We provide one .csv files, the raw dataset “sanitation.csv”.

The raw data were collected by the wrist smartwatch which was equipped with a triaxial accelerometer. An SD card and a SIM card were installed for storage and real-time data transmission, respectively.

The self-collected Sanitation dataset is collected from the open environment. When the sanitation workers were doing the daily work activities with the smartphone worn on the right hand or the left hand, the data were collected continuously at a frequency of 25 Hz and sent to the receiver server through the SIM card. These seven types of activity are: Walk, Run, Sweep, Bsweep (sweep using a big broom), Clean, Dump and Daily activities (like sitting and smoking).

The size of the whole dataset is 266555 x 3, which contains 266555 samples. Each sample contains X, Y and Z three axis acceleration values. There are 81739 samples of Bweep, 36502 samples of Clean, 45439 samples of Daily, 29518 samples of Dump, 3903 samples of Run, 60028 samples of Sweep and 9426 samples of Walk.

The first three columns of the sanitation.csv file represent the acceleration data of the X-axis, Y-axis and z-axis respectively. The acceleration data unit is g, that is, 9.81m/s. The fourth column represents the sampling point label.

Yong Zhang

March. 28, 2019