

Efficient Trajectory Retrieval

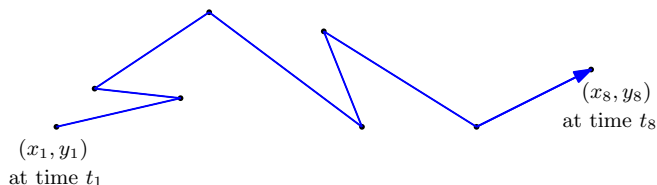
Hasan Pourmahmood

October 30, 2021

Trajectory

A spatial trajectory (also called a piece-wise linear curve) is a sequence of waypoints along with a time stamp, i.e.

$$T = \{(x_1, y_1, t_1), \dots, (x_n, y_n, t_n)\}.$$



Popular Distances

- ▶ Dynamic Time Warping distance
- ▶ Fréchet distance
- ▶ Discrete Fréchet distance

Popular Distances

- ▶ Dynamic Time Warping distance
- ▶ Fréchet distance
- ▶ Discrete Fréchet distance

Complexity: Quadratic in number of waypoints $O(mn)$.

Popular Distances

- ▶ Dynamic Time Warping distance
- ▶ Fréchet distance
- ▶ Discrete Fréchet distance

Complexity: Quadratic in number of waypoints $O(mn)$.

Queryies

- ▶ **k -NN queries:** find k nearest trajectories from data to the given query trajectory Q .
- ▶ **Range queries:** find all trajectories within distance r from data to the given query trajectory Q .

Popular Distances

- ▶ Dynamic Time Warping distance
- ▶ Fréchet distance
- ▶ Discrete Fréchet distance

Complexity: Quadratic in number of waypoints $O(mn)$.

Queries

- ▶ **k -NN queries:** find k nearest trajectories from data to the given query trajectory Q .
- ▶ **Range queries:** find all trajectories within distance r from data to the given query trajectory Q .

Applications

Route extraction, route recommendation, ML on trajectory datasets, ...

Information retrieval approach

Information retrieval approach

Treating the problem as a “term-document” pair

Information retrieval approach

Treating the problem as a “term-document” pair

- ▶ Consider a grid structure with some resolution for the region R , where data lives in, in order to get term-document representation. Each grid will be treated as a term.

Information retrieval approach

Treating the problem as a “term-document” pair

- ▶ Consider a grid structure with some resolution for the region R , where data lives in, in order to get term-document representation. Each grid will be treated as a term.
- ▶ Map each trajectory (as a document) to a vector space of large but fixed dimension.

Information retrieval approach

Treating the problem as a “term-document” pair

- ▶ Consider a grid structure with some resolution for the region R , where data lives in, in order to get term-document representation. Each grid will be treated as a term.
- ▶ Map each trajectory (as a document) to a vector space of large but fixed dimension.
- ▶ Employ a dimensionality reduction technique to reduce the dimension (our [index](#)).

Information retrieval approach

Treating the problem as a “term-document” pair

- ▶ Consider a grid structure with some resolution for the region R , where data lives in, in order to get term-document representation. Each grid will be treated as a term.
- ▶ Map each trajectory (as a document) to a vector space of large but fixed dimension.
- ▶ Employ a dimensionality reduction technique to reduce the dimension (our [index](#)).
- ▶ Utilizing this vector space model in a lower dimensional space, consider a similarity measure between reduced vectors.

Information retrieval approach

Treating the problem as a “term-document” pair

- ▶ Consider a grid structure with some resolution for the region R , where data lives in, in order to get term-document representation. Each grid will be treated as a term.
- ▶ Map each trajectory (as a document) to a vector space of large but fixed dimension.
- ▶ Employ a dimensionality reduction technique to reduce the dimension (our [index](#)).
- ▶ Utilizing this vector space model in a lower dimensional space, consider a similarity measure between reduced vectors.
- ▶ Answer to the queries in an efficient way.

Experiments, Evaluation and Interface

Experiments, Evaluation and Interface

- ▶ Consider 2 or 3 real world trajectory data sets.

Experiments, Evaluation and Interface

- ▶ Consider 2 or 3 real world trajectory data sets.
- ▶ Use the vector space model on data sets and evaluate the performance of the search method using all data points as queries.

Experiments, Evaluation and Interface

- ▶ Consider 2 or 3 real world trajectory data sets.
- ▶ Use the vector space model on data sets and evaluate the performance of the search method using all data points as queries.
- ▶ Report some statistics like [accuracy](#), [precision](#), [recall](#), [nDCG](#).

Experiments, Evaluation and Interface

- ▶ Consider 2 or 3 real world trajectory data sets.
- ▶ Use the vector space model on data sets and evaluate the performance of the search method using all data points as queries.
- ▶ Report some statistics like [accuracy, precision, recall, nDCG](#).
- ▶ Provide an interactive [user-friendly interface](#) for arbitrary queries on the map, where a user can create a query trajectory by clicking on the map and specifying the value of k (or r) to get the the recommended k nearest trajectories from data on the map visually.

My contributions

My contributions

1. Using different datasets as benchmark,

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,
3. Using term-frequency type of trajectory vectorization,

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,
3. Using term-frequency type of trajectory vectorization,
4. Using a landmark-based vectorization (if time allows),

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,
3. Using term-frequency type of trajectory vectorization,
4. Using a landmark-based vectorization (if time allows),
5. Employing a kernel-based similarity measure for terms,

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,
3. Using term-frequency type of trajectory vectorization,
4. Using a landmark-based vectorization (if time allows),
5. Employing a kernel-based similarity measure for terms,
6. In addition to the metric used in the reference paper, I will try to apply DTW or discrete Fréchet distance as a ground truth,

My contributions

1. Using different datasets as benchmark,
2. Utilizing binary vectorization of trajectories,
3. Using term-frequency type of trajectory vectorization,
4. Using a landmark-based vectorization (if time allows),
5. Employing a kernel-based similarity measure for terms,
6. In addition to the metric used in the reference paper, I will try to apply DTW or discrete Fréchet distance as a ground truth,
7. Providing a user-friendly interactive visualization tool for retrieving target trajectories.

QUESTIONS/COMMENTS?