PyNKDV: An Efficient Network Kernel Density Visualization Library for Geospatial Analytic Systems

Tsz Nam Chan
Hong Kong Baptist University
edisonchan@comp.hkbu.edu.hk

Rui Zang
Hong Kong Baptist University
19251017@life.hkbu.edu.hk

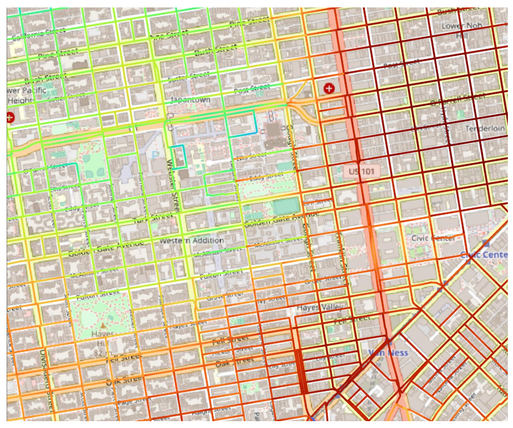
Pak Lon Ip
University of Macau
paklonip@um.edu.mo

Leong Hou U
University of Macau
ryanlhu@um.edu.mo

Jianliang Xu
Hong Kong Baptist University
xujl@comp.hkbu.edu.hk

Overview of Network Kernel Density Visualization (NKDV)

(a) 311-call data points



(b) NKDV

- Main applications
  - Crime hotspot detection
  - Traffic hotspot detection
  - Traffic accident hotspot detection

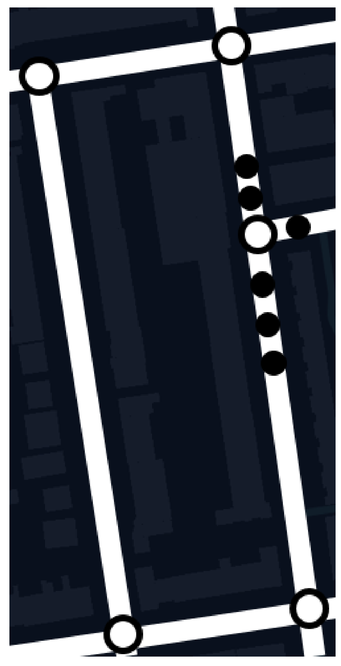
- Supported by famous software packages
  - spNetwork (an R package)
  - SANET (a plugin for QGIS and ArcGIS)

- How to generate NKDV?
  1. Divide each road into a set of lixels.
  2. Color each lixel q based on the network kernel density function $\mathcal{F}_P(q)$.

$$\mathcal{F}_P(q) = \overbrace{w}^{\text{constant}} \cdot \begin{cases} 1 - \frac{1}{b^2}\overbrace{d_G(q,p)}^{\text{shortest path distance}}{}^2 & \text{if } d_G(q,p) \le b \\ \underbrace{0}_{\text{bandwidth}} & \text{otherwise} \end{cases}$$

- The worst-case time complexity of generating NKDV is $O(L(T_{SP} + n))$
  - L is the number of lixels.
  - T_{SP} is the time complexity of the shortest path algorithm.
  - n is the number of data points.

- Generating NKDV is slow ☹

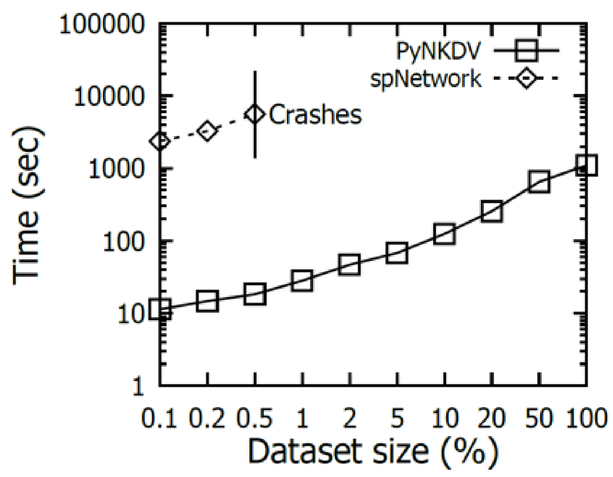


(a) Data points

(b) NKDV

Why PyNKDV?

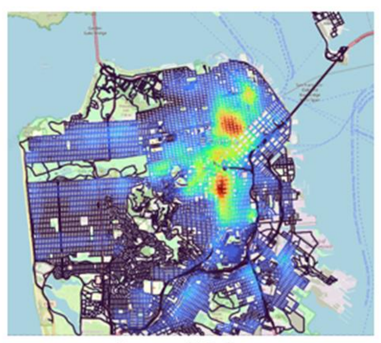
- Based on a fast solution, called aggregate distance augmentation (ADA), for generating NKDV ☺



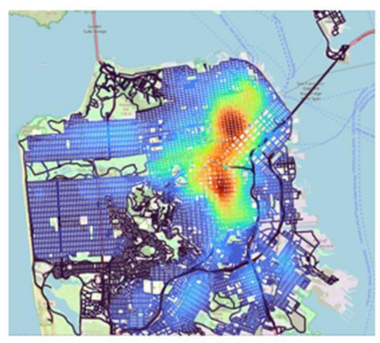
- The time complexity of ADA is $O\left(|E|T_{SP} + L|E|\log\left(\frac{n}{|E|}\right) + n\right)$.

- Efficiently support the bandwidth tuning operation ☺

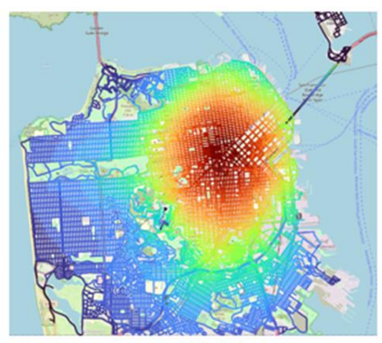
$b = 100$m



$b = 500$m



$b = 1000$m

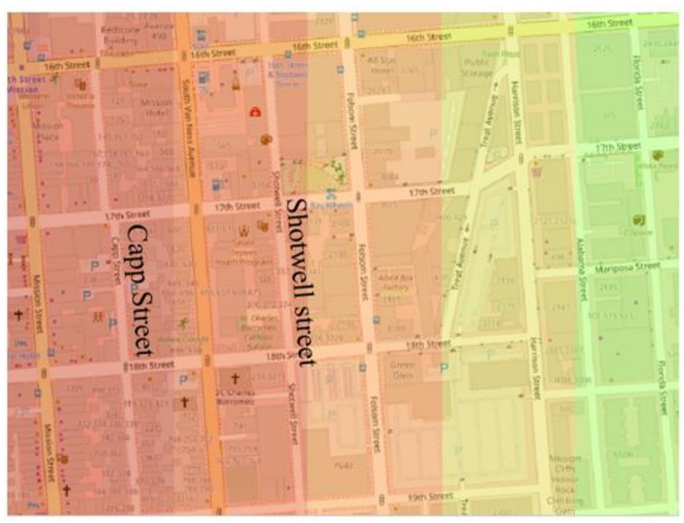


$b = 5000$m

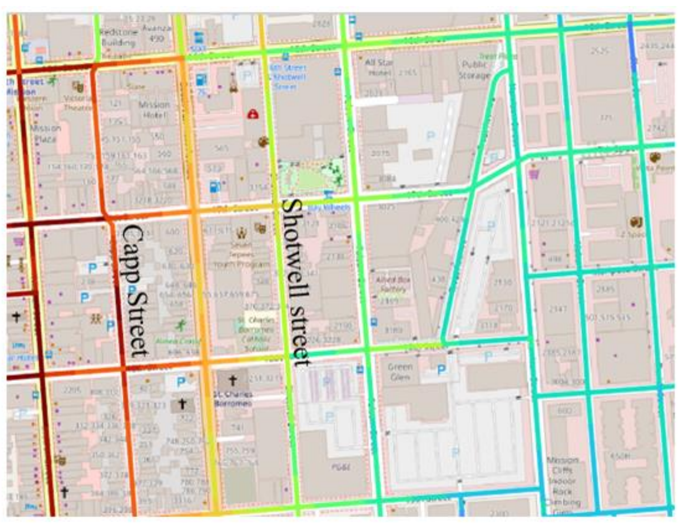
- Easy to use (Four lines of code) ☺

```
road_data = map_road_network(location_data)
model = PyNKDV(road_data, bandwidth=1000,
                    lixel_size=5, num_threads=8)
results = model.compute()
output(results, output_file_name)
```

Case Study: KDV v.s. NKDV



(a) KDV



(b) NKDV

- KDV tends to overestimate the density values in a road network.
- NKDV can provide more reasonable visualization results.

Github link for PyNKDV