The datasets are all in folder "datasets". The functions are all in folder "functions".

The steps of testing the similarity measures are the followings:

1. open the file "testcms.m"

2. uncomment one dataset

3. If you want to test k-mode, then uncomment the "test kmodes code";

If you want to test spectral or k-distance, uncomment one line of similarity matrix and uncomment the " test spectral clustering" or "test k-distance". The alpha value is 0.5, you can change in the code:"a=0.5".

The time of calculating similarity measure is very long. So if you want to test k-distance for several times, you can open the file "testdis.m". The similarity matrixes of cms have been stored in dataset folder. So you can uncomment the "read data code" and execute the file.

The file "draw.a" is the graph with different a. The kdis\_ire is the graph with different iretation times.

And very importantly, all algorithms have random factor, So you must execute the same code for several times and get the best result.(You can not execute the algorithm for many times by using loop due to the computer can only produce paseudorandom.) Sometimes, the kmode algorithm will go wrong for the iniatiating reasons. So if the execution time is longer then 10 minutes, you should consider stop the code and restart it.