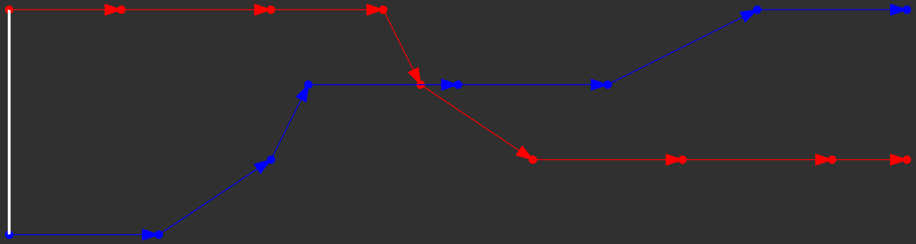
ANN for polygonal curves

Definition (Traversal of polygonal curves)

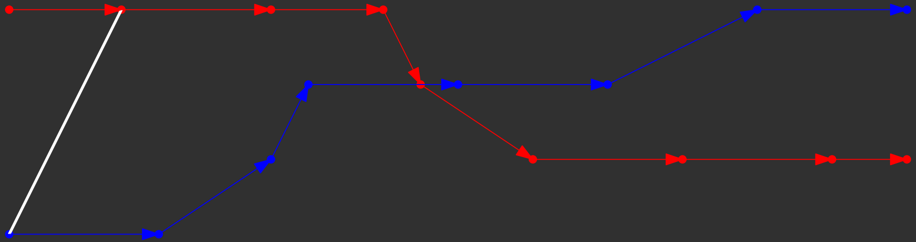
Given polygonal curves $V = v_1, \ldots, v_{m_1}$, $U = u_1, \ldots, u_{m_2}$, a traversal $T = (i_1, j_1), \ldots, (i_t, j_t)$ is a sequence of pairs of indices s.t.:

1. $i_1, j_1 = 1$, $i_t = m_1$, $j_t = m_2$.
2. $\forall (i_k, j_k) \in T : i_{k+1} - i_k \in \{0, 1\}$ and $j_{k+1} - j_k \in \{0, 1\}$.
3. $\forall (i_k, j_k) \in T : (i_{k+1} - i_k) + (j_{k+1} - j_k) \geq 1$.

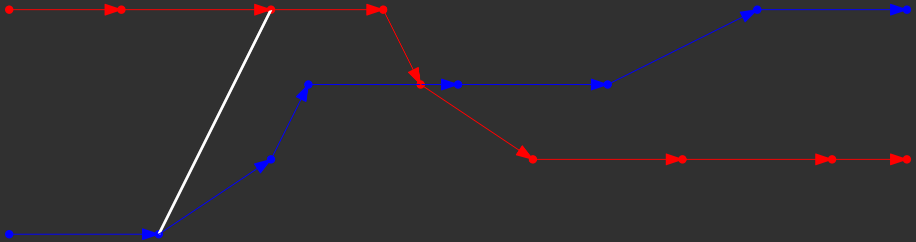
Discrete Metrics for curves



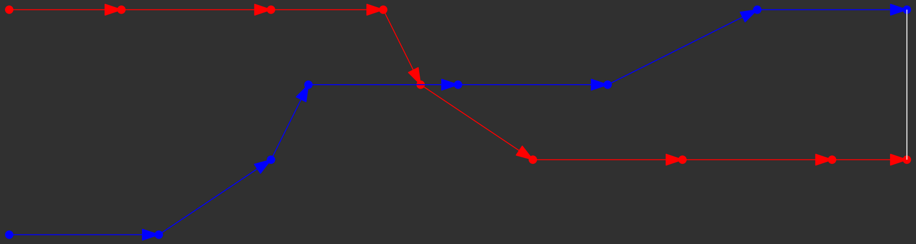
Discrete Metrics for curves



Discrete Metrics for curves



Discrete Metrics for curves



Discrete Metrics for curves

Definition (ℓ_p-distance of polygonal curves)

Given polygonal curves $V = v_1, \ldots, v_{m_1}$, $U = u_1, \ldots, u_{m_2}$, we define the ℓ_p-distance between V and U as the following function:

$$d_p(V,U) = \min_{T \in \mathcal{T}} \left(\sum_{(i_k, j_k) \in T} \|v_{i_k} - u_{j_k}\|_2^p \right)^{1/p},$$

where $\mathcal{T}$ denotes the set of all possible traversals for V and U.

Notice that,

$$d_\infty(V,U) \equiv \text{ Discrete Fréchet Distance,}$$

$$d_1(V,U) \equiv \text{Dynamic Time Warping.}$$

Point sequences

Definition (ℓ_p-products of ℓ_2)

We define the ℓ_p-product of ℓ_2 as the metric with domain $\left(\mathbb{R}^d\right)^k$ and distance function

$$d((x_1,\ldots,x_k),(y_1,\ldots,y_k)) = \left(\sum_{i=1}^{k} \|x_i - y_i\|_2^p\right)^{1/p}.$$

ANN for polygonal curves

The problem

ANN for the ℓ_p-distance of polygonal curves of complexity $\leq m$.

Sketch of solution [Emiris, P' 18]

- Build one data structure per traversal.
- Fix the traversal: ANN for ℓ_p-products of ℓ_2.
- Randomized embedding $\ell_2^d \mapsto \ell_p^{\tilde{O}(d)}$: ANN for $\ell_p^{\tilde{O}(md)}$.
- Invoke previous ANN results for $\ell_p^{\tilde{O}(md)}$.

Summary

	Space	Query	Approx.	Ref.
DFD	$O\left((m^2\|X\|)^m n^{2-o(1)}\right)$	$(m \log n)^{O(1)}$	$O(1)$	[4]
	$\tilde{O}(2^{4md} n)$	$\tilde{O}(2^{4md} \log n)$	$O(d^{3/2})$	[5]
	$\tilde{O}(n) \cdot 2^{\tilde{O}\left(m^{1/\epsilon} \cdot d \log \frac{1}{\epsilon}\right)}$	$\tilde{O}(dm^{1/\epsilon} 2^{4m} \log n)$	$1+\epsilon$	[EP18]
DTW	$\tilde{O}(mn)$	$O(m \log n)$	$O(m)$	[4]
	$\tilde{O}(n) \cdot 2^{O\left(m \cdot d \log \frac{1}{\epsilon}\right)}$	$\tilde{O}(d \cdot 2^{4m} \log n)$	$1+\epsilon$	[EP18]

Table: X denotes the domain set of the input metric. All previous results are tuned to optimize the approximation factor.

Recent: Space: $n \cdot O\left(\frac{1}{\epsilon}\right)^{dm}$, Query: $md \log\left(\frac{md}{\epsilon}\right)$ [Filtser, Filtser, Katz '19]

[4] any metric, deterministic, [Indyk '02]
[5] [Driemel, Silvestri '17]

Short queries

The problem

ANN for the DFD: dataset of complexity $\leq m$, queries of complexity $k \leq m$.

Sketch of solution

- Random partition of points: near points belong to the same parts with good probability, parts are of bounded diameter.
- Two "near" polygonal curves fall into the same sequence of $\leq k$ parts, with good probability.
- Discretize space and precompute answers for any sequence of $\leq k$ bounded-diameter parts.

Short queries

Theorem ([Driemel, P, Schmidt '19])

Given as input a set of n polygonal curves of complexity m in ℓ_2^d, and $\epsilon > 0$, there exists a randomized data structure with space in

$$n \cdot O\left(\frac{kd^{3/2}}{\epsilon}\right)^{kd} + O(dnm),$$

preprocessing time in $dnmk \cdot O\left(\frac{kd^{3/2}}{\epsilon}\right)^{kd}$, and query time in $O(dk)$, for the $(1+\epsilon, r)$-ANN problem under the discrete Fréchet distance. For any query curve q of complexity k, the preprocessing algorithm succeeds with constant probability.

Comparable to [Filtser, Filtser, Katz '19] when $m = k$.
Analogous result for arbitrary doubling metrics.

1 Random projections with false positives

2 Weak dimension reduction for doubling subsets of ℓ_1

3 High dimensional approximate r-nets

4 Approximate nearest neighbors for polygonal curves

5 Vapnik–Chervonenkis dimension for polygonal curves

Range spaces

Each range space can be defined as a pair of sets $(X, \mathcal{R})$, where X is the *ground set* and $\mathcal{R}$ is the *range set*. For $Y \subseteq X$, we denote:

$$\mathcal{R}_{|Y} = \{R \cap Y \mid R \in \mathcal{R}\}.$$

If $\mathcal{R}_{|Y}$ contains all subsets of Y, then Y is *shattered* by $\mathcal{R}$.

Definition (Vapnik-Chernovenkis dimension)

The Vapnik-Chernovenkis dimension (VC dimension) of $(X, \mathcal{R})$ is the maximum cardinality ν of a shattered subset of X.

Continuous Metrics for curves

Any polygonal curve V can be viewed as a parametrized curve $v: [0,1] \mapsto \mathbb{R}^d$.

Definition (Fréchet distance)

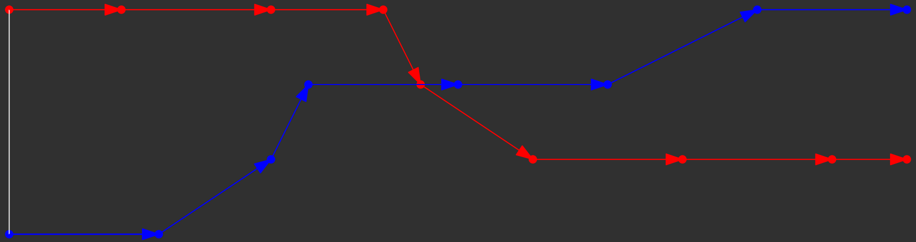
Given two parametrized curves $u, v: [0,1] \mapsto \mathbb{R}^d$, their Fréchet distance is defined as follows:

$$\mathrm{d}_F(u,v) = \min_{f:[0,1]\mapsto[0,1]} \max_{\alpha\in[0,1]} \|v(\alpha) - u(f(\alpha))\|,$$

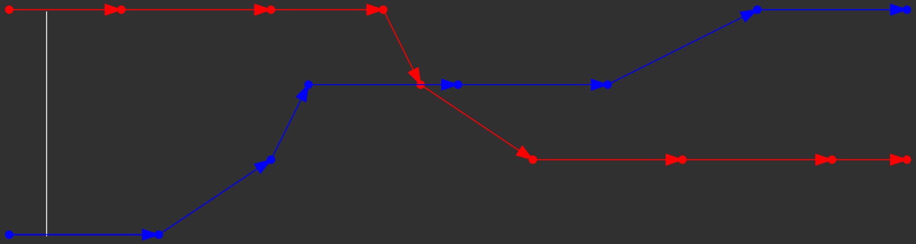
where f ranges over all continuous and monotone bijections with $f(0)=0$ and $f(1)=1$.

Weak Fréchet distance: no monotonicity constraint.

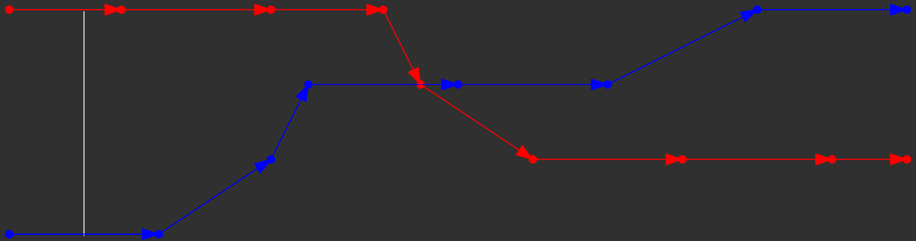
Continuous Metrics for curves



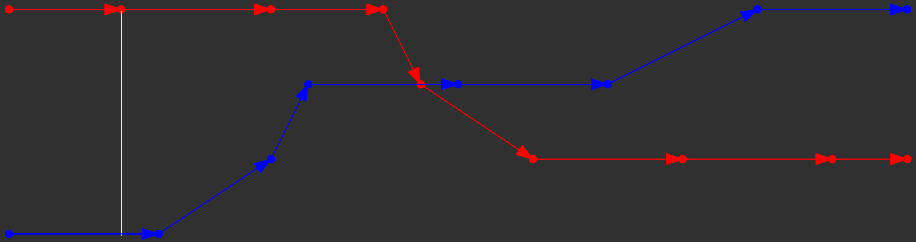
Continuous Metrics for curves



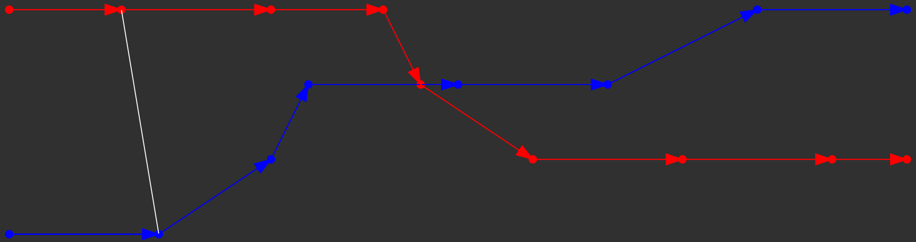
Continuous Metrics for curves



Continuous Metrics for curves



Continuous Metrics for curves



Range spaces for curves

Let (M, d) be a pseudometric space. We define the *ball* of radius r and center p, under the distance measure d, as the following set:

$$b_{\mathrm{d}}(p, r) = \{x \in M \mid \mathrm{d}(x, p) \leq r\},$$

where $p \in M$.
We study the VC dimension of range spaces $(X, \mathcal{R})$ induced by pseudometric spaces (M, d) by setting $X = M$ and

$$\mathcal{R} = \{b_{\mathrm{d}}(p, r) \mid r \in \mathbb{R}_+, p \in M\}.$$

Fréchet Distance Predicates

$P_{v,v}$ *(Vertex-Vertex)* **Input**: vertices v, u. **Returns** true iff $\|v - u\|_2 \leq r$.

$P_{v,e}$ *(Vertex-edge)* **Input**: edge e, vertex v. **Returns** true iff $\exists p :\in e : \ \|v - p\|_2 \leq r$.

P_m *(Monotonicity)* **Input**: ordered vertices v_i, v_j with $i < j$, and directed edge e. **Returns** true if $\exists p_1, p_2$ on the line supported by e, such that

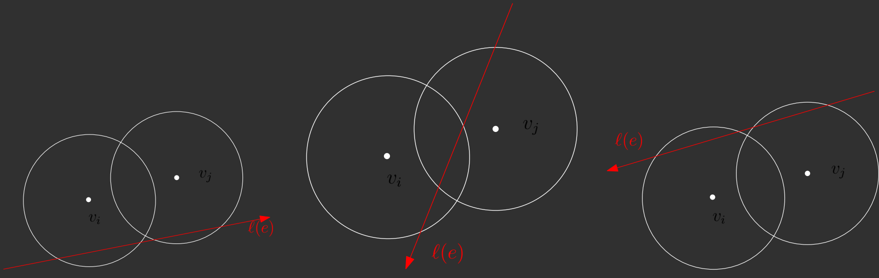
$$\|p_1 - v_i\|_2 \leq r, \ \|p_2 - v_j\|_2 \leq r,$$

and p_1 appears before p_2 on this line.

Lemma

Given the truth values of all predicates $P_{v,v}, P_{v,e}, P_m$ of two curves p, q for a fixed value of r, one can determine if $\mathrm{d}_F(s, q) \leq r$.

Fréchet Distance Predicates-Monotonicity



(a) True (b) True (c) False

Figure: The monotonicity predicate. Assume $i < j$.

Overall

The overall approach:

1. Break down to *primitive* range spaces: For any predicate $h : \mathbb{R}^n \times \mathbb{R}^d \mapsto \{0,1\}$, consider range space $(\mathbb{R}^n, \mathcal{H})$, where

$$\mathcal{H} = \{r_h(\alpha) \mid \alpha \in \mathbb{R}^d\},$$

   where $r_h(\alpha) = \{x \in \mathbb{R}^n \mid h(x,\alpha) = 1\}$.
2. Analyze each individual range space:
   - either by geometric arguments,
   - or by a powerful theorem of [Anthony, Bartlett] for *simple* predicates.
3. Extend the argument by composition.

Results

D. Fréchet	$O(dk \log(dkm))$	$\Omega(\max(dk \log k, \log dm))$, $(d \geq 4)$
weak Fréchet	$O(d^2 k \log(dkm))$	
Fréchet	$O(d^2 k^2 \log(dkm))$	$\Omega(\max(k, \log m))$, $(d \geq 2)$

Table: [Driemel, Phillips, P' 19] Ground set: polygonal curves of complexity m. Range set: balls centered at polygonal curves of complexity k.

Similar results for the Hausdorff distance.
Several implications by known sampling results.

References I

- Avarikioti, G., Emiris, I. Z., Kavouras, L. & Psarros, I. *High-dimensional approximate r-nets.* in *SODA* (2017).
- Anagnostopoulos, E., Emiris, I. Z. & Psarros, I. Randomized Embeddings with Slack and High-Dimensional Approximate Nearest Neighbor. *ACM Trans. Algorithms.* Prel. version in SoCG '15. (2018).
- Emiris, I. Z. & Psarros, I. *Products of Euclidean Metrics and Applications to Proximity Questions among Curves.* in *SoCG* (2018).
- Driemel, A., Phillips, J. M. & Psarros, I. *The VC Dimension of Metric Balls under Fréchet and Hausdorff Distances.* in *SoCG, To appear.* (2019).
- Driemel, A., Psarros, I. & Schmidt, M. Sublinear data structures for short Fréchet queries. **Submitted.** (2019).

References II

Emiris, I. Z., Margonis, V. & Psarros, I. Near neighbor preserving dimension reduction for doubling subsets of ℓ_1. *CoRR* **abs/1902.08815, Submitted** (2019).

Thank you!