Threshold-Based Graph Reconstruction Using Discrete Morse Theory

Brittany Terese Fasy[1,2] **Sushovan Majhi**[3]
Carola Wenk[4]

[1]Department of Computer Science, Montana State University

[2]Department of Mathematics, Montana State University

[3]**Department of Mathematics, Tulane University**

[4]Department of Computer Science, Tulane University

FWCG, 2018

PROBLEM STATEMENT

Given a (noisy) sample S taken around a (hidden) embedded graph G, how one can "reconstruct" the topology and geometry of G from S.

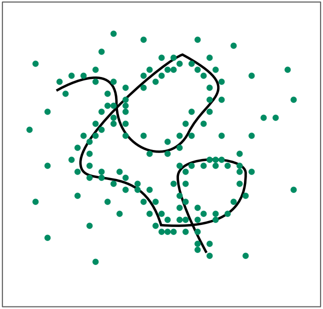

FIGURE: Sample around an embedded graph

Application: Map Reconstruction from GPS traces

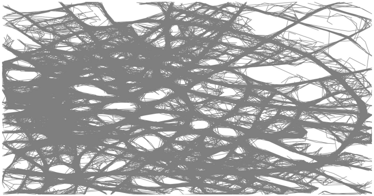

Figure: GPS traces of Berlin (mapconstruction.org)

Application: Map Reconstruction from GPS traces

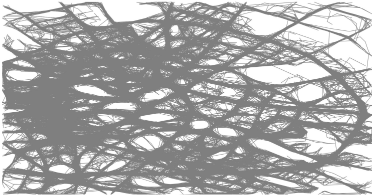

Figure: GPS traces of Berlin (mapconstruction.org)

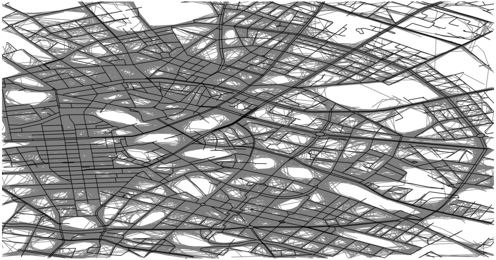

Figure: A reconstruction

Noise Models

Noise Models

1. Hausdorff noise
2. Non-Hausdorff noise

Noise Models

Noise Models

1. Hausdorff noise
2. Non-Hausdorff noise

What to reconstruct?

1. Topology (same homotopy type)
2. Geometry (small Hausdorff-distance)

Generic Density-Based Algorithm

Given a discretized domain $\tilde{D}$ and a sample S around G.

1. Compute density f of S over $\tilde{D}$.

GENERIC DENSITY-BASED ALGORITHM

Given a discretized domain $\tilde{D}$ and a sample S around G.

1. Compute density f of S over $\tilde{D}$.
    - Histogram Computation
    - Kernel Density Estimate

$$K(x, y; b) := \exp\left(\frac{-\|x - y\|^2}{2b^2}\right)$$

$$f(x) = \frac{1}{2\pi|S|b^2} \sum_{X_i \in S} K(x, X_i; b)$$

Generic Density-Based Algorithm

Given a discretized domain $\tilde{D}$ and a sample S around G.

1. Compute density f of S over $\tilde{D}$.
   - Histogram Computation
   - Kernel Density Estimate

$$K(x, y; b) := \exp\left(\frac{-\|x - y\|^2}{2b^2}\right)$$

$$f(x) = \frac{1}{2\pi|S|b^2} \sum_{X_i \in S} K(x, X_i; b)$$

2. for an appropriate threshold t, $f^{-1}[t, \infty)$ is considered.

Generic Density-Based Algorithm

Given a discretized domain $\tilde{D}$ and a sample S around G.

1. Compute density f of S over $\tilde{D}$.
   - Histogram Computation
   - Kernel Density Estimate

$$K(x, y; b) := \exp\left(\frac{-\|x-y\|^2}{2b^2}\right)$$

$$f(x) = \frac{1}{2\pi|S|b^2} \sum_{X_i \in S} K(x, X_i; b)$$

2. for an appropriate threshold t, $f^{-1}[t, \infty)$ is considered.
3. heuristic pruning methods are applied this super-level set to approximate G.

Related Work

Recent Works

1. Choosing thresholds systematically using Persistent Homology. (AFGW15)

Related Work

Recent Works

1. Choosing thresholds systematically using Persistent Homology. (AFGW15)
2. Graph reconstruction by Discrete Morse theory. (DWW18)

Related Work

Recent Works

1. Choosing thresholds systematically using Persistent Homology. (AFGW15)
2. Graph reconstruction by Discrete Morse theory. (DWW18)

Limitations

1. Thresholds are chosen heuristically.
2. Theoretical guarantees on the topological/geometric correctness is not proved.
3. The output is often a thick region around the hidden graph.

$(\omega, \beta_1, \beta_2, \nu)$-approximation of G

Let $\omega > 0$ such that G^ω has a deformation retract onto G.

$$f(x) \in \begin{cases} [\beta_1, \beta_1 + \nu], & x \in V^\omega \\ [\beta_2, \beta_2 + \nu], & x \in G - V^\omega \\ [0, \nu], & \text{otherwise} \end{cases}$$

V denotes the set of vertices of G.

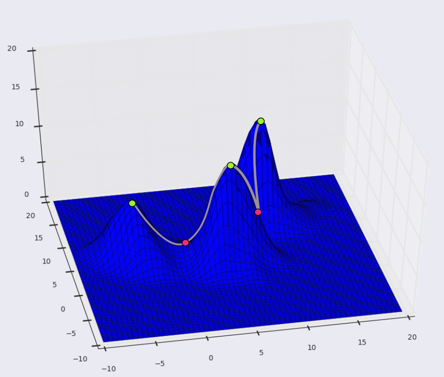

FIGURE: KDE

If the samples are concentrated around a graph, then the mountain ridges on the graph of the density function are expected to capture it.

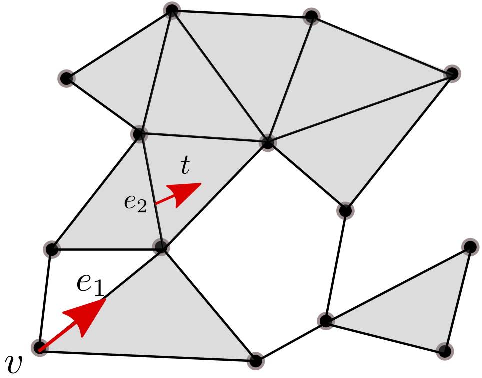

Figure: Simplicial Complex K and discrete vector

Discrete Vector

(σ, τ) is a discrete vector in K if $\sigma < \tau$ i.e. σ is a boundary of τ.
(v, e_1) and (e_2, t) are two discrete vectors on K.

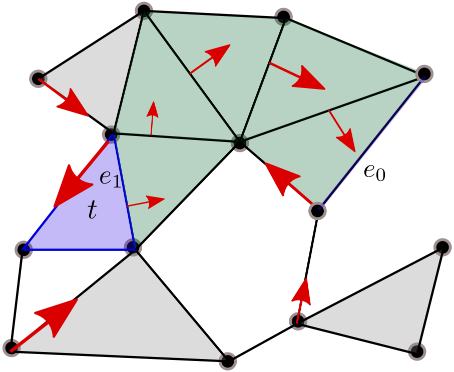

FIGURE: Simplicial Complex K and discrete vector

DISCRETE VECTOR FIELD

A discrete vector field V is a collection of discrete vectors such that every simplex of K is head/tail of at most one vector.

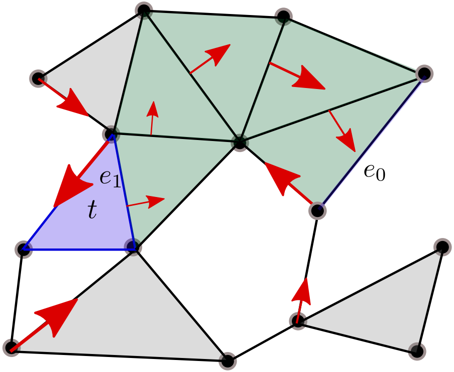

Figure: V-path

V-path

$$\sigma_0, \tau_0, \sigma_1, \tau_1, \ldots, \sigma_{l+1}$$

where (σ_i, τ_i) is a vector and $\tau_i < \sigma_{i+1}$.

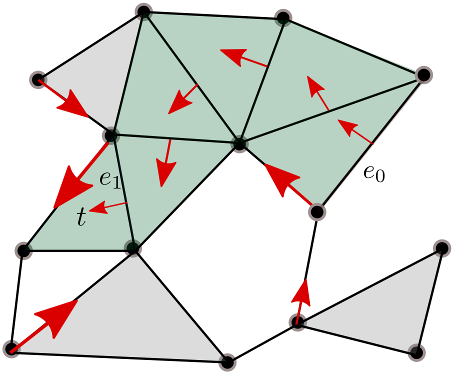

Figure: Morse Cancellation

Morse cancellation cancels critical points.

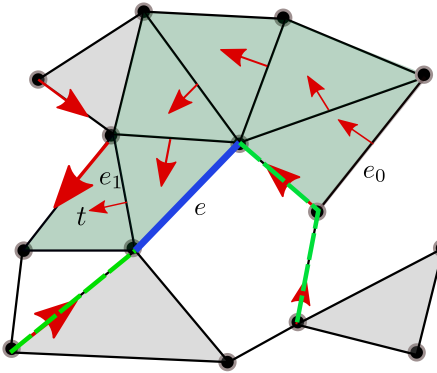

FIGURE: Morse Cancellation

STABLE MANIFOLD

For a critical edge *e*, its stable manifold is the set of all V-paths ending at the boundary of *e*.

Our Algorithm

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

1. Initialize V as the trivial vector field on K and $\hat{G} = \emptyset$.

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

1. Initialize V as the trivial vector field on K and $\hat{G} = \emptyset$.
2. Run persistence on the super-level set filtration of f to get the persistence pairs $P(K)$.

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

1. Initialize V as the trivial vector field on K and $\hat{G} = \emptyset$.
2. Run persistence on the super-level set filtration of f to get the persistence pairs $P(K)$.
3. For each $(\sigma, \tau) \in P(K)$ with $Pers(\sigma, \tau) < \delta$
   Try to perform a Morse cancellation for the pair and update V.

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

1. Initialize V as the trivial vector field on K and $\hat{G} = \emptyset$.
2. Run persistence on the super-level set filtration of f to get the persistence pairs $P(K)$.
3. For each $(\sigma, \tau) \in P(K)$ with $Pers(\sigma, \tau) < \delta$
   Try to perform a Morse cancellation for the pair and update V.
4. For each $(v, e) \in P(K)$ and $(e, t) \in P(K)$ with $Pers(v) \geq \delta$, $\hat{G} = \hat{G} \cup \{$ stable manifold of $e\}$.

Input: The discretized domain K, the density function f, the threshold δ
Output: The reconstructed graph $\hat{G}$

1. Initialize V as the trivial vector field on K and $\hat{G} = \emptyset$.
2. Run persistence on the super-level set filtration of f to get the persistence pairs $P(K)$.
3. For each $(\sigma, \tau) \in P(K)$ with $Pers(\sigma, \tau) < \delta$
   Try to perform a Morse cancellation for the pair and update V.
4. For each $(v, e) \in P(K)$ and $(e, t) \in P(K)$ with $Pers(v) \geq \delta$, $\hat{G} = \hat{G} \cup \{$ stable manifold of $e\}$.
5. output $\hat{G}$

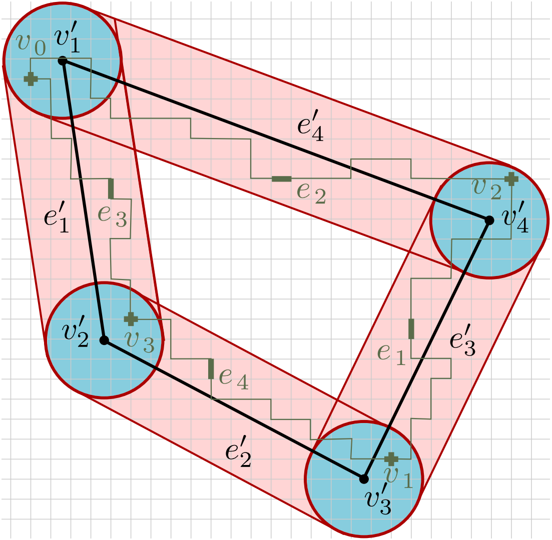

FIGURE: Algorithm in Picture

THEOREM

If G is a connected, embedded planar graph in a cubical complex K and f is an $(\omega, \beta_1, \beta_2, \nu)$-approximation then the output $\widehat{G}$ has the same homotopy type as G. Moreover, $d_H(G, \widehat{G}) < \omega$.

Future Work

1. Extend the result to higher dimensions.
2. What condition on the density function gives up a small Fréchet distance between the edges of the output and the edges of the graph.
3. How to circumvent the heavy persistence computation for all nodes.

Thanks

and

Questions?

References I

Mahmuda Ahmed, Brittany Terese Fasy, Matt Gibson, and Carola Wenk, *Choosing thresholds for density-based map construction algorithms*, Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems (New York, NY, USA), SIGSPATIAL '15, ACM, 2015, pp. 24:1–24:10.

Tamal K. Dey, Jiayuan Wang, and Yusu Wang, *Graph reconstruction by discrete Morse theory*, 34th International Symposium on Computational Geometry, 2018, pp. 31:1–31:15.