

HOW TO RUN THE EXPERIMENTS BY YOURSELF

(GeoHighlight Algorithm)

1st step (generate the ds.csv file)

- From the prepared datasets, you may run the *dsgenerator.py* to generate the *ds.csv* file. This file (*ds.csv*) is the input for the *iuga.py* script (the main algorithm for the GeoHighlight proposal).
- The *ds.csv* file can be generated from the original data set, for instance, a subset of the original dataset with 1,000 lines or 10,000 lines.
- The input .csv files are in the inputs folder, for example *input/nyctaxi10000.csv*
- To change the input you must change the line 83 in the *dsgenerator.py*.

2nd step (run iuga.py isolated)

- Run the *iuga.py*, changing the parameters from line 8 to line 14.

3rd step (running the experiments)

- The script shell *experiments.sh* generates the experiments results by running the *iuga_experiments.py* (an adapted version for the shell script) for K values from 2 to 500 in the *sigma* range from 0.1 to 10.
- The files are generated *outputs/execution/result<<\$sigma>>.csv*
- This files have omitted headers. The features of the produced csv file is:

k, lowest_acceptable_similarity, current_records, nb_iterations, total_time

Features description

k - K similar points

lowest_acceptable_similarity - The sigma value with the lowest acceptable similarity, a parameter of iuga algorithm.

current_records - the set P of point p points.

nb_iterations - number of iterations

total_time - execution time.

OBS: If you need change the input point, it is necessary to change the *iuga_experiments.py* files
(line 12)

OBS2: *It takes a lot of time to generate the experiments.*