COMMUNITY DETECTION IN TRIPARTITE NETWORKS OF UNIVERSITY STUDENT MOBILITY FLOWS

Vitale Maria Prosperina[1] , Vincenzo Giuseppe Genova[2], Giuseppe Giordano[1] and Giancarlo Ragozini[3]

[1] Department of Political and Social Studies, University of Salerno, (e-mail: mvitale@unisa.it, ggiordano@unisa.it)

[2] Department of Economics, Business, and Statistics, University of Palermo, (e-mail: vincenzogiuseppe.genova@unipa.it)

[3] Department of Political Science, Federico II University of Naples, (e-mail: giragoz@unina.it)

ABSTRACT: The purpose of this study is to explore how the multimode network approach can be used to analyse network patterns derived from student mobility flows. We define a tripartite network based on a three-mode data structure, consisting of Italian provinces of residence, universities and fields of study, with student exchanges representing the links between them. A comparison of algorithms for detecting communities from tripartite networks based on modularity optimization is provided, revealing relevant information about the phenomenon under analysis over time. The findings are applied to a real dataset containing micro-level longitudinal information on Italian university students' careers.

KEYWORDS: student mobility, tripartite networks, modularity optimisation

1 Introduction

The analysis of intra- and international student mobility has become a vibrant research field in migration literature and a key concern for national policy-making on tertiary education systems (Van Mol & Timmerman, 2014; Riaño *et al.*, 2018). Usually, European mobility in higher education is described by considering the dynamics of the Erasmus programme. From a national perspective, Italian student mobility from high school to bachelor and master degrees is analysed as a crucial step in determining future migration choices. Such analysis shows an unbalanced migration of students from the southern to the northern regions of the country (Genova *et al.*, 2019), which is influenced by the attractiveness of universities, related to the socio-economic characteristics and the job market opportunities in the geographic areas where they are located (Giambona *et al.*, 2017; Impicciatore & Panichella, 2019). Given the na-

ture of the student mobility data (i.e. flows of students connecting provinces of residence and universities of destination), network analysis has been adopted as one of the most appropriate methodological approach to interpret this phenomenon (Santelli *et al.*, 2019; Genova *et al.*, 2019; Columbu *et al.*, 2021).
Based on this theoretical framework and the intrinsic complexity of student mobility flows, this study analyses the data at hand using the framework of multimode networks (Fararo & Doreian, 1984). More specifically, we define a tripartite network based on a three-mode data structure, consisting of Italian provinces of residence, universities and fields of study, with student exchanges representing the links between them. A comparison of algorithms for detecting communities from tripartite networks or k-partite modularity (Neubauer & Obermayer, 2009; Ikematsu & Murata, 2013; Melamed *et al.*, 2013; Ignatov *et al.*, 2017; Feng *et al.*, 2019), mainly based on modularity optimisation, is applied to reveal relevant information about the phenomenon under analysis. The algorithms are applied to the MOBYSU.IT dataset which contains micro-level longitudinal information on university students' careers from 2008 to 2017 in Italy.*

2 Community detection algorithms in tripartite networks

Many real-world networks have a natural multimode network structure in which vertices of different types are linked together. Without reducing generalisability, in the case of tripartite networks, three types of vertices are defined and links can be present only between vertices of distinct types (Fararo & Doreian, 1984). Several approaches can be pursued to disentangle the inherent complexity of such kinds of data. Recently, Everett & Borgatti (2019) suggested that, in the case of multimode data, the collection of all bipartite networks should be examined.
In our case study, a tripartite network is considered in which $\mathscr{V}_P$ is the set of provinces of residence of Italian students enrolled in the first academic year of any bachelor/master degree, $\mathscr{V}_U$ is the set of public and private universities, and $\mathscr{V}_F$ is the set of educational fields of study. The tripartite network $\mathscr{T}$ can be defined as consisting of a pair $(\mathscr{V}, \mathscr{E})$, being $\mathscr{V} = \{\mathscr{V}_P, \mathscr{V}_U, \mathscr{V}_F\}$ the collection of three sets of vertices, one for each mode, and being $\mathscr{E} = \{\mathscr{E}_{PUF}\}$, $\mathscr{E}_{PUF} \subseteq \mathscr{V}_P \times \mathscr{V}_U \times \mathscr{V}_F$, with $\mathscr{E}_{PP}, \mathscr{E}_{UU}, \mathscr{E}_{FF} = \emptyset$, the collection of links among

*This study was supported by the Italian Ministerial grant PRIN 2017 'From high school to job placement: micro-data life course analysis of university student mobility and its impact on the Italian North-South divide', n. 2017HBTK5P - CUP B78D19000180001.

the vertices belonging to the three modes. Given $\mathscr{T}$, a unique supra-adjacency matrix $\mathbb{A}$ could be defined by combining the sociomatrices in a block matrix $\mathbf{A}_{PU}$, $\mathbf{A}_{UF}$, and $\mathbf{A}_{PF}$, where the links are the number of students enrolled, and the corresponding bipartite networks are weighted. Thus, the related supra-adjacency matrix is:

$$\mathbb{A} = \begin{bmatrix} \mathbf{0} & \mathbf{A}_{PU} & \mathbf{A}_{PF} \\ \mathbf{A}_{PU}^T & \mathbf{0} & \mathbf{A}_{UF} \\ \mathbf{A}_{PF}^T & \mathbf{A}_{UF}^T & \mathbf{0} \end{bmatrix}.$$

Over the past two decades, a growing number of studies have been devoted to community detection algorithmic solutions in tripartite graphs. The first and simplest proposed method consists of applying on the matrix $\mathbb{A}$, or on its version built up after matrices' transformation, the usual community detection algorithms (Melamed *et al.*, 2013; Everett & Borgatti, 2019). Other methods adopting an optimisation of tripartite networks (Neubauer & Obermayer, 2009; Ikematsu & Murata, 2013), extending the idea of bipartite modularity.
Given the nature of our data, the approaches which maximise the bipartite modularity seem more appropriate. A detailed comparison of proposed algorithms could be of interest in understanding how tripartite community detection can be used to interpret the network patterns underlying the Italian student mobility phenomenon.

References

COLUMBU, SILVIA, PORCU, MARIANO, & SULIS, ISABELLA. 2021. University choice and the attractiveness of the study area: Insights on the differences amongst degree programmes in Italy based on generalised mixed-effect models. *Socio-Economic Planning Sciences*, **74**, 100926.

EVERETT, MARTIN G, & BORGATTI, STEPHEN P. 2019. Partitioning multimode networks. *Pages 251–265 of: Advances in network clustering and blockmodeling*. John Wiley and Sons.

FARARO, THOMAS J, & DOREIAN, PATRICK. 1984. Tripartite structural analysis: Generalizing the Breiger-Wilson formalism. *Social Networks*, **6**(2), 141–175.

FENG, LIANG, ZHAO, QIANCHUAN, & ZHOU, CANGQI. 2019. An efficient method to find communities in K-partite networks. *Pages 534–535 of: 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE.

GENOVA, VINCENZO GIUSEPPE, TUMMINELLO, MICHELE, ENEA, MARCO, AIELLO, FABIO, & ATTANASIO, MASSIMO. 2019. Student mo-

bility in higher education: Sicilian outflow network and chain migrations. *Electronic Journal of Applied Statistical Analysis*, **12**(4), 774–800.

GIAMBONA, FRANCESCA, PORCU, MARIANO, & SULIS, ISABELLA. 2017. Students mobility: Assessing the determinants of attractiveness across competing territorial areas. *Social indicators research*, **133**(3), 1105–1132.

IGNATOV, DMITRY I, SEMENOV, ALEXANDER, KOMISSAROVA, DARIA, & GNATYSHAK, DMITRY V. 2017. Multimodal clustering for community detection. *Pages 59–96 of: Formal Concept Analysis of Social Networks*. Springer.

IKEMATSU, KYOHEI, & MURATA, TSUYOSHI. 2013. A fast method for detecting communities from tripartite networks. *Pages 192–205 of: International Conference on Social Informatics*. Springer.

IMPICCIATORE, ROBERTO, & PANICHELLA, NAZARENO. 2019. Internal migration trajectories, occupational achievement and social mobility in contemporary Italy. A life course perspective. *Population, Space and Place*, **25**(6), e2240.

MELAMED, DAVID, BREIGER, RONALD L, & WEST, A JOSEPH. 2013. Community structure in multi-mode networks: Applying an eigenspectrum approach. *Connections*, **33**(1), 1823.

NEUBAUER, NICOLAS, & OBERMAYER, KLAUS. 2009. Towards community detection in k-partite k-uniform hypergraphs. *Pages 1–9 of: Proceedings of the NIPS 2009 Workshop on Analyzing Networks and Learning with Graphs*.

RIAÑO, YVONNE, VAN MOL, CHRISTOF, & RAGHURAM, PARVATI. 2018. New directions in studying policies of international student mobility and migration. *Globalisation, Societies and Education*, **16**(3), 283–294.

SANTELLI, FRANCESCO, SCOLORATO, CONCETTA, & RAGOZINI, GIANCARLO. 2019. On the determinants of student mobility in an interregional perspective: a focus on Campania region. *Statistica Applicata - Italian Journal of Applied Statistics*, **31**(1), 119–142.

VAN MOL, CHRISTOF, & TIMMERMAN, CHRISTIANE. 2014. Should I stay or should I go? An analysis of the determinants of intra-European student mobility. *Population, Space and Place*, **20**(5), 465–479.