A multiplex network approach for analyzing university students' mobility flows

Ilaria Primerano, Francesco Santelli and Cristian Usala

Abstract This paper proposes a multiplex network approach to analyze the Italian students' mobility choices from bachelor's to master's degrees. We rely upon administrative data on Italian students' careers by focusing on those who decide to enroll in a different university for their master's studies once they graduate in a bachelor's program. These flows are explored by defining a multiplex network approach where the ISCED-F fields of education and training are the layers, the Italian universities are the nodes, and the weighted and directed links measure the number of students moving between nodes. Network centrality measures and layers similarity indexes are computed to highlight the presence of core universities and verify if the network structures are similar across the layers. The results indicate that each field of study is characterized by its network structure, with the most attractive universities usually located in the Center-North of the country. The community detection algorithm highlights that graduates' mobility between universities is encouraged by the geographical proximity, with different intensities depending on the field of study.

Key words: Community Detection, Layer Similarity, Network Centrality Measures, Students' Mobility

Ilaria Primerano
Department of Political and Social Studies, University of Salerno, Italy e-mail: iprimerano@unisa.it

Francesco Santelli
Department of Political Sciences, University of Naples Federico II, Italy e-mail: francesco.santelli@unina.it

Cristian Usala
Department of Political and Social Sciences, University of Cagliari, Italy e-mail: cristian.usala@unica.it

1 Introduction

Understanding the determinants and the main patterns of students' migration flows has become increasingly important to define the policies related to the attractiveness of the university systems in different areas [1]. Indeed, students' mobility is crucial since it affects several aspects related to regional competitiveness and growth capacity [2, 3] by also anticipating the future migration choices of highly educated individuals [4, 5]. These elements are particularly important if we consider the Italian case, characterized by substantial regional disparities in terms of educational outcomes, access to tertiary education, labor market conditions [6, 7, 8], and an unbalanced flow of students from southern to northern regions [9].

Several scholars have studied the determinants of students' mobility pathways by following different approaches and with different aims. In particular, some scholars have focused on the different factors affecting of mobility flows by investigating the role played by the hosting areas' characteristics [8], students' socio-economic and family background [10], peers' effects [11], and the quality of research [12]. These works have highlighted that universities' attractiveness positively depends on the quality of local amenities and job market opportunities, the quality of research and educational services.

Despite the growing interest in this phenomenon, most of the studies focused on the first level mobility (from high school to bachelor) [9, 13]. In contrast, only a few contributions have analyzed the second level mobility flows (from bachelor to master) by means of longitudinal analysis [14] and by focusing mainly on the choices of southern students [15]. Other studies have analyzed students' mobility flows by means of Social Network Analysis (SNA) [16] considering both international [17] and national levels [18], as well as aggregate data referred to macro-areas such as provinces and regions, and the specific field of study chosen by students [19]. A network perspective is considered to explore the structural patterns of students' mobility flows among Italian geographical areas focusing on the well-known South-North route [20]. Moreover, network measures, such as hub and authority scores, have also been used to analyze the direction and the intensity of student flows in a country and to label territories as good importers and/or good exporters [21]. These techniques have also been applied to reveal the presence of chain migrations phenomenons that link destination and origin geographical areas [22], and to analyze the mobility choices of Italian graduates that are choosing to attend a master's program [23].

Moving from this framework, an important element usually neglected is the presence of students' mobility flows between universities located within the same territory (e.g. city, province or region). In fact, even if a relevant part of the overall flows between universities in Italy is observed among institutions located in the same territory, most of the literature has focused only on the flows between macro areas.

In this paper, we aim to explore the role of Italian universities in the second level mobility network through a multiplex network approach [24]. In particular, students are considered in mobility when, after obtaining a bachelor's degree, they decide to change their university to attend a master's degree program regardless of the distance between origin and destination universities[1]. Moreover, to highlight the differences existing among the fields of study, we define a multiplex network data structure by considering these latter as layers[2] [25], the Italian universities as the nodes, and the flows of students between universities as the weighted and directed links. Based on this network structure, we assess the similarities across fields by computing a set of layer similarities indexes [26] in order to highlight if there is a common pattern or if such fields are rather heterogeneous in terms of observed network structures.

The paper is structured as follows. Section 2 provides the definition and analysis of the multiplex mobility network. Section 3 describes the data used in the analysis and the normalization procedure adopted. Section 4 presents the main results, and concluding remarks are given in Section 5.

2 Multiplex network definition and analysis

A multiplex network is a special case of multilayer networks [24, 27] where each layer holds a common set of nodes connected through two different kinds of relationships: intra-layer connections (i.e. edges linking nodes within the same layer) and inter-layer connections (i.e. edges crossing layers linking nodes in different layers).

Formally, let $\mathcal{M}$ be a multiplex network defined by a set of K different graphs $G(V, E_k)$, with $k = 1, \ldots, K$; where V is the set of common nodes and E_k is the set of both intra-layer edges (E_{kk}) and inter-layers edges (E_{kh}). For each layer k the corresponding adjacency matrix is $A_k = (a_{ijk})$, with $a_{ijk} = 1$ if $(v_i, v_j) \in E_k$, and $a_{ijk} = 0$ otherwise [24].

We define the layers of the Italian second level mobility network by grouping the degree programs in fields of study according to the ISCED-F classification. Since this latter is based on the similarities between programs in terms of disciplinary contents, it allows gathering in each layer the great part of the exchange flows of students between universities.

Thus, each observed layer holds a field-specific network where the set of nodes is represented by the Italian universities and the edges by the flows of students between them. In the second level mobility framework, intra-layer edges are given by the flows of students that change university after graduation but remain in the same field

[1] The definition of mobility adopted in this paper does not depend on the geographical distance between universities or territorial boundaries. For this reason, the Italian online universities are also included in the analysis.

[2] The educational fields are derived according to the 'International Standard Classification of Education: Fields of education and training' (ISCED-F) that was developed by the United Nations Educational, Scientific and Cultural Organization (UNESCO). In this contribution, the following ten broad fields are considered to define the layers of the multiplex network: Agriculture, forestry, fisheries and veterinary ('Agriculture'); Arts and humanities ('Arts'); Education; Engineering; manufacturing and construction ('Engineering'); Health and welfare ('Health'); Information and Communication Technologies ('ICTs'); Natural sciences, mathematics and statistics ('Sciences'); Services, Social sciences, journalism and information ('Social Sciences').

of study; while, inter-layer edges, consist of the flows of graduated students that change both university and field of study.

The resulting network is a weighted multiplex network [28], with the number of students involved in the flows defining the weights of both intra- and inter-layers edges. The edges of these networks are heavily influenced by universities' overall size which, in turn, depends on the number of students enrolled in the considered layers. Indeed, universities that provide programs in the most populated fields are more likely to have a higher number of students in mobility than those that provide programs in less chosen ones.

By construction, these networks are very rich in terms of edges (i.e. the networks are very dense) and thus tend to form a complete graph. Most connections likely consist of just a few students, while big universities will import and export a greater number of students because they can enroll more freshmen. For this reason, comparing layers holding universities with different sizes may not be appropriate since the most chosen universities are likely to rank both as top importers and top exporters in many fields. In order to get a simplified structure for each layer, a two-step procedure is applied. The first step normalizes the edges by defining a link-specific weight that depends on the size of the universities. In the second step, a cutting threshold value is set to identify the most relevant paths in the networks.[3]

Once all the layers have been processed, in depth analysis can be carried out to describe this structure through network comparisons in terms of similarity measures. Comparing networks is an important way to analyze multiplex network data structures, where the definition of similarity measures between layers is a key factor in appreciating their main characteristics. In its simplest form, this comparison consists of flattening all the layers into a single-layer network, where each actor is represented through a unique node linked by weighted edges whose weight depends on the number of layers on which the actors are connected. By contrast, the layer-by-layer methods process each layer separately and then compare the results of the SNA measures. Otherwise, ad hoc measures for multilayer networks have also been proposed by considering the difference between intra-layer and inter-layer edges and making a numerical distinction between layers or by analyzing edges involving different layers together but not mixing up with each other.

In particular, when dealing with layer similarity measures, several approaches have been proposed aiming to highlight the existence of the same structures across the layers and to identify their different characteristics [26]. A different approach has been developed for a visual exploration of the multiplex networks based on the use of factorial methods [29]. In this case, the authors analyzed the adjacency matrices derived by a multiplex network by using the DISTATIS technique [30]. This procedure allows showing the common structure of all layers (intra-layer perspective) and the variation of nodes across layers (inter-layer perspective).

In order to quantify the shared characteristics of these structures across the layers, different similarity measures have been considered in terms of both actor-centered as well as layer-centered perspectives.

[3] See Section 3 for details on the normalization procedure.

In this paper, a layer-by-layer perspective of analysis is applied [26] by focusing only on intra-layer edges to: *i)* highlight the core universities for each field of study using network centrality measures [31]; and *ii)* compare the results obtained across the layers by applying layer similarity measures.

Network centrality indexes have been computed to identify the universities that act as good importers or good exporters. Specifically, as for universities' attractiveness, we have considered the in-degree index, defined as the number of incoming edges. Instead, as a proxy of the export attitude of universities, we have computed the out-degree index, defined as the number of outgoing edges.

Furthermore, to compare the results across the layers, we take advantage of two indexes: the Pearson Degree Similarity coefficient [32] and the Jaccard Layer Correlation Coefficient [26]. The Pearson Degree Similarity Coefficient quantifies the similarity among nodes' degrees across layers and allows to assess universities' centrality across different disciplinary fields. The Jaccard Layer Correlation Coefficient measures the overlapping between pairs of layers. It varies between 0 and 1, with 0 indicating no overlapping and 1 indicating perfect overlapping. This coefficient highlights the presence of edges among the same nodes on different layers, i.e. the presence of edges linking the same group of universities in different disciplinary fields. It is used to identify the presence of common structures between fields and, at the same time, to determine the turnover that takes place between the layers.

Finally, in each field of study, the Clauset-Newman-Moore (CNM) community detection algorithm [33] has been adopted to identify the presence of groups of universities which are tightly connected in knit groups within communities and loosely connected with universities belonging to other communities.

3 Data description and normalization procedure

Italian students' second level mobility network is constructed starting from the database MOBYSU.IT [34] that holds the administrative data regarding all Italian university students.[4] In particular, this database includes information on students' careers that allows us to identify the universities in which the students have obtained their bachelor's degrees and those chosen for their master's studies.

We extracted the data regarding the cohorts of Italian students who started their university careers in a bachelor's program between a.y. 2011/2012 and a.y. 2015/2016, and have enrolled in a master's degree program between a.y. 2014/2015 and a.y. 2018/2019. Moreover, since students can enroll in different universities during their careers, in the network definition we consider the origin university the one in which she/he has obtained her/his last bachelor's degree. Namely, if a student has gained two or more bachelor's degrees in his/her career, we consider as the origin

[4] Data drawn from the Italian 'Anagrafe Nazionale della Formazione Superiore' has been processed according to the research project 'From high school to the job market: analysis of the university careers and the university North-South mobility' carried out by the University of Palermo (head of the research program), the Italian 'Ministero Università e Ricerca', and INVALSI.

node only the last university in which the student has been observed. Then, since students may also enroll in several master's programs, we consider as destination node only the first university in which the student has started her/his master's programs.

Therefore, starting from the population of 1,171,006 Italian bachelor's students, the dataset holds information about 621,075 (53%) students that have graduated in the time frame considered. Secondly, we keep the information regarding only those students that have enrolled in a master's program. Our analysis is based on 367,725 students, belonging to 92 universities (of which 11 are online universities).

Table 1 presents the descriptive statistics of the distribution of students among fields of study. In particular, for each ISCED-F field, the table shows the number of graduated students enrolled in a master's program in Italy (Master's students); the number and the percentage of students that have changed university for attending a master's program (Students in mobility); and the percentage of students in mobility that have chosen to stay in the same field of study (same field).

Table 1: Distribution of students in mobility according to ISCED-F fields

Code	ISCED - F 2013	Master's students N.	Students in mobility N.	Students in mobility (%)	Same field %
L1	Agriculture, forestry, fisheries and veterinary	10,959	2,631	(24.0%)	65.8%
L2	Arts and humanities	63,756	19,440	(30.4%)	72.4%
L3	Business, administration and law	59,949	15,849	(26.4%)	78.1%
L4	Education	12,728	3,245	(25.4%)	83.4%
L5	Engineering, manufacturing and construction	82,753	12,319	(14.8%)	94.3%
L6	Health and welfare	11,258	4,309	(38.2%)	61.1%
L7	Information and Communication Technologies (ICTs)	4,339	816	(18.8%)	80.6%
L8	Natural sciences, mathematics and statistics	41,195	10,945	(26.5%)	93.1%
L9	Services	12,484	3,979	(31.8%)	37.0%
L10	Social sciences, journalism and information	68,304	23,713	(34.7%)	75.4%

Table 1 shows that the two most chosen groups are 'Engineering' and 'Social sciences' while the least chosen is 'ICTs'. Regarding the second level mobility, we observe that at least 14.8% of students in each field have changed university after graduation, with a maximum of 38.2% in 'Health'. Considering the last two columns, it is noticeable that the tendency to change the field of study varies depending on the field considered, with the 94.3% of 'Engineering' graduates that have not changed their field after graduation, whereas the majority of students in the 'Services' field (63.0%) have decided to enroll in a different field.

In order to analyze the second level mobility networks we apply a two-step procedure. The first step consists in normalizing the weights given by the number of students involved in the flows. Following Slater (2009) [35], we apply a Multidimensional Iterative Proportional Fitting Procedure (MIPFP) [36]. This procedure sets as seed the original adjacency matrix, where the rows' and columns' marginals are, respectively, the total number of incoming and outgoing students for a given university, and then performs the reshaping. The set of known desired target marginal counts M is a non-empty subset (2–dimensional) of:

$$\tau = \left\{(a_{\bullet j}), (a_{i\bullet})\forall i, j\right\} \tag{1}$$

where $\bullet$ is the summation over the corresponding university. The procedure iteratively updates the values of the cells depending on the targets. The first iterations for both marginals are:

$$a_{ij}^{1} = a_{ij}^{0} \cdot \frac{a_{\bullet j}}{a_{\bullet j}^{0}} \quad \forall i, j \quad ; \quad a_{ij}^{2} = a_{ij}^{1} \cdot \frac{a_{i\bullet}}{a_{i\bullet}^{1}} \quad \forall i, j \tag{2}$$

The adjustments at a generic iteration l stop when the stopping criterion given by the *tol* parameter (a small constant) is reached:

$$max\left|a_{ij}^{l-1} - a_{ij}^{l}\right| \leq tol \quad \forall i, j \tag{3}$$

This procedure defines a value for each edge ranging from 0 to 1 accounting for nodes' attractiveness (columns marginal) and nodes' export attitude (rows marginal). Thus, at the end of the normalization, each weight is a value that takes into account both the overall number of incoming and outgoing edges. Namely, edges' weights inversely depend on the number of students: higher weights are associated to universities with a small number of flows; lower weights are attached to universities characterized by a relatively large number of incoming and outgoing students.

Since the normalized network is very dense, a cut-off threshold is set at the median value of the non-zero normalized entries in the second step. This value is used to dichotomize the obtained weights in order to let the most relevant flows emerge.

4 Main results

The multiplex mobility network under analysis consists of 10 layers and 5.689 total intra-links. Table 2 shows the descriptive statistics on the main structural characteristics of the networks. Some network measures have been computed at the network level for each layer and the flattened network. In particular, the table shows the number of actors (n), the number of edges (m), the number of strong components (nc), the density (den), and the clustering coefficient (cc). For all the networks, given the normalization procedure, a low value of the density is observed, with maximum values of 0.13 for both 'Health' and 'Social Sciences' and a minimum value of 0.06 for 'Agriculture'. At the same time, the global clustering coefficient is almost homogeneous across the layers, ranging from 0.31 to 0.46.

Moreover, to highlight the role played by each university in all fields of study, we have computed the centrality measures for every single layer and the flattened network. In particular, in-degree and out-degree centrality measures have been com-

Table 2: Descriptive measures on Student Mobility Multiplex Network by fields of study

Layers		n	m	nc	den	cc
	Flattened network	92	5,689	1	0.68	0.66
L1	Agriculture, forestry, fisheries and veterinary	52	168	2	0.06	0.35
L2	Arts and humanities	88	868	1	0.11	0.31
L3	Business, administration and law	87	827	1	0.11	0.32
L4	Education	57	251	1	0.08	0.46
L5	Engineering, manufacturing and construction	81	498	2	0.08	0.35
L6	Health and welfare	71	663	1	0.13	0.40
L7	Information and Communication Technologies (ICTs)	79	562	1	0.09	0.32
L8	Natural sciences, mathematics and statistics	62	387	1	0.10	0.41
L9	Services	75	490	2	0.09	0.46
L10	Social sciences, journalism and information	80	835	1	0.13	0.40

puted to identify the most attractive universities (i.e. high in-degree value) and those losing students (i.e. high out-degree value).

Concerning the flattened network, the top five universities in terms of attractiveness are Pisa, Bologna, Florence, Milan, and Turin. In contrast, the top five exporters are Parma, Catania, Pisa, Florence and Modena. Although some of these are so-called big universities, for which it is customary to expect a larger number of enrolled students, it is worth noting that the two-step normalization procedure takes into account the overall size of universities. In fact, the flattened network results show the presence of smaller universities as well.

As regards single-layer networks, the in-degree centrality results show that the top five universities are very different among the 10 layers considered, indicating that each field is characterized by its peculiar structure. Specifically, few universities are present in more than one layer. For example, Pisa ranks in the top positions in five different layers: ‘Arts’, ‘Education’, ‘Engineering’, ‘ICTs’, and ‘Social sciences’. It is also remarkable that in the top positions for some layers, we find several online universities (UNITEL, Niccolò Cusano, Pegaso, and UniNettuno) and universities located in metropolitan areas (e.g. Turin, Milan, and Rome).

The out-degree results also emphasize the heterogeneity among the layers, but involving different groups of universities. Indeed, the out-degree values reflect the well-known pattern of Italian student migration, with many outgoing flows originating in Southern Italy universities that are directed towards the universities of central and northern Italy.

Considering, for instance, the two most chosen fields by Italian mover students, ‘Social Sciences’ and ‘Engineering’, the top five attractive universities in ‘Social Sciences’ field are IULM of Milan, Carlo Bo of Urbino, Perugia, Pisa and Siena, while the top five exporters are the universities of Bologna, Brescia, Modena, Parma and Pavia. Instead, the mobility flows in the ‘Engineering’ field show as the top five attractive universities La Sapienza of Rome, Bologna, Milan, Pisa, and Turin, while those losing students are Bicocca of Milan, Padua, Salerno, Trento and Udine.

Overall, the single-layer analysis shows that it is uncommon for a university to be in a leading position in all the fields. Indeed, the high position of a university in the ranking of the most attractive ones could be related to the presence of a 'department of excellence' that could act as a pole of attraction for students, making the university to which they belong a central node in that specific layer. What is instead very interesting is the centrality of the online universities in different layers in which they act primarily as importers rather than exporters in the second level mobility framework.

Moving to the multiplex mobility network analysis, the two similarity measures described in Section 3 have been computed to compare layers' properties.

Table 3: Correlation matrix of the Pearson Degree Similarity coefficient by fields of study

		L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
L1	Agriculture	1.00	-	-	-	-	-	-	-	-	-
L2	Arts	0.56	1.00	-	-	-	-	-	-	-	-
L3	Business	0.41	0.52	1.00	-	-	-	-	-	-	-
L4	Education	0.57	0.28	0.35	1.00	-	-	-	-	-	-
L5	Engineering	0.54	0.41	0.41	0.47	1.00	-	-	-	-	-
L6	Health	0.45	0.34	0.51	0.39	0.41	1.00	-	-	-	-
L7	ICTs	0.74	0.56	0.65	0.50	0.49	0.57	1.00	-	-	-
L8	Sciences	0.51	0.40	0.47	0.33	0.50	0.33	0.49	1.00	-	-
L9	Services	0.38	0.63	0.58	0.23	0.38	0.43	0.49	0.46	1.00	-
L10	Social sciences	0.55	0.51	0.55	0.32	0.43	0.48	0.60	0.46	0.63	1.00

For what concerns the Pearson Degree Similarity coefficient, reported in Table 3, the higher values are observed for the 'ICTs' and other fields: 'Agriculture' (0.74), 'Business' (0.65), and 'Social Sciences' (0.60). Other high correlations involve 'Services' with 'Arts' (0.63) and 'Social sciences' (0.63) fields. With respect to the Jaccard coefficient, it is noticeable that all the values reported in Table 4 are very close to 0. These results perfectly align with what emerged in the single-layer analysis, where the central actors differ among the observed layers. In other words, the mobility flows considered, rather than defining a common set of core universities for all the ISCED-F fields, highlight the presence of different typical structures in each specific field.

In the following, the results obtained for a specific layer are presented to show the application scope of our analysis procedure. Specifically, we consider the normalized network of the layer 'Education' whose representation is given in Figure 1. This graph shows the edges connecting the 57 Italian universities among which the most important exchange flows occur. The size of each node indicates the level of attractiveness of the university in this field, measured through the in-degree centrality index. Moreover, nodes are colored according to the cluster (i.e. communities) they lie in. In this layer six communities have been identified by means of The Clauset-Newman-Moore (CNM) community detection algorithm.

Table 4: Correlation matrix of the Jaccard edge similarity coefficient by field of study

		L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
L1	Agriculture	1.00	-	-	-	-	-	-	-	-	-
L2	Arts	0.15	1.00	-	-	-	-	-	-	-	-
L3	Business	0.09	0.11	1.00	-	-	-	-	-	-	-
L4	Education	0.10	0.06	0.07	1.00	-	-	-	-	-	-
L5	Engineering	0.08	0.05	0.06	0.08	1.00	-	-	-	-	-
L6	Health	0.11	0.10	0.12	0.06	0.06	1.00	-	-	-	-
L7	ICTs	0.16	0.15	0.12	0.08	0.08	0.12	1.00	-	-	-
L8	Sciences	0.12	0.11	0.09	0.05	0.07	0.09	0.12	1.00	-	-
L9	Services	0.13	0.18	0.10	0.06	0.04	0.10	0.13	0.12	1.00	-
L10	Social sciences	0.14	0.17	0.10	0.06	0.04	0.11	0.13	0.11	0.17	1.00

By exploring the patterns outlined by the flows of students who change university for their master's studies while remaining in the Education field, the geographical aspect of mobility clearly emerges. This element is mainly due to the fact that students also move into the same geographical boundaries. In fact, looking at the six communities, this aspect is evident. The dark green group on the top-left of Figure 1 consists of the universities that are located in the central and southern regions of the eastern part of Italy, such as Apulia, Molise, Abruzzo and Marche regions. Some of these universities, i.e. Foggia and Bari, have many incoming and outgoing flows.

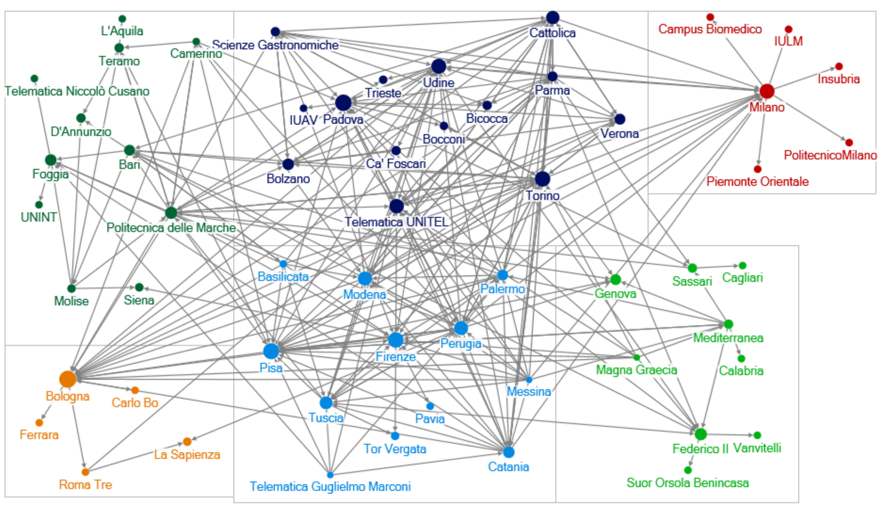

Fig. 1: Single Layer Visualizations of the Student Mobility Network for the Education field of study. Nodes' colors depend on the Community; nodes' size depends on the in-degree centrality index

Instead, the Politecnica of Marche has a strong attitude toward attracting students, even from universities located outside the community.

The first community (orange), on the bottom-left of the Figure 1, has Bologna university as the attractive pole of the community that also includes other universities from Emilia-Romagna, Lazio, and the Urbino University.

Indeed, on the top-right, the second community (red) groups most of the universities located in the Lombardy region, whose edges form a star-shaped configuration around the university of Milan. Thus, in this community, the centrality role of the university of Milan clearly emerges. At the same time, the university of Milan acts as a local hub by attracting students from all the other regions and yielding a quota of its graduates to other universities in the same geographic area.

The third community (green), on the bottom-right, is mainly composed of the universities located in Campania, Sardinia, and Calabria. This is a dense community, with many flows within it, in which emerges the central role of the Federico II University of Naples characterized by many incoming flows.

Finally, the two big communities in the center of Figure 1 are related to North universities (top-center, dark blue color) and Center and South universities (top-center, light blue color). These communities show many incoming and outgoing flows without a specific single region involved but showing high values for both in-degree and out-degree indexes.

As the last remark, Figure 2 shows, for each field of study, the skeleton graph of the communities, where each community is a vertex. This visualization highlights the relationships among different communities (i.e. flows of students among communities) and the isolated ones characterized only by internal flows of students (i.e. flows of students within the same community). Finally, the size of the vertexes is proportional to the number of internal flows. The skeleton representation in Figure 2 of the six communities belonging to the field 'Education' allows us to notice that the first and the second communities are reached by all the other ones, thus meaning that these communities include the most attractive universities for the considered layer. Indeed, the second community holds the university of Milan, while the first the one of Bologna, which are the two most popular universities in this layer.

5 Concluding remarks

In this contribution, we have introduced the study of second level student mobility into the framework of multiplex network analysis. Our results show clear hints on the students' flows at the moment of enrollment in a master's degree program. Such kind of mobility has not been explored deeply in the Italian context. To the best of our knowledge, this is the first study on the Italian second level mobility that accounts, at the same time, for several aspects: including online institutions, considering the differences among fields of study, and neglecting the geographical setting.

Furthermore, in line with previous studies [37, 21], the evidence obtained through the community detection analysis shows that one of the most important factors

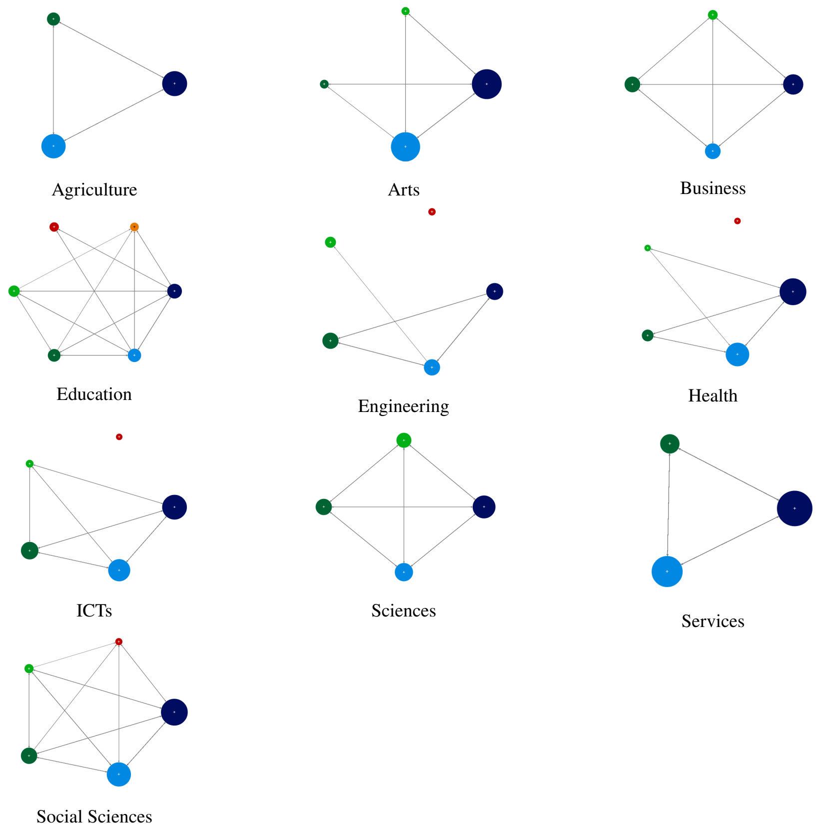

Fig. 2: Skeleton graph representation of the communities detected for each field of study.

that encourage second level mobility is the geographic proximity. Moreover, most attractive universities are located in the northern and central regions. However, such geographical influences on students' flows have a different intensity depending on the field of study considered. In fact, the layer comparison analysis has clearly shown that the layers differ in terms of their network structures and their flow dynamics. This element may be due also to the fact that the supply of degree programs in Italy is heterogeneous and some universities provide only a few programs in specific fields.

Several aspects could be considered when dealing with students' mobility data organized into multiplex structures, such as the analysis of the inter-layers connections to understand the determinants of students' decisions to change their field after graduation. Moreover, future lines of research include the study of the relationship between the overall universities' attractiveness scores, their supply of educational services and hosting areas' characteristics, and the assessment of the sensitivity of the results to different normalization procedures and cutting thresholds.

Acknowledgements This contribution has been supported from 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.

References

1. Attanasio, M.: Students' university mobility patterns in Europe: an introduction. Genus, **78(17)**, (2022)
2. Krugman, P.: Increasing Returns and Economic Geography. J Polit Econ, **99(3)**, 483-99, (1991)
3. Valero, A., & Van Reenen, J.: The economic impact of universities: Evidence from across the globe. Econ Educ Rev, **68**, 53–67, 68, (2019)
4. Dotti, N. F., Fratesi, U., Lenzi, C., & Percoco, M.: Local Labour Markets and the Interregional Mobility of Italian University Students. Spat Econ Anal, **8(4)**, 443–468, (2013)
5. Oggenfuss, C., & Wolter, S. C.: Are they coming back? The mobility of university graduates in switzerland. Rev Reg Stud, **39**, 189–208, (2019)
6. Türk, U.: Socio-Economic Determinants of Student Mobility and Inequality of Access to Higher Education in Italy. Netw and Spat Econ, **19(1)**, 125–148, (2019)
7. Fratesi, U., & Percoco, M.: Selective Migration, Regional Growth and Convergence: Evidence from Italy. Reg Stud, **48(10)**, 1650–1668, (2014)
8. Giambona, F., Porcu, M., & Sulis, I.: Students Mobility: Assessing the Determinants of Attractiveness Across Competing Territorial Areas. Soc Indic Res, **133(3)**, 1105–1132, (2017)
9. Attanasio, M., Enea, M. & Priulla, A.: Quali atenei scelgono i diplomati del Mezzogiorno d'Italia?. Neodemos, ISSN: 2421-3209, (2019)
10. Impicciatore, R., & Tosi, F.: Research in Social Stratification and Mobility Student mobility in Italy: The increasing role of family background during the expansion of higher education supply. Res Soc Stratif Mobil, **62**, 100409, (2019)
11. Porcu, M., Sulis, I. & Usala, C.: Estimating the peers effect on students' university choices. In R. Lombardo, I. Camminatello, V. Simonacci, editors, Book of short Papers IES 2022: Innovation & Society, 134-139, (2022)
12. Bratti, M., & Verzillo, S.: The 'gravity' of quality: research quality and the attractiveness of universities in Italy. Reg Stud, **53(10)**, 1385–1396, (2019)
13. D'Agostino, A., Ghellini, G., & Longobardi, S.: Exploring determinants and trend of STEM students internal mobility. Some evidence from Italy. Electron J Appl Stat Anal **12(4)**, 826–845, (2019)
14. Enea, M.: From South to North? Mobility of southern italian students at the transition from the first to the second level university degree. In C. Perna, M. Pratesi, and A. Ruiz-Gazen, editors, Studies in Theoretical and Applied Statistics, 239–249, (2018)
15. Attanasio, M., Enea, M., & Albano, A.: Dalla triennale alla magistrale: continua la 'fuga dei cervelli' dal Mezzogiorno d'Italia. Neodemos, ISSN: 2421-3209, (2019)
16. Wasserman, S., & Faust, K.: Social network analysis: Methods and applications, Cambridge University Press, (1994)
17. Restaino, M., Vitale, M., & Primerano, I.: Analysing international student mobility flows in higher education: A comparative study on european countries. Soc Indic Res, **149(3)**, 947–965, (2020)
18. Santelli, F., Scolorato, C., & Ragozini, G.: On the determinants of students mobility in an interregional perspective: A focus on campania region. J Appl Stat, **31(1)**, 119–142, (2019)
19. Columbu, S. & Primerano, I.: A multilevel analysis of university attractiveness in the network flows from bachelor to master's degree. In A. Pollice, N. Salvati, and F. Schirippa Spagnolo, editors, Book of short Papers SIS 2020, pages 480–485, (2020)

20. Vitale, M., Giordano, G., & Ragozini, G.: University student mobility flows and related network data structure. In A. Pollice, N. Salvati, and F. Schirippa Spagnolo, editors, Book of short Papers SIS 2020, pages 515–520, (2020)
21. Columbu, S., Porcu, M., Primerano, I., Sulis, I., & Vitale, M. P. Analysing the determinants of Italian university student mobility pathways. Genus, **78(6)**, (2022)
22. Genova, V. G., Tumminello, M., Enea, M., Aiello, F., & Attanasio, M.: Student mobility in higher education: Sicilian outflow network and chain migrations. Electron J Appl Stat Anal **12(4)**, 774–800, (2019)
23. Columbu, S., Porcu, M., Primerano, I., Sulis, I., & Vitale, M. P.: Analysing the determinants of Italian university student mobility pathways. Genus, 78(6), (2022)
24. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A.: Multilayer networks. J Complex Netw, **2**, 203-271, (2014)
25. UNESCO Institute for Statistics. ISCED Fields of Education and International Standard Classification of Education 2011, Montrèal (2014)
26. Bródka, P., Chmiel, A., Magnani, M., & Ragozini, G.: Quantifying layer similarity in multiplex networks: a systematic study. R Soc Open Sci, **5**, 171747, (2018)
27. Dickison, M. E., Magnani, M., & Rossi, L.: Multilayer social networks. Cambridge University Press, (2016)
28. Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R. J., & Bianconi, G.: Weighted multiplex networks. PloS one, **9(6)**, e97857, (2014)
29. Giordano, G., Ragozini, G., & Vitale, M. P.: Analyzing multiplex networks using factorial methods. Soc Netw **59**, 154-170, (2019)
30. Abdi, H., Dunlop, J.P., Williams, L.J.: How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the bootstrap and 3-way multidimensional scaling (DISTATIS). NeuroImage **45**(1), 89–95 (2009)
31. Freeman, L. C.: Centrality in social networks conceptual clarification, Soc Netw, **1(3)**, 215-239, (1978)
32. Berlingerio, M., Coscia, M., Giannotti, F.: Finding and characterizing communities in multidimensional networks. In: 2011 International Conference on Advances in Social Networks Analysis and Mining. IEEE. pp. 490–494, (2011)
33. Clauset, A., Newman, M. E., & Moore, C: Finding community structure in very large networks. Phys Rev E, **70(6)**, 066111. (2004)
34. Database MOBYSU.IT [Mobilità degli Studi Universitari in Italia], research protocol MUR - Universities of Cagliari, Palermo, Siena, Torino, Sassari, Firenza, Cattolica and Napoli Federico II, Scientific Coordinator Massimo Attanasio (UNIPA), Data Source ANS-MUR/CINECA.
35. Slater, P. B.: Multiscale network reduction methodologies: Bistochastic and disparity filtering of human migration flows between 3,000+ us counties. arXiv preprint arXiv:0907.2393 (2009)
36. Barthélemy, J., & Suesse, T.: mipfp: An R package for multidimensional array fitting and simulating multivariate Bernoulli distributions. J Stat Softw, **86(2)**, (2018)
37. Santelli, F., Ragozini, G., & Vitale, M. P. Assessing the effects of local contexts on the mobility choices of university students in Campania region in Italy. Genus, **78(5)**, 1-25, (2022)