

A COMPOSITIONAL STOCHASTIC BLOCK MODEL FOR THE ANALYSIS OF THE ERASMUS PROGRAMME NETWORK

Iuliia Promskaia^{1,2}, Adrian O'Hagan^{1,2} and Michael Fop¹

¹ School of Mathematics and Statistics, University College Dublin,
(e-mail: iuliia.promskaia@ucdconnect.ie, adrian.ohagan@ucd.ie,
michael.fop@ucd.ie)

² SFI Insight Centre for Data Analytics

ABSTRACT: The Erasmus Programme is one of the most well-known student exchange programmes in the world, with over 200,000 higher education students availing of its benefits each year. We explore the Erasmus exchange data, aiming to identify clustering structure in the mobility patterns of students between countries. A directed weighted network of country-to-country student exchanges is constructed, with edge weights representing the percentage of students travelling from one country to any other. These edge weights are compositional in nature, so they cannot be assumed independent. We propose an extension of the stochastic block model for clustering network data with compositional edge weights, and compare its performance to that of other models.

KEYWORDS: clustering, networks, compositional data, stochastic block model.

References

- GADÁR, L., KOSZTYÁN, Z.T., TELCS, A. ET AL. 2020. A multilayer and special description of the Erasmus mobility network. *Scientific Data*, **7**, 41.
- LEE, C., WILKINSON, D.J. 2019. A review of stochastic block models and extensions for graph clustering. *Applied Network Science*, **4**, 122.
- GREENACRE, M. 2021. Compositional Data Analysis. *Annual Review of Statistics and Its Applications*, **8:1**, 271-299.