Statistical methods for spatially-embedded relational data

Supervisor: Dr. Isabella Gollini

January 2019

This project will aim to develop new models for the analysis of complex spatial and relational data. In particular it will focus on assessing the effect of spatial dependence on network formation processes by adopting flexible modelling approaches including latent variable models and Gibbs random fields.

The relational structure of spatial data can be represented by links between observations. In this context, spatial locations are treated as an important source of information to measure the probability of the presence of a link between observations.

The models proposed will improve the capacity to identify clusters and relational patterns in the data which could not be captured by standard methods and to enhance prediction procedures.

Spatial and relational data are typically of large size and complexity and the likelihood functions of these proposed models will involve integrals that cannot be evaluated analytically. To overcome this problem it will be necessary to develop efficient estimation algorithms (such as using variational inference or likelihood-free methods).

Visualization is an important tool where complex data information can be transformed into an easily understood format, enhancing not only data comprehension, but also allowing for greater coverage of patterns within data. If information is more accessible and visible, it will allow for better insights into the various aspects of the research application. For this reason the design of a widely usable software (for example a Python or R package) is crucial to enhance the usability, comprehension and functionality of the models proposed.