

# **FLEXIBLE MODELLING OF HETEROGENEOUS POPULATIONS OF NETWORKS: A BAYESIAN NONPARAMETRIC APPROACH**

Francesco Barile <sup>1</sup>, Simón Lunagómez <sup>2</sup> and Bernardo Nipoti<sup>1</sup>

<sup>1</sup> Department of Economics, Management and Statistics, University of Milano-Bicocca, Italy

(e-mail: francesco.barile@unimib.it, bernardo.nipoti@unimib.it)

<sup>2</sup> Department of Statistics, Instituto Tecnológico Autónomo de México, México

(e-mail: simon.lunagomez@itam.mx)

**ABSTRACT:** A popular approach to the problem of clustering multiple network data makes use of distance metrics that measure the similarity among networks based on some of their global or local characteristics. In this context, we propose a novel Bayesian nonparametric approach to model undirected labelled graphs sharing the same set of vertices, which allows us to identify clusters of networks characterized by similar patterns in the connectivity of nodes. Our construction relies on the definition of a location-scale Dirichlet process mixture of centered Erdős–Rényi kernels. An efficient Markov chain Monte Carlo scheme is proposed to carry out posterior inference and provide a convenient clustering of the multiple network data, while the number of clusters in the population is not set a priori but inferred from the data. The performance of our approach is investigated by means of an extensive simulation study and illustrated with the analysis of a dataset on brain networks.

**KEYWORDS:** Bayesian nonparametrics, centered Erdős–Rényi model, Dirichlet process, mixture model, multiple network data.