Bayesian analysis of flow data on a network

An application to bike sharing data in Milan

Bissoli, Principi, Rinaldi

January 2019

Flow model

Let E be the set of arcs.

For any $(i, j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta \log(\mathbf{X}_{ij})$$

where $\beta = (\beta_1, ..., \beta_p)$ and $\mathbf{X}_{ij} = (X_{ij1}, ..., X_{ijp})$

Preliminary issues

Positioning problem:

real positions of the stations

Complexity problem:

too much stations

Localization of the stations

Positioning solution:

Cropping vs Clustering

Complexity solutions:

cropping arcs with <10 trips

dbscan clustering of stations

Distance model on cropped network

Let E be the set of arcs with more than 10 trips.

For any $(i, j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10)$$

where d_{ij} is the distance between the nodes computed using latitude and longitude

Distance and degrees model on cropped network

Let E be the set of arcs with more than 10 trips.

For any $(i, j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10) + \beta_2 \log(S_i) + \beta_3 \log(T_j)$$

where S_i is out-degree of node i and T_j is in-degree of node j

Distance and degrees model on cropped network

Distance model on clustered network

Let E be the set of arcs among the clusters.

For any $(i,j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10)$$

where d_{ij} is the distance between the baricenters of the clusters computed using latitude and longitude

Distance model on clustered network

Distance and degrees model on clustered network

Let E be the set of arcs among the clusters.

For any $(i, j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10) + \beta_2 \log(S_i) + \beta_3 \log(T_j)$$

where S_i is out-degree of node i and T_j is in-degree of node j

Distance and degrees model on clustered network

Distance, degrees and accessibility model on clustered network

Let E be the set of arcs among the clusters.

For any $(i, j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10) + \beta_2 \log(S_i) + \beta_3 \log(T_j) + \beta_4 \log(A_{ij}???)$$

where A is the accessibility matrix of elements $A_{ij} = formula!!!$

Distance, degrees and accessibility model on clustered network

Distance, degrees and center-periphery model on clustered network

Let E be the set of arcs among the clusters.

For any $(i,j) \in E$

$$Y_{ij} \sim Poisson(\mu_{ij})$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(d_{ij} + 10) + \beta_2 \log(S_i) + \beta_3 \log(T_j)$$
$$+\beta_4 \log(CC_{ij}) + \beta_5 \log(CP_{ij})$$
$$+\beta_6 \log(PC_{ij}) + \beta_7 \log(Auto_{ij})$$

where CC, CP, and PC are boolean matrices to store if trips are from center (C) or periphery (P) to C or P, and $Auto$ is an identity matrix to store if the trip is an autoarc.
mettiamo solo il risultato migliore : d s t cc auto oppure d s t cc cp pc auto?

Distance, degrees and center-periphery model on clustered network

Distance, degrees and center-periphery model on clustered network

Further developments

- mixed effects of center and periphery
- inflated Poisson

Thank you