

## Supplemental Material

### Appendix A. Comparison of Uniform Representation

The running time of the uniform representation is related to the size of the sliding window and the number of trajectories. Figure 1 shows the running time spent under different parameter settings, where the x-axis represents the number of trajectories and the y-axis represents the running time in milliseconds. We can see that the k-gram method (i.e., the blue, yellow, green and red lines) runs significantly faster than dynamic time warping (i.e., the purple line). For example, when analyzing 1,000 trajectories, k-gram takes 690 milliseconds, while dynamic time warping takes 3,500 milliseconds, resulting in a five times difference. In addition, the size of the sliding window has little effect on the running time. As the number of trajectories increases, the running time constantly increases, and it takes about 2,500 milliseconds (i.e., 42 minutes) to analyze 2,000 trajectories at a time.

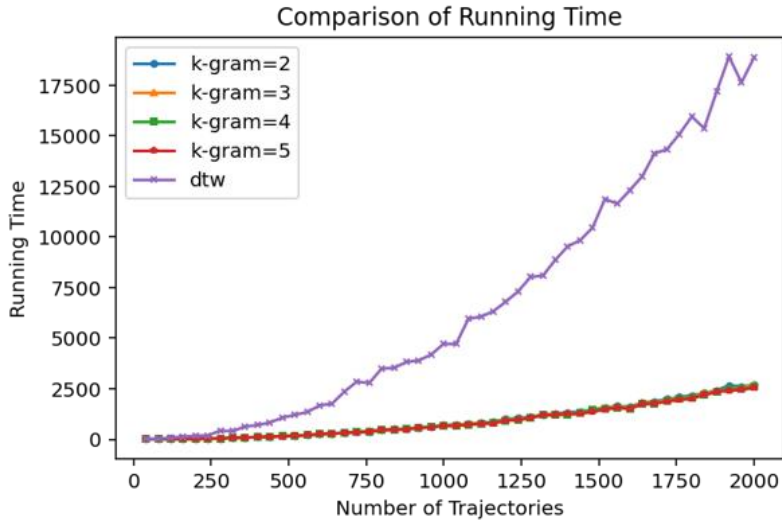


Figure 1: Running time of k-gram and dynamic time warping.

### Appendix B. Evaluation of Clustering Effects with Three Metrics

In preliminary experiments, we compare the effects of k-Means, hierarchical clustering (abbreviated as AC), balanced iterative reducing and clustering using hierarchies (abbreviated as BIRCH) and spectral clustering (abbreviated as Spectral), respectively. Figure 2 illustrates their performance on silhouette coefficient, Calinski Harabasz index (abbreviated as CHI) and Davies Bouldin index (abbreviated as DBI). The larger values of silhouette coefficient and CHI, the better the clustering effect, while DBI is the closer to 0, the better. Overall, the hierarchical clustering (i.e., the orange line) outperforms the other methods in both silhouette coefficient and DBI indexes (Figure 2(a,b,c)) and the performance on CHI is not optimal, but the difference is not significant.

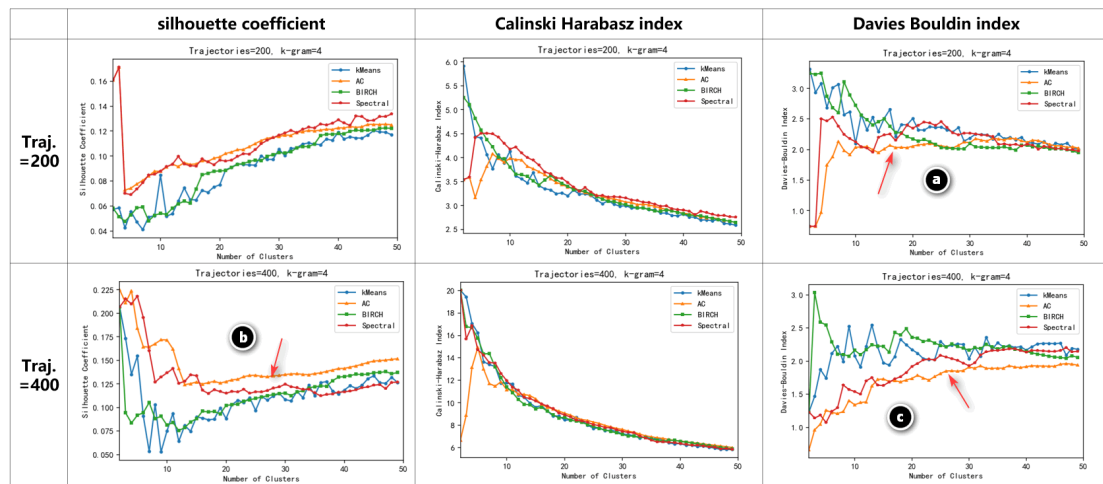


Figure 2: Performance of different clustering methods on each metric.

### Appendix C. Questionnaire for Preference of Travel Route Choice

Dear Sir/Madam,

This questionnaire focuses on investigating preferences for route choice, which will be used to improve the urban travel experience. The results are for research use only, and the personal information involved will be strictly protected. Thank you for your active participation!

**Q1:** What's your gender?

- A. Male
- B. Female

**Q2:** What's your age?

- A. Under 18
- B. 18~25
- C. 25~30
- D. 31~35
- E. 36~40
- F. 41~50
- G. Over 50

**Q3:** What factors do you usually consider when you choose a specific route (Note that 1~10 indicates the priority from highest to lowest)?

- A. Distance
- B. Congestion
- C. Duration
- D. Destination
- E. Road familiarity
- F. Cost

- G. Signal
- H. Weather
- I. Departure time
- J. Road level

**Q4:** Suppose you are going to work (or school), which factors will you give priority to?

- A. Distance
- B. Signal
- C. Congestion
- D. Road level
- E. Duration
- F. Weather
- G. Road familiarity
- H. Destination
- I. Departure time
- J. Cost

**Q5:** Suppose you are going to the train station (or airport), which factors will you give priority to?

- A. Distance
- B. Signal
- C. Congestion
- D. Road level
- E. Duration
- F. Weather
- G. Road familiarity
- H. Destination
- I. Departure time
- J. Cost

**Q6:** Suppose you are going to the shopping mall, which factors will you give priority to?

- A. Distance
- B. Signal
- C. Congestion
- D. Road level
- E. Duration
- F. Weather
- G. Road familiarity

H. Destination

I. Departure time

J. Cost

**Q7:** Suppose you are going to the hospital, which factors will you give priority to?

A. Distance

B. Signal

C. Congestion

D. Road level

E. Duration

F. Weather

G. Road familiarity

H. Destination

I. Departure time

J. Cost

**Q8:** Except for the above factors, what are the additional considerations you take into account?