Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Tracking Structural Evolution in Higher Education Mobility -- A Comparative Analysis of Graph Distance Metrics (Hungary, 2006--2024)

Zsolt T. Kosztyán

kosztyan.zsolt@gtk.uni-pannon.hu

University of Pannonia

András Hosznyák

University of Pannonia

Tünde Király

University of Pannonia

Attila I. Katona

University of Pannonia

Dénes M. Kornél

Eötvös Loránd University

Gergő Hornák

Eötvös Loránd University

Research Article

Keywords: network, graphon, mobility, early career

Posted Date: July 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9148720/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Tracking Structural Evolution in Higher Education Mobility – A Comparative Analysis of Graph Distance Metrics (Hungary, 2006–2024)

Zsolt T. Kosztyán[1,2,3*], András Hosznyák[1,4], Tünde Király[1], Attila I. Katona[1], Dénes M. Kornél[5], Gergő Hornák[6]

[1]Department of Quantitative Method, University of Pannonia, Veszprém, Hungary, 8200, Egyetem str 10.
[2]Institute of Advanced Studies Kőszeg (iASK), Kőszeg, Hungary, 9730, Charnel str 14.
[3]Wekerle International University, Budapest, Hungary, 1083, Jázmin str 10.
[4]Education Office, Budapest, Hungary, 1363, Szalay str 10-14.
[5]Faculty of Law, Institute of Political Science, Eötvös Loránd University, Budapest, Hungary, 1097, Tóth Kálmán str 10.
[8]Faculty of Informatics, Eötvös Loránd University, Budapest, Hungary, 1117, Pázmány Péter sqr 1/C.

*Corresponding author(s). E-mail(s): kosztyan.zsolt@gtk.uni-pannon.hu;
Contributing authors: hosznyak.andras@gtk.uni-pannon.hu; kiraly.tunde@gtk.uni-pannon.hu; katona.attila@gtk.uni-pannon.hu; dkornel98@gmail.hu; hornak.gergo4@gmail.hu;

Abstract

Higher education mobility is a key driver of regional development and human capital formation. While traditional approaches—gravity models, regression-based analyses—capture aggregate determinants of student flows, they are limited in revealing the structural architecture of mobility networks and its evolution over time. This study introduces a comprehensive framework for tracking structural changes in domestic higher education application networks using graph-theoretic and graphon-based distance metrics. We model yearly application data from Hungary (2006–2024) as weighted, directed graphs aggregated over 175 micro-regions and analyze them along three dimensions: total applications, preference-ranked

applications (1st, 2nd, 3rd choice), and field-of-study disaggregation. We compute an extensive suite of network-level and node-level indicators, and extend the analysis by estimating graphon representations for each yearly network to quantify structural distances via spectral, cut, Frobenius, Wasserstein, Jensen–Shannon, and total variation metrics. Our results show that graphon-based distances capture global structural shifts—such as those induced by the Bologna transition, STEM-favouring funding reforms, COVID-19, and the foundation-model university transformation—that traditional network statistics detect only partially or with delay. We identify which metric families are most sensitive to different types of policy-induced structural change and demonstrate that graphon distances provide a unifying lens for comparing mobility network evolution. Furthermore, graphon-based community detection reveals a hierarchical layering of the mobility network—distinguishing Budapest, major university cities, and smaller institutions as structurally distinct tiers—that conventional modularity-based methods fail to capture. This work contributes to the emerging literature on temporal network analysis in education systems and offers a replicable methodological blueprint for applied network scientists.}

Keywords: network, graphon, mobility, early career

JEL Classification: D8 , H51

MSC Classification: 35A01 , 65L10

1 Introduction

The spatial mobility of students constitutes one of the most consequential forms of human capital movement (Tan et al, 2024), shaping regional development trajectories (Chankseliani and Kwak, 2025), institutional competitiveness, and long-term demographic patterns (Findlay et al, 2012; Choudaha, 2017). Understanding how students select higher education institutions—and how these choices aggregate into system-wide mobility patterns—has attracted sustained scholarly attention from economists, geographers, sociologists, and, more recently, network scientists.

1.1 Approaches to studying student mobility

Traditional studies of student mobility rely on *gravity models* that relate flows between origins and destinations to masses (e.g., population, institutional capacity) and distances (Sá et al, 2004; Di Pietro and Perez-Encinas, 2025). *Econometric approaches* estimate the determinants of mobility using individual- or aggregate-level regressions (Pérez et al, 2025), identifying factors such as tuition fees, institutional prestige, labour-market returns, and regional amenities (Beine et al, 2014; González et al, 2011). *Sociological accounts* (Carlson et al, 2025) emphasise the role of cultural and social capital (Van Mol and Timmerman, 2014; Findlay et al, 2012), while *geographic perspectives* highlight spatial inequality and access (Avola and Giuga, 2025; Vickers and Rees, 2017; Perkins and Neumayer, 2015).

These approaches have been invaluable for isolating causal mechanisms, yet they share a common limitation: they treat each origin–destination dyad as an independent observation, thereby discarding the *relational structure* that emerges when all flows are considered simultaneously. A university that attracts students from a few nearby regions plays a fundamentally different structural role than one that draws from the entire country, even if their total enrollment figures are comparable. Capturing such distinctions requires methods that model the *entire architecture* of the mobility system.

1.2 Network perspectives on mobility

Network analysis offers precisely this system-level perspective. By representing regions (or institutions) as nodes and student flows as weighted, directed edges (Chankseliani and Kwak, 2025), one obtains a graph whose topological properties encode information about hierarchy, clustering, resilience, and structural change (Newman, 2018; Barabási, 2016). Pioneering applications include Shields (2013), who analysed international student flows as a global network, revealing core–periphery structures and regional clustering; Kondakci et al (2018), who used social network analysis to identify emerging regional hubs; and Barnett et al (2016), who traced the evolution of international student flows over two decades. At the European level, Derzsy et al (2012) and Breznik and Gologranc (2017) and later Kropp and Larsen (2025) investigated Erasmus mobility using network methods.

Despite these advances, the existing network literature on student mobility exhibits several important gaps.

RG_1 *Limited application of graphon theory to mobility networks.* Graphon theory (Lovász and Szegedy, 2006; Lovász, 2012; Borgs et al, 2008) provides a powerful nonparametric framework for representing and comparing networks of different sizes and densities. Although graphon methods have seen rapid development in statistics and machine learning (Gao et al, 2015; Klopp et al, 2017; Airoldi et al, 2013), they have not been applied to higher education mobility networks. Without graphon-based representations, one cannot rigorously quantify structural similarity or dissimilarity between networks observed at different time points or across different subpopulations.

RG_2 *Lack of systematic comparison of graphon distance metrics for detecting structural changes in mobility.* Several graphon distance metrics exist—spectral, cut, Frobenius, Wasserstein, Jensen–Shannon—each capturing different aspects of structural (dis)similarity (Wills and Meyer, 2020; Donnat and Holmes, 2018; Tantardini et al, 2019). Yet no study has systematically compared their sensitivity to policy-induced structural changes in mobility networks. It remains unclear which metric best captures which type of change.

RG_3 *Insufficient sub-national, longitudinal network analysis of domestic mobility.* Most network studies focus on international mobility or adopt a cross-sectional design. Longitudinal, sub-national analyses that track how the *internal structure* of a country's higher education mobility network evolves—year by year, preference by preference, field by field—are rare (Kosztyán et al, 2021).

RG$_4$ *Missing integrative framework linking traditional network statistics to graphon-based structural analysis.* Traditional network metrics (centralization, modularity, efficiency, resilience) and graphon distances address complementary aspects of network structure. No existing study provides a unified comparison showing which traditional metrics align with—or diverge from—graphon-based structural distances, thereby limiting our understanding of what each approach uniquely reveals.

To address these gaps, we formulate the following research questions:

RQ (Main): How can structural changes in domestic higher education mobility networks be most effectively captured and compared using graph-theoretic and graphon-based distance metrics?

This overarching question decomposes into three sub-questions:

RQ$_1$ How do the structural patterns of Hungarian higher education application networks evolve over the period 2006–2024, and which policy events are associated with detectable structural shifts? (Addresses RG$_3$)

RQ$_2$ Which network-level and node-level indicators are most sensitive to different types of policy-induced structural changes, and what aspects of structural change does each metric family capture? (Addresses RG$_2$, RG$_4$)

RQ$_3$ To what extent do graphon-based distance metrics reveal structural changes that traditional network metrics cannot, and how do the two approaches complement each other? (Addresses RG$_1$, RG$_4$)

2 Background

2.1 Overview of approaches to studying student mobility

Research on student mobility spans multiple disciplines, each employing distinct methodological toolkits. Table 1 provides a structured overview.

This study privileges a network-analytic framework for several reasons. First, mobility data are inherently *relational*: each application connects an origin micro-region to a destination micro-region. A network representation preserves the full relational structure, enabling analyses of centrality, community structure, and resilience that would be impossible in a regression framework. Second, network metrics offer natural tools for detecting *structural change*: shifts in centralization indicate changing hierarchies; modularity changes reflect altered regional clustering; and resilience metrics reveal whether the system has become more or less vulnerable to disruptions. Third, the graphon framework provides a mathematically rigorous basis for comparing networks across time, preference levels, and fields of study.

2.2 Network approaches to mobility: State of the art

Table 2 summarizes key network-based studies of student mobility.

2.3 Graphon theory: A brief overview

A *graphon* (graph function) is a symmetric, measurable function $W : [0,1]^2 \to [0,1]$ that serves as the continuum limit of a sequence of dense graphs (Lovász and Szegedy,

Table 1: Approaches to studying student mobility.

Approach	Key aspect captured	Representative studies	Temporal analysis?	Limitations
Gravity models	Distance decay, mass effects, attractiveness	Sá et al (2004); Beine et al (2014); Civera et al (2023)	Rarely	Ignores relational structure; dyad-level only
Econometric / regression	Causal determinants of choice	González et al (2011); Beine et al (2014); Pérez et al (2025)	Panel designs	Cannot reveal structural topology
Sociological / qualitative	Cultural capital, motivations	Van Mol and Timmerman (2014); Findlay et al (2012); Chankseliani and Kwak (2025)	Limited	Small samples; not system-level
Geographic / spatial	Accessibility, spatial inequality	Vickers and Rees (2017); Perkins and Neumayer (2015); Litmeyer et al (2023)	Cross-sectional	Limited structural analysis
Network analysis	Relational structure, topology	Shields (2013); Kondakci et al (2018); Weber and Van Mol (2023)	Emerging	Metrics not yet unified; graphon methods unexplored

Table 2: Network-based studies of student mobility.

Study	Scope	Network type	Temporal?	Key metrics	Limitations
Shields (2013)	International flows	Directed, weighted	No	Centrality, core–periphery	Cross-sectional; no structural distances
Kondakci et al (2018)	International flows	Directed, weighted	Partial	Centrality, clustering	No formal structural comparison
Barnett et al (2016)	International flows	Directed, weighted	Yes	Centralization	No graphon; limited metric set
Derzsy et al (2012)	Erasmus (Europe)	Directed, weighted	No	Degree, clustering	Single year; no sub-national
Kosztyán et al (2021)	Hungary (domestic)	Directed, weighted	Yes	Community, centrality	No graphon distances
This study	Hungary (domestic)	Directed, weighted	Yes (2006–2024)	Full metric suite + graphon distances	Structural, but no spatio-structural analysis

2006; Lovász, 2012). Intuitively, $W(x, y)$ represents the probability (or intensity) of a connection between "types'' x and y in an infinite population. Every finite graph G on n vertices can be associated with a step-function graphon by partitioning $[0, 1]$ into n equal intervals and setting the value in each cell to the corresponding edge weight (or adjacency indicator).

The foundational insight of Borgs et al (2008, 2012) was that sequences of dense graphs converge if and only if the corresponding graphon sequence converges in the *cut metric*. This result was a theoretical breakthrough because it provided, for the first time, a rigorous framework for saying that two large networks are "close'' in a structural sense, even when they differ in size.

Extensions of graphon theory have broadened its applicability considerably:

- **Directed graphs**: The graphon generalises to a bi-variate kernel $W : [0, 1]^2 \to [0, 1]$ without the symmetry constraint, i.e., $W(x, y) \neq W(y, x)$ in general (Diaconis and Janson, 2008; Kunszenti-Kovács et al, 2022).
- **Weighted graphs**: The range of W extends to $[0, C]$ or even $\mathbb{R}_{\geq 0}$ for unbounded weights.
- **Sparse graphs**: Borgs et al (2019) developed an L^p theory that accommodates sparse graph sequences where edge density tends to zero.

Despite this intensive theoretical development, graphon methods have not been applied to mobility networks. This is a significant omission: mobility networks exhibit precisely the kind of evolving, size-varying, structurally rich patterns that graphon theory was designed to handle. Without graphon-based distances, researchers are confined to comparing networks via summary statistics (e.g., modularity, centralization), which can be identical for structurally very different networks. Graphon distances, by contrast, compare the *full generating mechanism* of the network.

2.4 Policy context: Hungarian higher education 2006–2024

The study period encompasses several major policy interventions that plausibly altered the structural configuration of the mobility network:

1. **Bologna transition (2006–2008):** Hungary adopted the two-cycle (BSc/MSc) system in 2006. Master's programmes launched in 2008, potentially shifting application patterns.
2. **Financial crisis (2008–2009):** The global financial crisis, which reached Hungary in 2009, may have altered students' geographic and institutional choices due to constrained household budgets.
3. **Funding reform favouring STEM (2011):** State-funded places in business, law, and social sciences were sharply reduced; minimum admission scores in these fields were raised. The explicit aim was to redirect applicants toward STEM fields.
4. **COVID-19 pandemic (2019–2021):** The shift to online education may have altered the perceived importance of geographic proximity, potentially expanding or compressing the mobility network.
5. **Foundation-model university transformation (2021–):** A growing share of institutions transitioned from state to foundation governance. By 2023, over 70% of applicants targeted a foundation-model university.

6. **Institutional bonus points (2024–):** Institutions were allowed to award up to 20% of the total admission score, partly based on prior affiliation, potentially distorting the merit-based allocation mechanism.

Building on the identified gaps and the policy context, this paper makes the following contributions:

- $\boldsymbol{C_1}$ First application of graphon distance metrics to domestic higher education mobility networks, demonstrating that graphon-based structural analysis captures global shifts invisible to traditional metrics. *(Addresses RG_1; linked to RQ_3)*
- $\boldsymbol{C_2}$ Systematic, comparative evaluation of six graph distance metrics (spectral, cut, Frobenius, Wasserstein, Jensen–Shannon, total variation) for detecting temporal structural changes. *(Addresses RG_2; linked to RQ_2, RQ_3)*
- $\boldsymbol{C_3}$ Comprehensive longitudinal analysis of sub-national mobility networks at the micro-region level, disaggregated by preference rank and field of study, covering nearly two decades (2006–2024). *(Addresses RG_3; linked to RQ_1)*
- $\boldsymbol{C_4}$ An integrative methodological framework that links traditional network statistics to graphon-based distances, clarifying what each approach uniquely reveals about structural change. *(Addresses RG_4; linked to RQ_2, RQ_3)*

$\mathbf{C}_1$ is important because, without graphon representations, one cannot distinguish between two networks that share the same summary statistics but differ in their underlying connectivity pattern. By estimating graphons, we obtain a continuous, non-parametric representation of the mobility network that enables principled structural comparison.

$\mathbf{C}_2$ addresses the practical question of which distance metric to use. Different metrics emphasise different aspects of structure (global vs. local, spectral vs. combinatorial), and practitioners need guidance on which metric is most informative for a given research question. Our systematic comparison fills this gap.

$\mathbf{C}_3$ provides the most detailed longitudinal account of Hungarian higher education mobility to date, exploiting three complementary network constructions (total, preference-ranked, field-disaggregated) to capture multiple dimensions of structural change.

$\mathbf{C}_4$ bridges two largely separate literatures: classical network analysis and graphon theory. By computing both traditional metrics and graphon distances on the same data, we can identify where they agree, where they diverge, and what complementary insights each offers.

3 Materials and Methods

The methodological design of this study is driven by the research questions formulated in Section 1. We construct our datasets to enable the computation of both traditional network indicators and graphon-based structural distances, facilitating the integrative comparison central to RQ_2 and RQ_3.

3.1 Employed Data

Our primary data source is the Hungarian Higher Education Admission Database, covering the period 2006–2024. Each record contains (among other fields) the applicant's micro-region of residence, the micro-region of the chosen institution's campus, the applicant's preference rank (1st, 2nd, 3rd), and the field of study. We aggregate applications to the level of 175 sub-regions (LAU 1, formerly NUTS4), which provides a consistent geographic unit throughout the study period despite administrative boundary changes.

We construct three families of weighted, directed, dynamic multilayer networks:

(i) Total application network (by year).

For each year $t \in \{2006, \ldots, 2024\}$, a directed graph $G_t = (V, E_t, w_t)$ where V is the set of 175 micro-regions, and $w_t(i, j)$ is the number of applications from micro-region i to micro-region j in year t.

(ii) Preference-ranked networks.

For each year t and preference rank $r \in \{1, 2, 3\}$, a directed graph $G_{t,r}$ with edges weighted by the number of r-th choice applications. We restrict attention to the first three preference ranks because (a) applicants could submit up to three applications free of charge, making 4th and 5th choices exceedingly rare, and (b) the majority of admitted students are accepted at their first-choice institution (Kosztyán et al, 2020; Telcs et al, 2016).

(iii) Field-of-study networks.

For each year t and field $f \in \{1, \ldots, 13\}$, a directed graph $G_{t,f}$ with edges weighted by the total number of applications in field f.

The edge list for each network follows the structure:

`Year | From | To | Weight`

example of a completed record

`2010 | 42 | 117 | 285`

Node attributes are recorded as:

`Node_ID | Node_Name | Latitude | Longitude`

example of a completed record

`24 | Budapest | 47.4979 | 18.9989`

3.2 Derived databases

From the raw graph data, we construct four derived databases:

Database 1: Multi-layer graph collection.

Each "layer'' corresponds to one network instance. Three multi-layer structures are defined:

- **ML-Total**: 19 layers (one per year, 2006–2024), each containing the total application graph.

- **ML-Pref**: $19 \times 3 = 57$ layers (year $\times$ preference rank).
- **ML-Field**: $19 \times 13 = 247$ layers (year $\times$ field of study).

This enables not only year-over-year comparisons but also cross-sectional comparisons across preference ranks and fields within the same year.

Database 2: Network-level indicator time series.

For each network instance, we compute a vector of network-level indicators (Section 3.3). The simplest table is two-dimensional (rows: years; columns: indicators) for the total application network. The preference-ranked and field-disaggregated variants yield three-dimensional tables, with the third dimension being preference rank or field, respectively.

Database 3: Node-level indicator time series.

For each network instance and each micro-region, we compute a vector of node-level indicators (Section 3.4). The total-application table is three-dimensional (year $\times$ micro-region $\times$ indicator). The preference-ranked and field-disaggregated variants are four-dimensional, with the fourth dimension being preference rank or field.

Database 4: Graphon representations and distance matrices.

For each network instance, we estimate a graphon and compute pairwise distances using six metrics (Section 3.6). The resulting distance matrices enable multidimensional scaling, clustering, and time-series analysis of structural proximity.

3.3 Network-Level Indicators

We employ a broad suite of network-level indicators to characterize the structural properties of each mobility network. These metrics are grouped by the aspect of structure they capture, as summarized in Table 3. This constitutes a *classical network analysis* that is indispensable both in its own right and as a benchmark for the graphon-based analysis (RQ2, RQ3). Detailed definitions and formulas for all metrics are provided in Appendix A.2.

Below, we briefly discuss each group's significance for structural analysis. Note that low/high values and the direction of change carry specific interpretive meaning.

Degree & Strength centralization.

These metrics quantify the inequality of the degree and strength distributions. A centralization value near 1 indicates a star-like topology dominated by a single hub; values near 0 indicate uniformity. In the mobility context, increasing strength centralization (Instrength Centralization (SZI), Outstrength Centralization (SZO), Strength Centralization (SZA)) signals that application flows are becoming more concentrated around a few destination (or origin) regions. Decreasing values suggest a more balanced distribution. Disaggregating into in- and out-components reveals whether the concentration is driven by receiving regions (institutional attractors) or sending regions (population centres).

Table 3: Network-Level Indicator Groups and the Structural Changes They Detect in Temporal Analysis.

Metric group	**Interpretation**	**Structural change detected in time-series analysis**
Degree & Strength Centralisation (DZI, DZO, DZ, SZI, SZO, SZA)	Measures how unevenly connections (degree) or application flow volumes (strength) are distributed across micro-regions. Values near 1 indicate a hub-dominated network.	Increasing centralisation signals growing dominance of a few destination (or source) micro-regions; decreasing values indicate a more balanced distribution of application flows.
Path & Distance-based Centralisation (BZ, CZ, HZ)	Captures the extent to which shortest-path-based importance is concentrated in specific micro-regions.	Rising betweenness centralisation (BZ) indicates that certain micro-regions become critical intermediaries in application flow routes; changes in closeness (CZ) and harmonic (HZ) centralisation reflect shifts in overall reachability across the mobility network.
Prestige Centralisation (EZ, PRZ, AZ, PZ)	Reflects the concentration of recursive, prestige-based importance measures—where a micro-region's importance depends on the importance of the regions connected to it.	Detects shifts in the prestige hierarchy: whether the importance distribution among micro-regions becomes more concentrated (a few dominant hubs) or more balanced.
Community & Modularity (LeM, IM, SGM)	Quantifies the strength of community-like organisation in the mobility network.	Increasing modularity signals regionalisation (self-contained clusters of micro-regions with strong internal application flows); decreasing modularity suggests integration or blurring of regional boundaries.
Connectivity & Efficiency (Arcs, Dens, Diam, AVPL, ALE, GLE)	Describes the basic topology of the mobility network: how many application routes exist, how dense the network is, and how efficiently students can reach institutions.	Changes reflect whether the network is becoming sparser or denser, more or less efficient. A shrinking diameter with increasing density suggests the network is becoming more compact; growing average path length with decreasing efficiency suggests fragmentation.
Edge & Higher-order Structure (BEZ, Mot, Tra)	Captures meso-scale patterns: the importance of specific application routes (BEZ), triangular clustering among micro-regions (Tra), and recurring local connectivity patterns (Mot).	Motif count changes signal shifts in local connectivity patterns (e.g., triangular relationships among micro-regions); transitivity changes indicate evolving clustering behaviour in application flows.
Resilience & Robustness (RRes, SRes, NRI)	Measures the mobility network's ability to maintain connectivity if micro-regions (and their institutions) are removed.	Declining resilience suggests increased vulnerability (e.g., overdependence on a few hub regions whose closure would fragment the network); increasing resilience indicates a more robust structure with redundant pathways.
Asymmetry & Hierarchy (VAs, RCC, Assort)	Vertex asymmetry quantifies the average directional imbalance between incoming and outgoing application flows. The rich-club coefficient measures whether high-degree micro-regions preferentially connect to each other.	Growing asymmetry signals that micro-regions become increasingly specialised as either senders or receivers. Rising rich-club coefficient indicates strengthening connections among major university micro-regions (an emerging "elite club"). Assortativity reveals whether institutions primarily attract applicants from micro-regions that also host institutions (high value) or from those without institutions (low value).
Complexity, Fractal & Entropy (FDIMALL, FDIMGC, KIGC, NE, PCTGC, NLAC)	Self-similarity (fractal dimension), structural complexity (Kirchhoff index), information content of the degree distribution (entropy), connectivity threshold (percolation), and gap heterogeneity (lacunarity).	Changes reflect whether the mobility network is becoming more or less self-similar, more or less complex, and whether critical thresholds for connectivity are shifting—signalling deeper structural transformations.

Path & Distance-based centralization.

Betweenness Centralization (BZ) measures whether a few nodes serve as critical bridges. In a mobility network, high BZ means that certain micro-regions are obligatory "waypoints' ' in the flow architecture. Closeness Centralization (CZ) and harmonic Harmonic Centralization (HZ) centralization capture how concentrated reachability is. Temporal increases in BZ may signal emerging bottleneck regions; decreases may indicate redundancy or network decentralization.

Prestige Centralisation.

The eigenvector (Eigenvector Centralization (EZ)), PageRank (PageRank Centralization (PRZ)), alpha (Alpha Centralization (AZ)), and power centralisation (Power Centralization (PZ)) measure the concentration of recursive, prestige-based importance among micro-regions. In a mobility context, increasing values signal that the prestige hierarchy is becoming more concentrated—a few destination micro-regions with prestigious institutions accumulate an ever-larger share of the network's recursive importance. Decreasing values indicate a more balanced prestige distribution across micro-regions.

Community & Modularity.

Leiden's Graph Modularity Value (LeM), Infomap's Modularity Value (IM), and Spin-Glass Modularity Value (SGM) measure the strength of community structure using different algorithmic foundations. High modularity indicates that the network decomposes into relatively self-contained clusters (e.g., regional education markets). Temporal increases signal *regionalisation*; decreases signal *integration*.

Connectivity & Efficiency.

The number of Number of Arcs (Arcs), Density (Dens), Diameter (Diam), Average Path Length (AVPL), and efficiency metrics (Average Local Efficiency (ALE), Global Efficiency (GLE)) describe the network's basic connective fabric. A shrinking diameter with increasing density suggests the network is becoming more compact; increasing average path length with decreasing global efficiency suggests fragmentation.

Edge & Higher-order Structure.

Edge Betweenness Centralization (BEZ), Motifs (Mot) counts, and Transitivity (Tra) probe meso-scale structure. Changes in motif profiles indicate shifts in local connectivity patterns (e.g., more reciprocal or more hierarchical triads); transitivity changes reflect evolving clustering behaviour.

Resilience & Robustness.

Resilience for Random Attacks (RRes), Resilience for Systematic Attacks (SRes) and the Normalized Resilience Index (NRI) quantify how well the network maintains connectivity under perturbation. Declining resilience over time suggests that the network is becoming structurally more fragile.

Asymmetry & Hierarchy.

Mean Value of Vertex Asymmetries (VAs) quantifies the average directional imbalance at the node level. The Rich Club Coefficient (RCC) tests whether high-degree nodes are disproportionately connected to each other. Rising RCC over time suggests an increasingly "elite' ' club of strongly interconnected hub regions. Assortativity (Assort) shows whether high-degree vertices are connected to high-degree nodes – or, in the case of a disassortative network, low-degree nodes. Usually, microregions with high (in)degree vertices are those where higher education institutions are located, so if this value is high, it shows that there is high mobility even between microregions that have higher education institutions. A low value indicates that universities are able to attract students primarily from microregions where there are no higher education institutions.

Complexity, Fractal Dimension & Entropy.

Network Box Counting Fractal Dimension for the Entire Graph (FDIMALL), FDIMGC Network Box Counting Fractal Dimension for the Giant Component (FDIMGC) quantifies self-similarity; the Kirchoff Index for the Giant Component (KIGC) captures the sum of effective resistances; Network Entropy (NE) measures the information content of the degree or weight distribution; the Percolation Treshold for the Giant Component (PCTGC) indicates the critical point at which the giant component disintegrates; and Networks' Lacunarity (NLAC) captures the heterogeneity of gaps in the structure. Changes in these metrics signal deep structural transformations that affect the network's complexity and scale-free or self-similar character.

3.4 Node-Level Indicators

In parallel with network-level indicators, we compute an extensive set of node-level indicators for each micro-region in each network instance. These enable us to track how *individual micro-regions' roles* change over time (RQ1) and to construct feature vectors for node-level clustering and comparison. Table 4 summarizes the metric groups. Detailed definitions and formulas appear in Appendix A.1.

Degree & Strength Centrality (SCI, SCO, SCA, DCI, DCO, DCA).

This group captures the most elementary aspect of a node's position: the volume and breadth of its direct connections. Instrength Centrality (SCI) measures the total number of applications *received* by a micro-region. A low SCI indicates a region that attracts few applicants—typically one without a major higher education institution or hosting only small, specialised programmes. A high SCI identifies major destination hubs: micro-regions containing large universities with broad programme portfolios (e.g., Budapest districts, Debrecen, Szeged, Pécs). Outstrength Centrality (SCO) measures the total number of applications *sent* from a micro-region. Low SCO characterizes sparsely populated or demographically declining regions whose residents generate few applications. High SCO identifies populous "feeder' ' regions whose residents apply in large numbers, often to multiple destinations. Strength Centraliy (SCA)

Table 4: Node-Level Indicator Groups and the Role Changes They Detect in Temporal Analysis.

Metric group	**Interpretation**	**Node role change detected in time-series analysis**
Degree & Strength Centrality (SCI, SCO, SCA, DCI, DCO, DCA)	Measures a micro-region's direct connectivity (degree: number of partner regions) and application flow volume (strength: total applications sent or received).	A micro-region gaining in-strength is attracting more applications; gaining out-strength means its residents are applying more widely. Comparing strength and degree reveals whether flows are concentrated on a few partners or dispersed across many.
Path-based Centrality (BC, CC, HC, GEOK, LINC)	Captures a micro-region's importance based on shortest paths, reachability, and geodesic accessibility within the application flow network.	Increasing betweenness signals an emerging bridge role (the micro-region becomes a critical intermediary between otherwise weakly connected regions); increasing closeness/harmonic centrality signals improved structural accessibility.
Prestige Centrality (EC, AUT, HBS, PRC, AC, PC)	Recursive importance: a micro-region is important if connected to other important regions. Authority vs. hub scores distinguish receiving (destination) from sending (source) roles.	Rising authority signals a growing attractor role (the micro-region's institutions are increasingly targeted by applicants from important source regions); rising hub score signals that the micro-region's residents increasingly apply to the most sought-after destinations.
Neighbourhood & Local Structure (KNN, SCR, CLR, DMNC, LCC, SLC, LE)	Characterises the local environment of a micro-region: the connectivity of its partner regions, the clustering among its neighbours, and local efficiency.	Changes signal shifts in a micro-region's local embedding within the mobility network—whether it operates within a dense, cohesive regional cluster or at the periphery with sparse local connections.
Influence & Leadership (LC, LR, VR, CI, LOB)	Identifies micro-regions with disproportionate influence on system-wide dynamics (e.g., trend propagation, maintaining network connectivity).	A micro-region gaining collective influence is becoming a key driver of the overall mobility system: its application patterns or institutional changes ripple through the network.
Positional & Structural Role (ECC, COR, BURT, CRIT, MC)	Captures structural position: eccentricity (peripherality), coreness (embedding depth), Burt's constraint (access to structural holes), criticality (impact of removal), Markov centrality (random-walk accessibility).	A micro-region moving to a lower core or increasing eccentricity is becoming more peripheral. Decreasing Burt's constraint indicates the micro-region is gaining a brokerage position—bridging otherwise disconnected parts of the mobility network.
Fractal, Resilience & Spectral (node) (LAC, RC, LPC)	Node-level contributions to the network's fractal structure, resilience (connectivity upon removal), and Laplacian energy (structural complexity).	Changes indicate whether a micro-region's removal would significantly affect the mobility network's overall resilience or structural complexity.

combines both directions; high values identify regions that are simultaneously significant senders and receivers (e.g., Budapest districts that both host universities and generate many applicants).

The degree variants—in-degree (Indegree Centrality (DCI)), out-degree (Outdegree Centrality (DCO)), and total degree (DCO)—count the *number* of distinct partner regions rather than flow volumes. A region may have high strength but low degree (concentrated flows to/from a few partners) or moderate strength but high degree (dispersed flows across many partners). Low DCI signals a narrow geographic catchment (the institution draws from only a few nearby regions); high DCI signals a broad, potentially national catchment. Low DCO indicates that residents target only a few nearby destinations (limited geographic mobility); high DCO indicates geographically diversified application behaviour. Comparing strength and degree variants reveals whether a region's flows are *concentrated* (high strength, low degree) or *dispersed* (moderate strength, high degree)—a distinction that is critical for understanding the equity and resilience implications of mobility patterns.

Path-based Centrality (BC, CC, HC, GEOK, LINC).

While degree and strength capture direct connections, path-based metrics reveal a node's importance in the *indirect* flow architecture. Betweenness Centrality (BC) counts the fraction of all shortest paths between other node pairs that pass through a given region. A low BC means the region is not located on critical routes—removing it would not disrupt the flow of applications between other parts of the country. A high BC identifies structural *bridges* or *gateways*: regions that mediate between otherwise weakly connected parts of the network. In the Hungarian context, county seats or regional centres that connect rural peripheries to major university cities often exhibit high BC. Temporal increases in a region's BC signal that it is becoming a more critical intermediary—potentially because alternative routes have disappeared or because its institution has gained a unique bridging role.

Closeness Centrality (CC) measures the inverse of the average shortest-path distance to all other nodes. Low CC characterizes regions that are structurally "far' ' from the rest of the network—typically peripheral areas with limited direct or indirect connections to major hubs. High CC identifies regions that can reach the entire system in few steps, reflecting excellent structural accessibility. Harmonic Centrality (HC) is a variant that handles disconnected components gracefully: while CC becomes undefined when some nodes are unreachable, HC assigns a contribution of zero for infinite distances, making it more robust. A region with high HC but moderate CC may be well-connected within its component but separated from a small isolated subnetwork.

Geodesic K-path centrality (**GEOK**) restricts the reachability measure to paths of length at most K, thereby capturing *local–regional* rather than national accessibility. Low GEOK indicates limited reach within K hops; high GEOK signals strong connectivity within the immediate regional neighbourhood. Lin Centrality (LINC) balances the size of the reachable set against the average distance to reachable nodes, rewarding regions that combine *many* reachable destinations with *short* distances. Low LINC indicates either a small reachable set or large distances; high LINC indicates efficient, broad access.

Prestige Centrality (EC, AUT, HBS, PRC, AC, PC).

This group captures *recursive* or *self-reinforcing* importance: a node is central not merely because it has many connections, but because it is connected to *other central nodes*. Eigenvector Centrality (EC) assigns each node a score proportional to the sum of its neighbours' scores, corresponding to the leading eigenvector of the adjacency matrix. Low EC characterizes regions connected primarily to peripheral, low-importance areas. High EC identifies regions embedded in the "prestige core'' of the network—connected to other well-connected regions in a mutually reinforcing structure.

The HITS framework decomposes recursive importance into two complementary roles. Authority Centrality (AUT) is high for regions that *receive* applications from many good hubs; Hubness Centrality (HBS) is high for regions whose residents *send* applications to many good authorities. In the mobility context, high-authority regions host the most sought-after institutions—those that the most active, geographically mobile applicant populations target. High-hub regions are "sophisticated senders'' whose residents consistently choose prestigious, high-authority destinations. A region with high AUT but low HBS is a pure attractor; a region with high HBS but low AUT is a pure feeder of talent toward top destinations. Temporal shifts in the AUT–HBS profile reveal how the sending–receiving balance of a region evolves.

PageRank Centrality (PRC) models a damped random walk on the network: a hypothetical "random applicant'' who, at each step, either follows an existing application link (with probability α) or jumps to a random region (with probability $1-\alpha$). High PRC identifies natural accumulation points—regions where the random walker spends the most time, reflecting both direct attractiveness and favourable structural position. Low PRC characterizes regions that the random walk rarely visits, either because few edges point to them or because their predecessors are themselves rarely visited.

Alpha Centrality (AC) generalises eigenvector centrality by introducing an exogenous input vector, allowing external factors (e.g., population size, institutional capacity) to modulate structural importance. Power Centrality (PC) further allows the sign of the neighbour-influence parameter β to vary: positive β means being connected to high-centrality neighbours *increases* one's centrality (cooperative regime), while negative β means it *decreases* it (competitive regime). In the competitive interpretation, a region surrounded by powerful competitors may itself have low power centrality, reflecting a "crowded market'' effect.

Neighbourhood & Local Structure (KNN, SCR, CLR, DMNC, LCC, SLC, LE).

This group characterizes the *local environment* in which a node operates, probing the structure of its immediate neighbourhood rather than its global position. Average nearest-neighbour degree (K-th Nearest Neighbourhood (KNN)) measures whether a region's partners are themselves well-connected. Low KNN indicates that the region's direct partners are peripheral (low-degree) nodes—the region is embedded in a sparse, peripheral zone. High KNN indicates that the region connects to well-connected hubs. The combination of a node's own degree and its KNN reveals the local mixing pattern:

high KNN with low own degree suggests *disassortative* mixing (a peripheral region attached to a hub), while high KNN with high own degree suggests an *assortative* elite cluster.

Weighted s-core value (S-core Value (SCR)) identifies the deepest core to which a node belongs in the nested k-core decomposition. Low SCR places the region in the outer shells—it would be among the first to be "peeled off'' in a progressive pruning process. High SCR places it in the innermost, densest core of the network. In mobility terms, the inner core typically comprises the largest university cities and their primary feeder regions, forming the backbone of the national mobility system.

Cluster Rank Value (CLR) combines local clustering with out-degree, capturing both the cohesiveness of a node's neighbourhood and its outward reach. Low CLR indicates a region with few outgoing connections in a loosely clustered neighbourhood; high CLR indicates a region with many outgoing connections embedded in a tightly knit local cluster. Density of Maximum Neighborhood Component (DMNC) measures how densely connected a node's neighbours are among themselves. Low DMNC signals a fragmented neighbourhood where the node's partners rarely interact with each other; high DMNC signals a cohesive local cluster—a group of micro-regions that exchange students among themselves.

Local Community Centrality (LCC) counts the number of edges among a node's neighbours, directly measuring the density of the local community. Semi Local Centrality (SLC) extends the view to two hops, summing the degrees of nodes at distance two. Low SLC indicates limited extended influence; high SLC indicates a large "influence footprint'' that extends beyond the immediate neighbourhood. Local Efficiency (LE) measures how well a node's neighbours can communicate with each other *if the node itself were removed*. High LE indicates a redundant local structure where the node's removal would cause little disruption; low LE indicates that the node is critical for local connectivity—its removal would fragment the neighbourhood.

Influence & Leadership (LC, LR, VR, CI, LOB).

This group identifies nodes with disproportionate influence on network-wide dynamic processes such as information spreading, collective action, or percolation. Leverage Centrality (LC) compares a node's degree to those of its neighbours. A region with positive (high) LC has significantly more connections than its neighbours on average, giving it structural *leverage*: it can influence many partners who have few alternative connections. Negative (low) LC indicates a region whose neighbours are more connected than itself—it is structurally dependent rather than influential.

Leader Rank in a Directed Graph (LR) is a PageRank variant that adds a bidirectional "ground node" connected to all regions, eliminating the need for a damping factor and improving convergence. Low LR identifies regions with minimal leadership potential; high LR identifies natural leaders in the flow hierarchy. Vote Rank (VR) uses an iterative voting process: in each round, neighbours vote for a node, the top-voted node is selected as a key spreader and removed, and the process repeats. Regions selected *early* (high VR) are the most effective independent spreaders; those

never selected (low VR) contribute little to diffusion dynamics. In the mobility context, high-VR regions are those whose application patterns most effectively "reach'' the entire network.

Collective Influence (CI) identifies nodes whose removal, together with their boundary at distance ℓ, would cause the largest drop in network connectivity (optimal percolation). Low CI characterizes structurally redundant regions; high CI characterizes regions that are critical for maintaining the network's percolation backbone. The Lobby Index (LOB) is analogous to the academic h-index: a region has lobby index l if it has at least l neighbours each with degree at least l. Low LOB indicates a poorly connected local elite; high LOB indicates a node surrounded by a strong, well-connected neighbourhood.

Positional & Structural Role (ECC, COR, BURT, CRIT, MC).

This group locates each node within the network's global architecture, identifying its structural role (core vs. periphery, broker vs. constrained, critical vs. redundant). Eccentricity Value (ECC) measures the maximum shortest-path distance from a node to any other node. Importantly, the interpretation is *reversed* relative to most centrality measures: *low* eccentricity places a region in the graph's centre (close to all other nodes), while *high* eccentricity identifies peripheral regions far from at least one other part of the network. Temporal increases in a region's eccentricity signal growing peripheralisation.

Coreness (COR) assigns each node the index of the deepest k-core to which it belongs. Low coreness places a region in the outer shells of the network—positions that are topologically peripheral and vulnerable to pruning. High coreness identifies regions in the innermost core, the dense nucleus of the mobility system. In the Hungarian context, the highest-coreness regions are typically the major university cities and their adjacent micro-regions.

Burt's Constraint (BURT) measures the extent to which a node's contacts are redundantly connected to each other. Again, the interpretation is *inverted*: *low* constraint is desirable, as it indicates access to *structural holes*—the node bridges between groups that would otherwise be disconnected, enjoying brokerage advantages. *High* constraint indicates that the node is locked into a single, dense cluster with no bridging opportunities. In the mobility context, a region with low Burt's constraint connects applicant populations from different geographic or disciplinary segments, potentially serving as a gateway between otherwise separate parts of the system. A region with high constraint is embedded in a single regional cluster with limited inter-regional brokerage capacity.

Criticality Value (CRIT) measures the drop in global efficiency caused by removing a node. Low criticality characterizes structurally redundant regions whose removal barely affects overall network performance. High criticality identifies *irreplaceable* regions: removing them would significantly degrade the system's ability to efficiently route flows. Markov Centrality (MC) measures the inverse of the mean first-passage time to a node in a random walk. High MC means the region is a natural "attractor'' in the random-walk sense: starting from any other region, a random walker reaches it

quickly. Low MC characterizes regions that the random walker visits only after many steps, reflecting poor structural accessibility.

Fractal, Resilience & Spectral—node level (LAC, RC, LPC).

This group captures each node's contribution to the network's overall complexity, resilience, and spectral energy. Local Lacunarity (LAC) measures the heterogeneity of gaps in a node's neighbourhood at a given scale. Low LAC indicates a homogeneous, uniform local structure—the connections around the node are evenly distributed. High LAC indicates an irregular, "patchy' ' local connectivity with heterogeneous gaps, suggesting that the node's neighbourhood contains both densely and sparsely connected zones.

Resilience Centrality (RC) quantifies the fractional decrease in the size of the giant component caused by removing a given node. Low RC indicates that the node's removal would have little effect on the network's overall connectivity—the giant component remains largely intact. High RC identifies structurally critical nodes whose removal would cause a significant loss of connectivity. In the mobility context, high-RC regions are those on which the entire system's connectedness depends; their closure or decline would fragment the mobility network into disconnected components.

Laplacian Centrality (LPC) measures the fractional change in the Laplacian energy of the network upon node removal. The Laplacian energy, defined as the sum of squared Laplacian eigenvalues, captures the overall "vibrational energy' ' of the network structure. Low LPC characterizes structurally redundant nodes whose removal barely alters the eigenvalue structure. High LPC identifies nodes that are essential to the network's spectral profile—their removal would fundamentally alter the structural modes of the system. Together, RC and LPC provide complementary views: RC focuses on connectivity (giant-component integrity), while LPC focuses on the spectral complexity of the overall structure.

Because the six graphon distance metrics operate on different scales, we apply min–max normalization across the time series for each metric d_m:

$$\hat{d}_m(t) = \frac{d_m(t) - \min_\tau d_m(\tau)}{\max_\tau d_m(\tau) - \min_\tau d_m(\tau)}, \tag{1}$$

where τ ranges over all observed time points. After normalization, 0 corresponds to the temporal minimum and 1 to the temporal maximum of each metric. This facilitates direct cross-metric comparison and helps identify years when multiple metrics simultaneously signal large structural change.

To identify years in which the network structure changed most relative to adjacent years, we detect local maxima (peaks) in the normalized distance time series. A normalized distance value $\hat{d}_m(t)$ is classified as a local peak if

$$\hat{d}_m(t) > \hat{d}_m(t-1) \quad \text{and} \quad \hat{d}_m(t) > \hat{d}_m(t+1). \tag{2}$$

This discrete local maximum criterion identifies time points at which a given metric indicates the greatest structural shift relative to neighboring years.

3.5 Standardized Growth Rate and Growth-Rate-Based Anomaly Detection

Let x_t denote the value of a network metric at time t (e.g., degree, betweenness centrality, or clustering coefficient). The simple growth rate is defined as:

$$g_t = \frac{x_t - x_{t-1}}{x_{t-1}} \tag{3}$$

This measure captures the relative change in the network characteristic between two consecutive time points. The growth rate is positive when the metric increases and negative when it decreases. Since the network metrics under study may take negative values, the logarithmic growth rate cannot be applied (the logarithm is undefined for negative numbers), making the simple growth rate the appropriate choice.

Different network metrics operate on different scales, so the growth rates are standardized to enable cross-metric comparability. Since the growth rate distributions did not exhibit significant skewness, the classical z-score standardization is applicable:

$$z_t = \frac{g_t - \bar{g}}{\sigma_g} \tag{4}$$

where $\bar{g}$ is the temporal mean and σ_g is the standard deviation of the growth rate series. As a result of standardization, the transformed values have zero mean and unit variance, enabling anomaly signals from different network metrics to be compared using a uniform threshold.

Anomaly detection performed on the standardized growth rate time series reveals **dynamic** irregularities in the network. Using a simple threshold-based approach, a time point t is flagged as anomalous if:

$$|z_t| > \tau \tag{5}$$

where τ is a predetermined threshold (typically $\tau = 2$ or $\tau = 3$).

As an alternative to the threshold-based approach, we used a model-based anomaly detection procedure (Chen and Liu, 1993). The function operates through the following steps:

1. *STL Decomposition.* The function first performs an STL decomposition (Seasonal and Trend decomposition using Loess) on the time series, using robust fitting:

$$x_t = T_t + S_t + R_t \tag{6}$$

   where T_t is the trend component, S_t is the seasonal component, and R_t is the remainder. If the time series is non-seasonal, a simple trend fit is applied instead. The robust fitting procedure ensures that outliers do not distort the decomposition itself.
2. *Examination of the Remainder.* Anomaly detection is performed exclusively on the remainder component R_t, from which the effects of trend and seasonality have already been removed.
3. *IQR-Based Threshold.* The interquartile range (IQR) rule is applied to the remainder. Let Q_1 and Q_3 denote the first and third quartiles, and let IQR =

$Q_3 - Q_1$. A time point t is classified as an outlier if:

$$R_t < Q_1 - 3 \cdot \text{IQR} \qquad \text{or} \qquad R_t > Q_3 + 3 \cdot \text{IQR} \tag{7}$$

This $3 \times$ IQR rule is a more conservative variant of the standard boxplot rule ($1.5 \times$ IQR). Under a normal distribution, it corresponds approximately to the $\pm 4.7\sigma$ range, meaning that only truly extreme values are flagged.

4. By default, the function performs two iterations: outliers identified in the first pass are replaced via linear interpolation, the decomposition is recomputed, and a second round of outlier detection is carried out. This is necessary because multiple adjacent outliers can mask each other (the masking effect): when several extreme points are present simultaneously, they inflate the IQR, causing none of them to be detected.

Growth-rate-based anomaly detection is capable of uncovering the following phenomena:

1. *Sudden structural reorganization:* If a node's betweenness centrality increases dramatically in a single time step, it suggests the node has suddenly assumed a key role in information flow. The absolute level may still be within a normal range, but the *magnitude* of change is abnormal.
2. *Cascade-like network events:* A rapid increase (or decrease) in a node's degree may indicate node failure, an attack, or the emergence of a new node attracting many connections.
3. *Sharp shocks in the absence of slow drift:* The growth rate is insensitive to slowly unfolding trends but highly sensitive to abrupt jumps from one time point to the next. This is advantageous when the focus is specifically on short-term shocks rather than gradual structural evolution.
4. *Cross-metric comparability:* Thanks to standardization, a uniform threshold (τ) can be used to flag anomalies regardless of whether the underlying metric is degree, centrality, or clustering coefficient.

It is important to emphasize that this approach is **complementary** to anomaly detection on the original time series. While anomalies in the original series reveal "unusual states" (e.g., abnormal absolute levels), anomalies detected on the growth rate reveal "unusual changes" (Blázquez-García et al, 2021), which is sought in this study.

3.6 Graphon Theory and Distance Metrics

3.6.1 Graphon Estimation for Directed, Weighted Graphs

Given a directed, weighted graph $G = (V, E, w)$ on n nodes, we associate it with a step-function graphon $\hat{W}_G : [0,1]^2 \to \mathbb{R}_{\geq 0}$ as follows. Let $\pi : V \to \{1, \ldots, n\}$ be a node ordering (obtained, e.g., by degree sorting or spectral methods). Partition $[0,1]$ into n equal intervals $I_1, \ldots, I_n$. Define:

$$\hat{W}_G(x,y) = \frac{w(\pi^{-1}(i), \pi^{-1}(j))}{w_{\max}} \quad \text{for } (x,y) \in I_i \times I_j, \tag{8}$$

where $w_{\max} = \max_{(i,j)\in E} w(i,j)$ is a normalisation factor ensuring $\hat{W}_G \in [0,1]^2$. For directed graphs, we do *not* impose symmetry, so $\hat{W}_G(x,y) \neq \hat{W}_G(y,x)$ in general.

To obtain a smooth graphon estimate, we apply a stochastic block model approximation or universal singular value thresholding (Chatterjee, 2015; Airoldi et al, 2013).

Beyond measuring distances between graphons, the estimated graphon block structure itself provides a natural basis for community detection. In the step-function graphon $\hat{W}_G$, nodes are grouped into *blocks* (node types) such that nodes within the same block exhibit similar connectivity patterns. These blocks define the *node-level communities* of the mobility network. If the graphon value between two blocks is non-zero, the corresponding block pair identifies a *flow pathway group*—a systematic application flow route connecting two structurally distinct groups of micro-regions.

This approach differs fundamentally from modularity-based methods (e.g., Leiden, Infomap), which partition the network into densely connected subgroups. Graphon-based community detection instead groups nodes by *structural equivalence*: micro-regions that play similar roles in the mobility network are assigned to the same block, regardless of whether they are directly connected. The inter-block flow pathways then reveal the hierarchical organisation of the mobility system. The detailed algorithmic steps are described in Appendix A.4.

3.6.2 Graphon distance metrics

Given two graphon estimates W_1 and W_2, we compute six distance metrics. Table 5 provides an overview; full mathematical definitions appear in Appendix A.3.

4 Results

This section presents the results needed to address our research questions (RQ_1–RQ_3). We report and describe the findings without providing detailed interpretation; the implications, underlying causes, and connections to research questions are discussed in Section 5.

4.1 Descriptive Statistics

Figure 1 shows the total number of first-preference applications and their distribution across 13 fields of study from 2006 to 2024.

Table 5: Graphon Distance Metrics: Overview and Structural Interpretation

Metric	**What it measures**	**Structural interpretation**	**Analogous traditional metrics**
Cut distance ($\delta_{\square}$)	Maximum discrepancy in total application flow between any pair of micro-region subsets.	Captures global, large-scale rearrangements of mobility flows; insensitive to small local perturbations. A large cut distance indicates that the two networks differ in how they partition flows across macro-regions.	Modularity, degree centralisation
Spectral distance (δ_{spec})	Difference in eigenvalue spectra of the graphon operators.	Detects changes in the network's dominant structural modes (e.g., community structure, core–periphery hierarchy). Sensitive to reorganisation of the mobility network's overall architecture.	Eigenvector centralisation, community modularity
Frobenius distance (δ_F)	Root-mean-square difference in edge weights across all micro-region pairs.	Sensitive to both local and global weight changes; penalises large deviations strongly. Captures both macro-level restructuring and localised shifts in specific application routes.	Density, strength centralisation
Wasserstein distance (δ_W)	The magnitude of structural rearrangement needed to transform one year's mobility pattern into another's, accounting for both the extent and the spatial reach of changes.	Captures not only whether application flows changed, but also whether the changes occurred between nearby or distant parts of the network. Small values indicate minor local adjustments; large values indicate that the mobility pattern required extensive reconfiguration.	AVPL, efficiency, resilience
Jensen–Shannon (δ_{JS})	Divergence between the proportional distributions of application flows in two consecutive years.	Small JS distance indicates that the relative shares of application flows among micro-region pairs remained largely stable—even if absolute volumes changed. Large JS distance indicates that the distributional pattern of flows was fundamentally reshaped.	Network entropy, asymmetry
Total Variation (δ_{TV})	The largest local discrepancy in the proportional distribution of application flows between two consecutive years.	Small TV distance indicates that flow proportions remained stable across all micro-region pairs. Large TV distance signals at least one point in the network where flow proportions changed dramatically—particularly useful for detecting sudden, concentrated structural shifts (e.g., the rapid rise or decline of a specific institution or route).	Resilience, fractal dimension

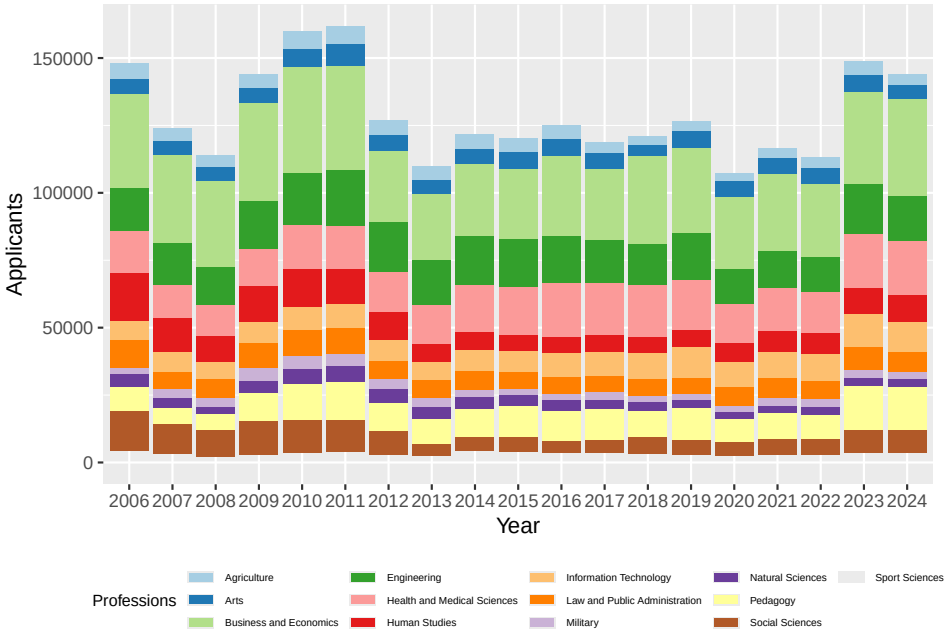


Fig. 1: First-Preference Applications: Total Volume and Distribution by Field of Study (2006–2024)

The overall volume peaked in 2011 at 161,731 and reached lows in 2008 (107,897) and 2024 (107,620). Two notable dips occurred in 2007–2008 and 2012–2013. Business and Economics remained the largest field throughout the period, while Information Technology approximately doubled its share (from 5% in 2006 to 8% in 2018). Human Studies and Social Sciences experienced sustained declines of approximately 58% and 57%, respectively.

4.1.1 Network-Level Indicators

Table 6 shows the 37 network-level indicators for 2006, 2013, and 2024, arranged according to the 9 indicator groups presented in Table 3. These three years represent the beginning, midpoint, and end of the study period.

Table 6: Calculated Network Indicators on 2006, 2013, 2024

	2006	**2013**	**2024**
DZI	0.8868	0.9098	0.8939
DZO	0.0186	0.0190	0.0181
DZ	0.4524	0.4642	0.4557
SZI	0.9847	0.9875	0.9877
SZO	0.9276	0.9023	0.9075
SZA	0.9747	0.9779	0.9784

Table 6: Calculated Network Indicators on 2006, 2013, 2024 *(continued)*

	2006	2013	2024
BZ	0.0012	0.0014	0.0013
CZ	0.8514	0.8847	0.8699
HZ	665.7695	695.4135	587.3003
EZ	0.6792	0.7163	0.6999
PRZ	46.5806	57.0458	52.0597
AZ	17,222.7010	3,844.3263	114,069.8749
PZ	92.2569	79.2871	152.9437
LeM	0.0176	0.0198	0.0227
IM	0.0000	0.0000	0.0000
SGM	0.0021	0.0022	0.0018
Arcs	3,175.0000	2,747.0000	2,606.0000
Dens	0.1132	0.0902	0.0999
Diam	268.0000	442.0000	339.0000
AVPL	22.7667	28.6016	22.4275
ALE	0.1132	0.1079	0.1321
GLE	0.0314	0.0264	0.0300
BEZ	2,459,900.4929	631,621.2000	1,180,080.1667
Mot	210,336.0000	186,963.0000	167,611.0000
Tra	0.2986	0.2526	0.2688
RRes	126.0000	131.5000	121.5000
SRes	38.9765	34.4432	34.4024
NRI	0.9613	1.0000	0.9254
VAs	0.8595	0.8784	0.8810
RCC	0.9152	0.9307	0.9316
Assort	-0.1638	-0.1861	-0.1495
FDIMALL	0.3645	0.8764	0.5573
FDIMGC	0.7025	0.8764	1.4727
KIGC	29.5281	38.5632	21,757,650,543,606.7031
NE	6.5989	6.5475	6.4996
PCTGC	0.1310	0.1360	0.1370
NLAC	1.2960	1.2760	1.3870

The most notable changes from the 2006 baseline include AZ increased more than 6-fold (from 17,223 to 114,070), indicating an extreme concentration of iterative prestige at a few dominant nodes. PZ increased by 65%. The number of arcs decreased by 18% (3,175 → 2,606), density fell by 12%. At the same time, the average path length decreased below the 2006 value in 2024 after an increase, suggesting fewer but more efficiently organised connections. BEZ 53%, and motif count fell by 21%, indicating

a flattening of local triangular clusters. LeM increased by 47%, signalling stronger community separation. The FDIMGC more than doubled (+110%), while the Kirchoff Index increased even more radically, indicating increasing structural complexity.

4.1.2 Node-Level Indicators for Micro-Regions Hosting the Most Students

Table 7 shows node-level indicators for Budapest and three major university micro-regions (Debrecen, Szeged, Pécs) in 2006 and 2024.

Table 7: Node-Level Indicators for Several Micro-Regions on 2006, 2024

	Budapest		Debrecen		Szeged		Pécs	
	2006	**2024**	**2006**	**2024**	**2006**	**2024**	**2006**	**2024**
SCI	114,490.0000	153,982.0000	29,338.0000	29,701.0000	25,459.0000	29,525.0000	28,547.0000	28,175.0000
SCO	24,208.0000	20,413.0000	4,463.0000	4,987.0000	3,058.0000	4,787.0000	2,627.0000	799.0000
SCA	138,698.0000	174,395.0000	33,801.0000	34,688.0000	28,517.0000	34,312.0000	31,174.0000	28,974.0000
DCI	167.0000	160.0000	167.0000	159.0000	166.0000	158.0000	167.0000	160.0000
DCO	21.0000	18.0000	21.0000	16.0000	21.0000	18.0000	21.0000	17.0000
DCA	188.0000	178.0000	188.0000	175.0000	187.0000	176.0000	188.0000	177.0000
BC	0.0000	0.0000	27.2619	6.6667	2.0000	0.3333	0.0000	136.1667
CC	0.0006	0.0022	0.0017	0.0019	0.0017	0.0019	0.0017	0.0022
HC	0.2804	0.7602	5.0208	2.2653	2.5692	2.2185	3.2793	4.4895
GEOK	0.0000	0.0000	151.0000	73.0000	40.0000	31.0000	62.0000	113.0000
LINC	3.0134	8.2551	57.7412	41.4073	38.7347	36.2531	40.5953	48.4505
EC	1.0000	1.0000	0.3423	0.2954	0.3404	0.3628	0.3392	0.2278
AUT	1.0000	1.0000	0.1849	0.1002	0.1727	0.1159	0.1850	0.1214
HBS	0.4590	0.1928	0.4013	0.4379	0.2889	0.4292	0.2809	0.0605
PRC	0.2815	0.3236	0.0633	0.0665	0.0792	0.0936	0.0830	0.0798
AC	-2.8958	0.1946	-15.0518	-9.0076	17.7365	22.1794	-28.2965	-49.9884
PC	-1.1230	-0.7268	-1.1230	-1.1768	-1.1230	-0.7268	-1.1230	-1.0883
LE	0.0317	0.0304	0.0311	0.0270	0.0315	0.0301	0.0314	0.0267
KNN	86.7196	74.0567	77.1771	75.2990	86.5972	77.9632	79.1356	60.6240
SCR	22.0000	19.0000	21.0000	16.0000	19.0000	18.0000	20.0000	17.0000
CLR	429.0095	278.5948	429.0095	289.3375	429.0095	278.5948	429.0095	297.4632
DMNC	0.2539	0.2292	0.3778	0.1058	0.4954	0.2644	0.5202	0.2203
SLC	960,751.0000	767,165.0000	960,751.0000	751,870.0000	959,081.0000	763,140.0000	960,751.0000	753,480.0000
LCC	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
LC	0.6538	0.6863	0.6538	0.6854	0.6512	0.6818	0.6538	0.6870
LR	7.6274	8.1299	7.6274	8.0756	7.6236	8.1174	7.6274	8.1076
LOB	22.0000	19.0000	22.0000	19.0000	22.0000	19.0000	22.0000	19.0000
VR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CI	0.0000	0.0000	564,179.0000	109,098.0000	21,390.0000	25,725.0000	43,571.0000	276,672.0000
ECC	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	1.0000	2.0000
MC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
COR	37.0000	30.0000	37.0000	30.0000	37.0000	30.0000	37.0000	30.0000
BURT	0.0582	0.0582	0.1937	0.2247	0.2163	0.2943	0.2138	0.2587
CRIT	0.0060	0.0062	0.0060	0.0062	0.0060	0.0062	0.0060	0.0062
LAC	1.0000	1.0000	2.5882	1.8438	1.4746	2.0388	1.0647	2.5221
RC	0.0237	0.0273	0.0237	0.0254	0.0234	0.0265	0.0237	0.0265
LPC	47,856.0000	41,930.0000	47,856.0000	40,680.0000	47,460.0000	41,170.0000	47,856.0000	41,404.0000

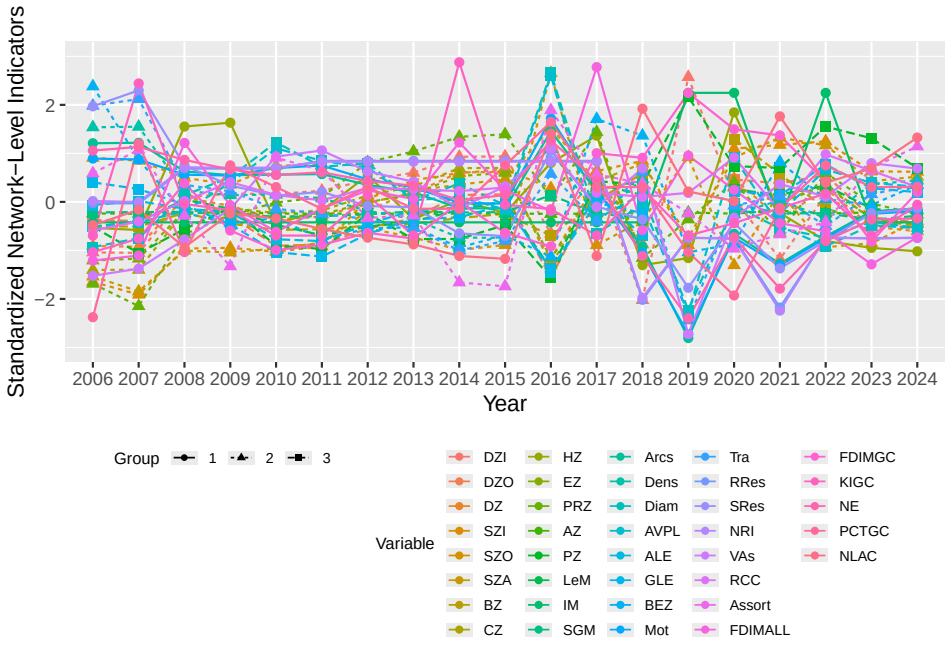
Budapest strengthened its in-strength (SCI: 34%) but lost 16% of its out-strength (SCO), while closeness centrality increased by 359% and alpha centrality. Pécs experienced the most dramatic decline: out-strength fell by 70% (SCO: 2,627 → 799).

Szeged was the only micro-region where out-strength increased dramatically (+56%), with hub score rising by 48%. Debrecen's out-strength increased by 12%, and its collective influence decreased by 81%. Core number decreased uniformly across all four micro-regions (37 → 30).

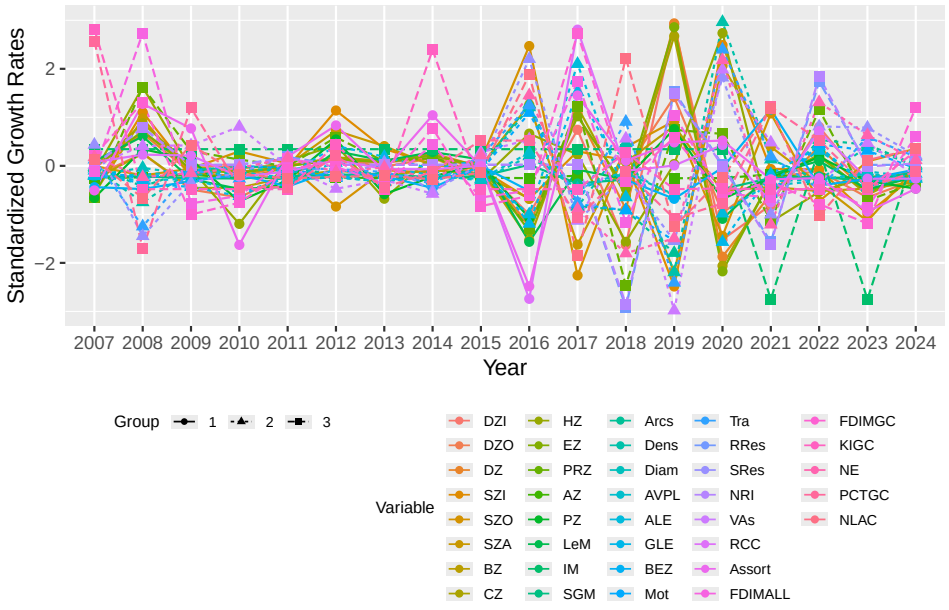
4.2 Anomaly detection for the Time-Series of Growth Rates and Differences

4.2.1 Analysis of Network-Level Indicators

Figure 2 shows the standardized network-level indicator values (2a) and their standardized growth rates (2b). Values beyond the ±3 range indicate statistically unusual observations. The indicators are arranged by logical group; in addition, correlation analysis identified three correlation clusters for standardized values and two for growth rates, distinguished by line type and marker shape.



(a) Standardized Network-Level Indicators (2006-2024)

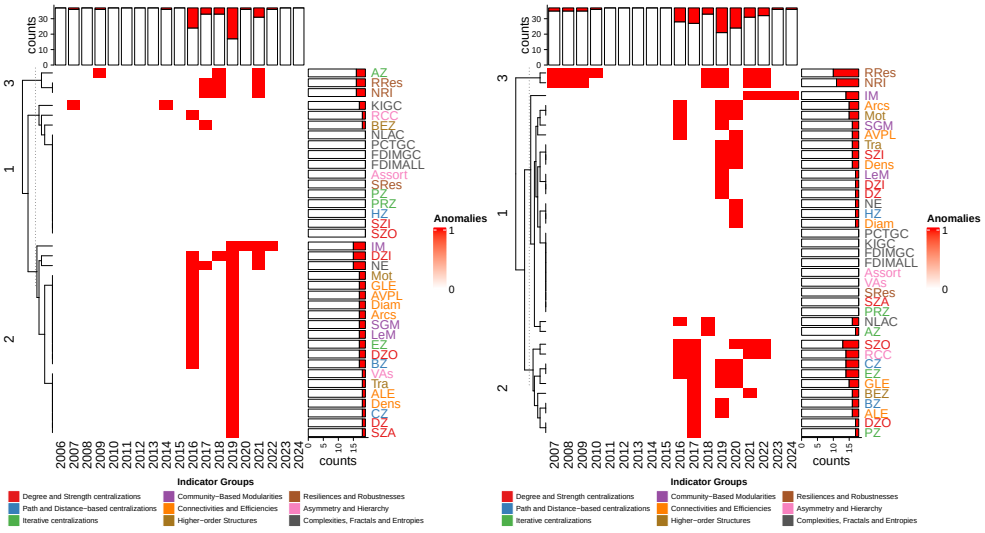


(b) Standardized Growth Ratios (2007-2024)

Fig. 2: Standardized Value of Network-Level Indicators and Growth Ratios

Several indicators exceed the ±3 range in multiple years. Indegree Centralization (DZI), NE, IM are the most volatile in standardized values (anomalous in 2009, 2020, 2022–2024). For growth rates, the SRes is most frequently anomalous (8 years), followed by NRI (7), RRes (6), RCC (4), and IM (4 years each).

Figure 3 shows the temporal pattern of anomalies for standardized values (3a) and growth rates (3b). Red cells indicate anomalies detected by the model-based procedure described in Section 3.5. Indicators are clustered by value similarity, which does not necessarily coincide with the logical grouping.



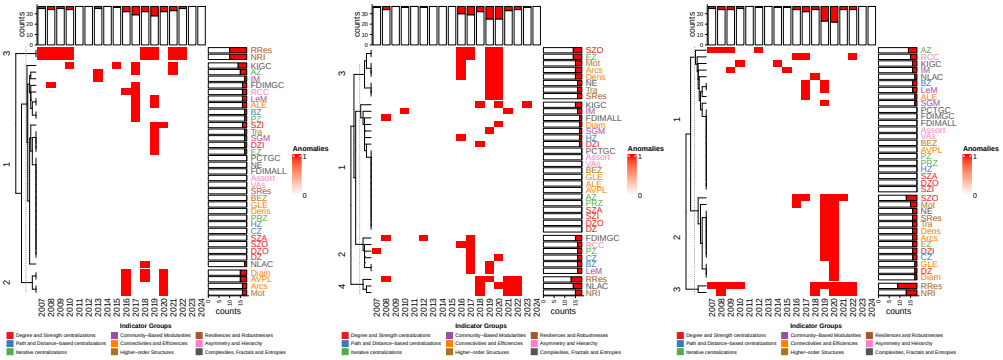
(a) Network-Level Indicators (b) Growth Ratios

Fig. 3: Anomalies of the Standardized Network-Level Indicators (2006-2024) and Growth Ratios (2007–2024)

The years with the most anomalies (growth rates) are 2019 (16 indicators), 2020 (13), 2017 (10), and 2016 (9). The most stable indicators (zero anomalies) include SZA, PRZ, FDIMALL, FDIMGC, and PCTGC. The most volatile are resilience indicators, such as RRes and NRI (anomalous in 2007-2011, 2018-2019, 2021-2022).

Comparison of Application Orders

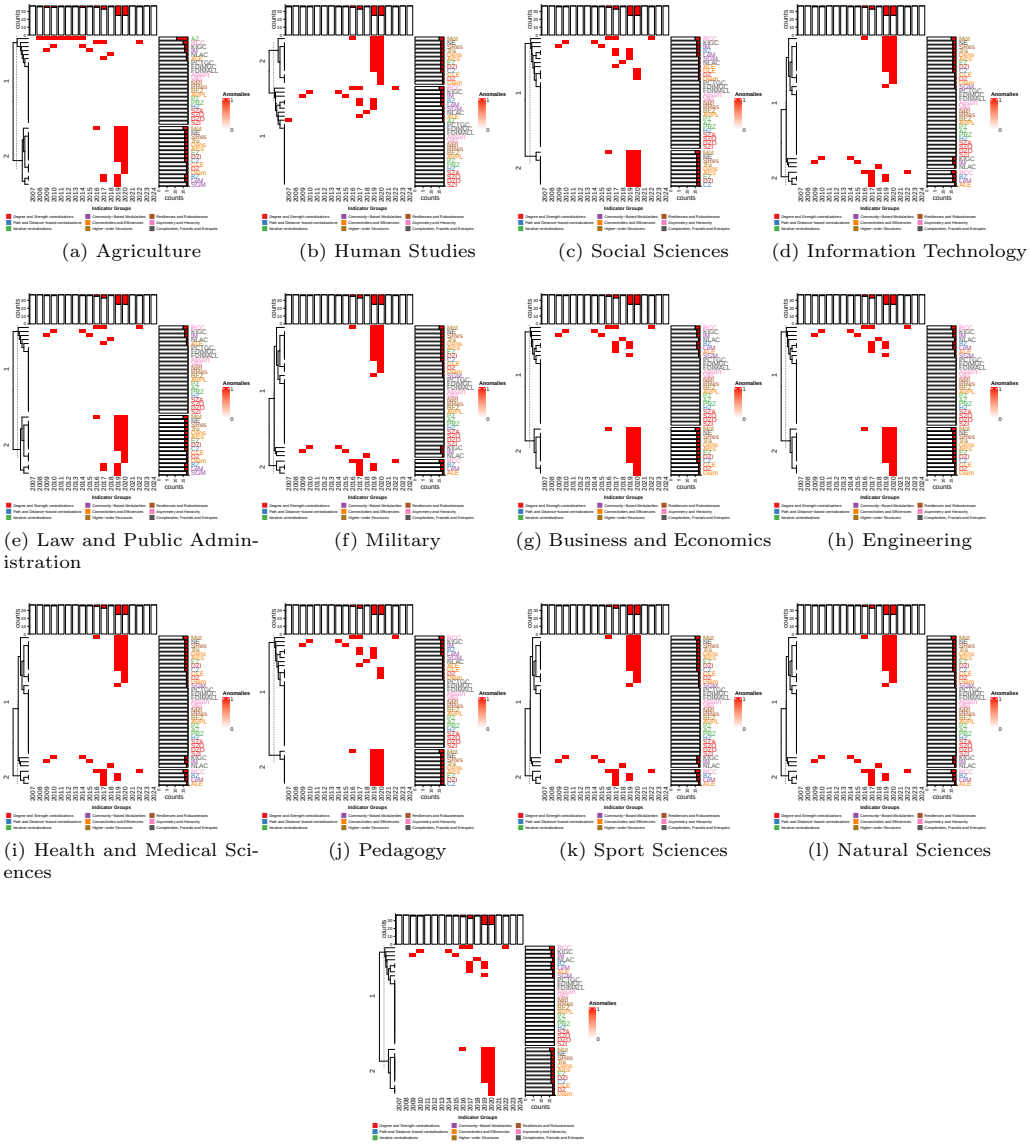
Figure 4 shows the anomalies of growth rates by application preference: first (4a), second (4b), and third (4c) preference. First-preference networks display different anomalies than second- or third-preference networks.



(a) 1st place (b) 2nd place (c) 3rd place

Fig. 4: Anomalies of the Growth Ratios (2007–2024) by Application Preferences

Figure 5 shows anomalies of growth ratios by field of study. The panels correspond to the 13 fields listed in Table 1. Indicator names are colour-coded by logical group.



(a) Agriculture (b) Human Studies (c) Social Sciences (d) Information Technology

(e) Law and Public Administration (f) Military (g) Business and Economics (h) Engineering

(i) Health and Medical Sciences (j) Pedagogy (k) Sport Sciences (l) Natural Sciences

(m) Arts

Fig. 5: Anomalies of Network-Level Growth Rates by Field of Study (2007–2024)

4.2.2 Graphon Differences

Figure 6 shows the six normalized (0–1) graphon distance metrics from 2007 to 2024. The normalisation procedure enables direct comparison across metrics with different native scales.

Three periods stand out: (i) **2019:** where all distance metrics has a local maxima. Cut (1.0) and Frobenius (1.0) reach their maxima, with Wasserstein near maximum (0.989), indicating the largest global restructuring of application flows in the entire study period; (ii) **2009–2015:** JS and TV remain at or near 1.0, while Cut and Frobenius are moderate; (iii) **2019:** Spectral distance peaks at 1.0.

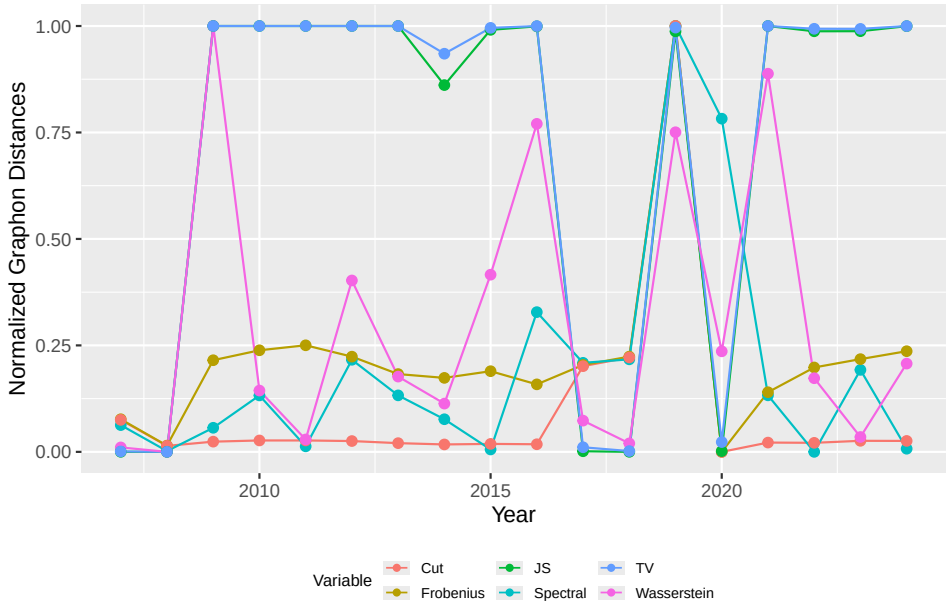


Fig. 6: Normalized Graphon Distances (2007–2024)

Figure 7 shows the local peaks of graphon distances over time. The years with the most simultaneous peaks are 2016 (Spectral, Wasserstein, JS, TV), 2019 (all), and 2021 (Cut, Wasserstein, JS, TV).

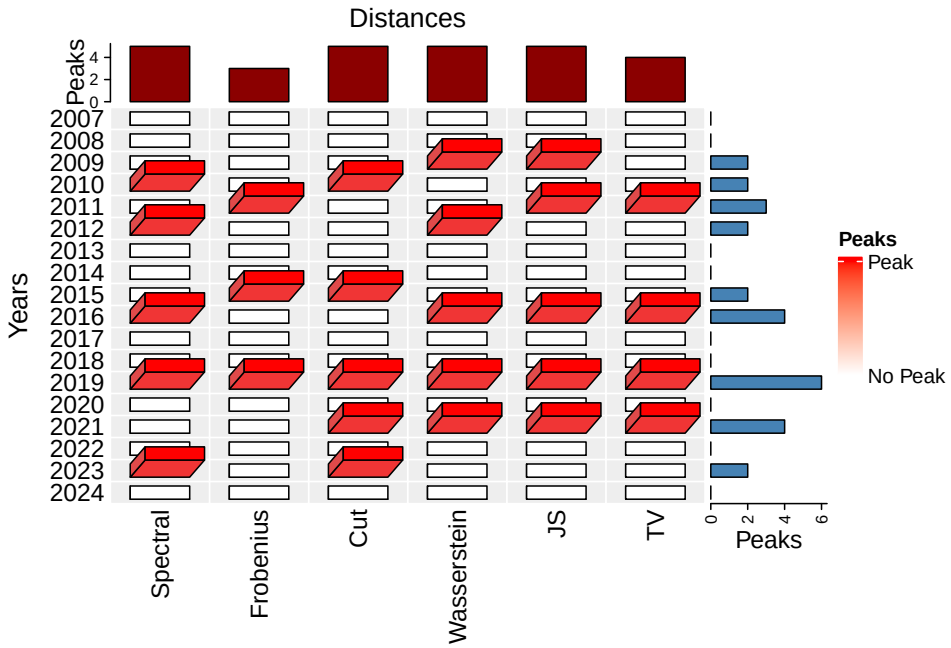
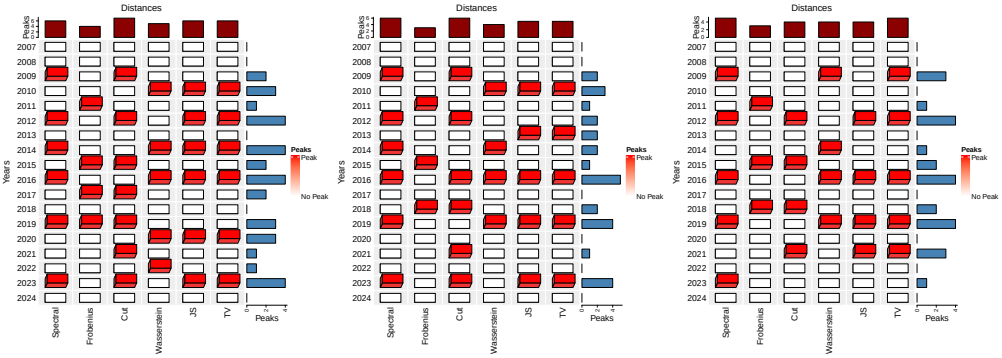


Fig. 7: Peaks of Graphon Distances

Figure 8 shows graphon distance peaks disaggregated by preference rank. First- and second-preference networks share similar peak patterns, while third-preference networks show some deviations.

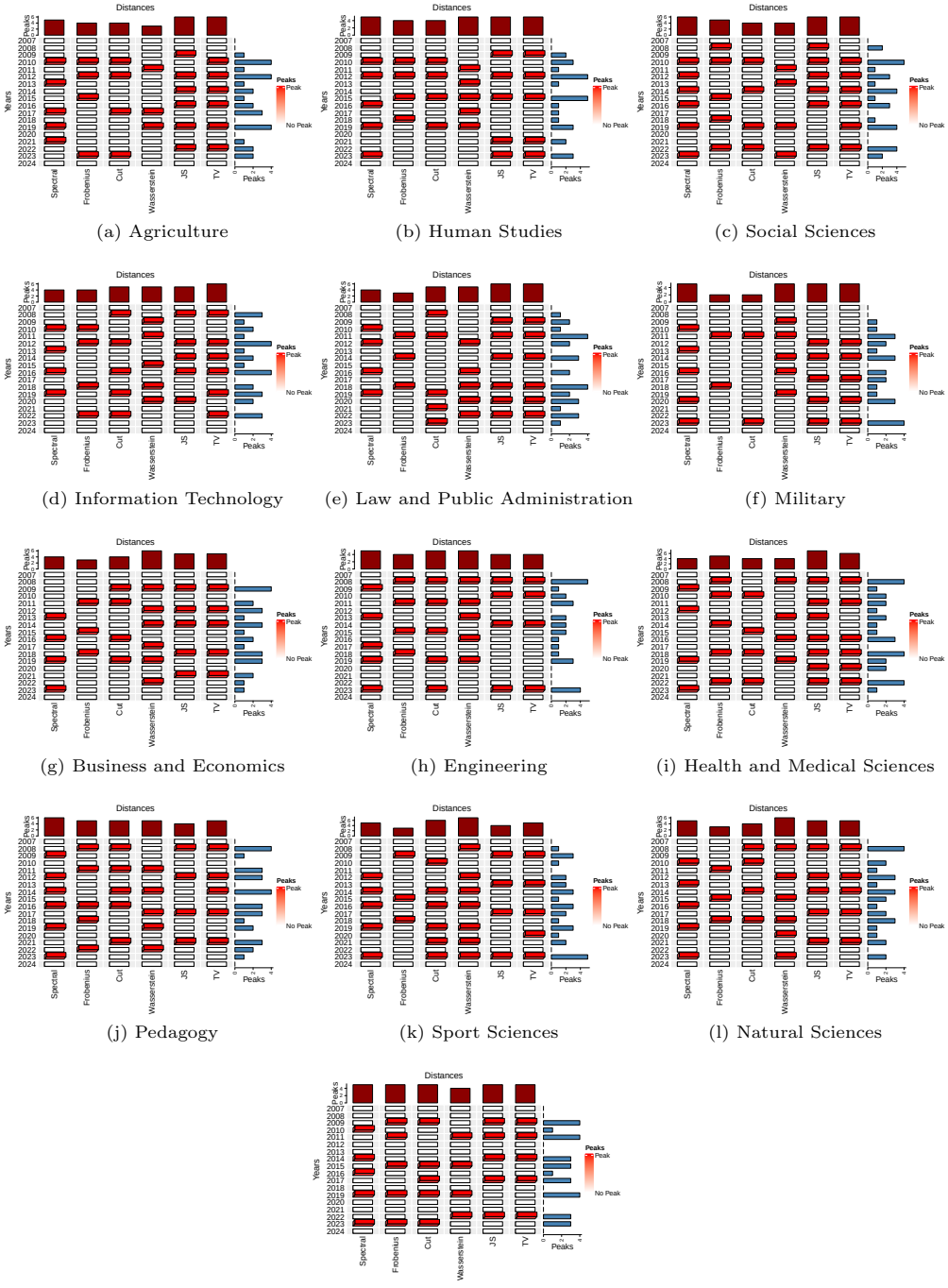


(a) 1st place (b) 2nd place (c) 3rd place

Fig. 8: Peaks of Graphon Distances

Figure 9 shows graphon distance peaks by field of study. STEM fields (Engineering, IT, Natural Sciences) and non-STEM fields (Business and Economics, Social Sciences)

show largely similar peak patterns, reinforcing the finding that structural shifts are primarily system-wide rather than field-specific.



(a) Agriculture (b) Human Studies (c) Social Sciences

(d) Information Technology (e) Law and Public Administration (f) Military

(g) Business and Economics (h) Engineering (i) Health and Medical Sciences

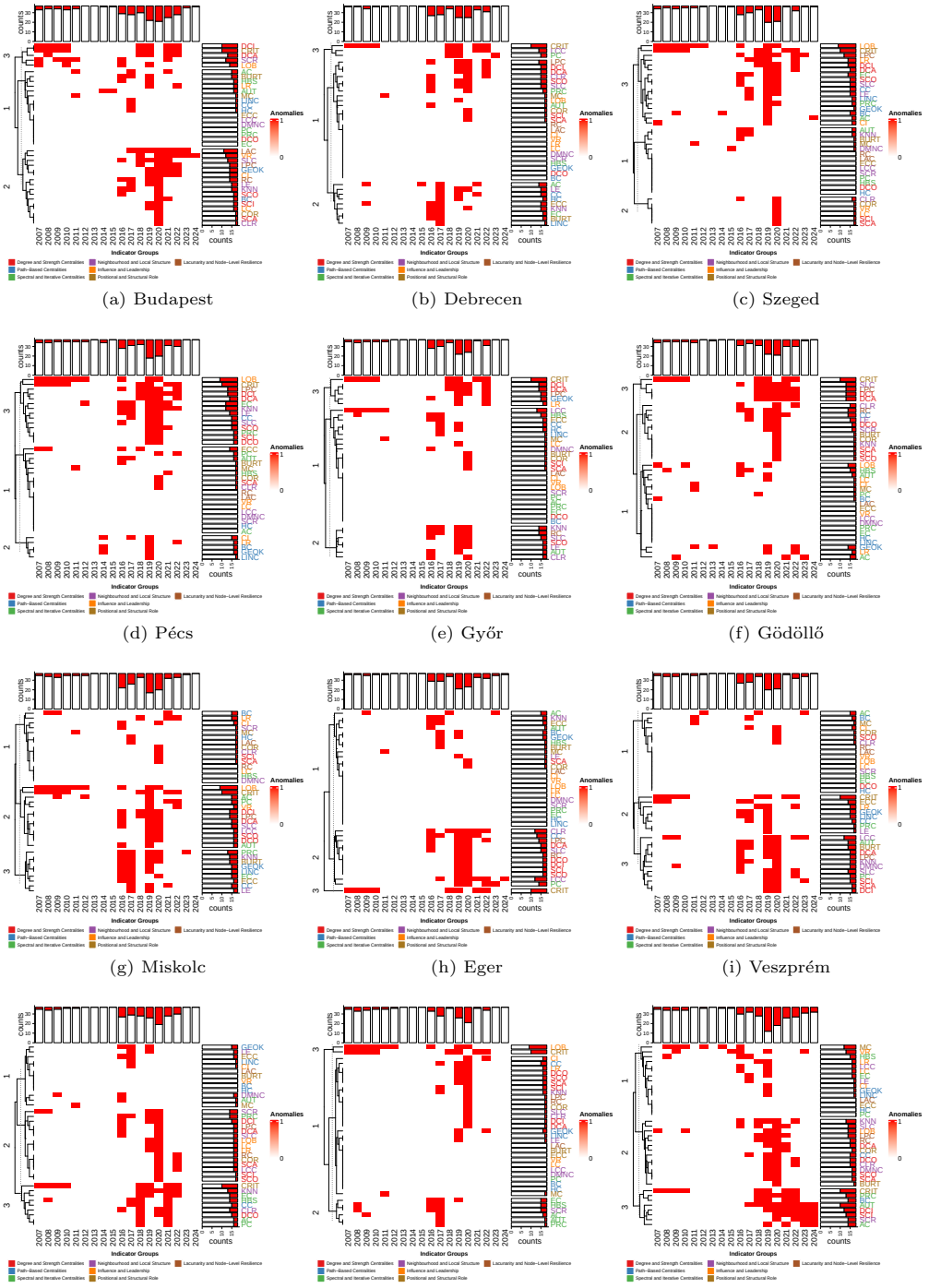
(j) Pedagogy (k) Sport Sciences (l) Natural Sciences

(m) Arts

Fig. 9: Peaks of Graphon Distances

4.2.3 Analysis of Node-Level Indicators

Figure 10 displays anomalies of the standardized growth ratios for node-level indicators in the top twelve incoming-student micro-regions, ordered by descending student volume: Budapest, Debrecen, Szeged, Pécs, Győr, Göodöllő, Miskolc, Eger, Veszprém, Nyíregyháza, Kecskemét, Székesfehérvár. Indicator names are colour-coded by the 7 logical groups defined in Table 4.



(a) Budapest (b) Debrecen (c) Szeged

(d) Pécs (e) Győr (f) Gödöllő

(g) Miskolc (h) Eger (i) Veszprém

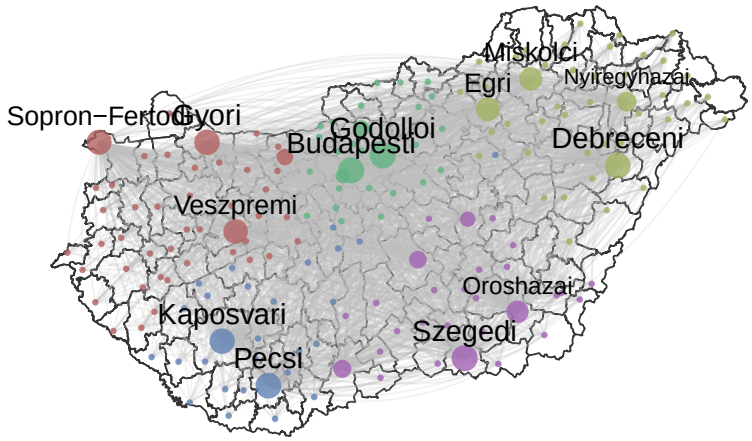
(j) Nyíregyháza (k) Kecskemét (l) Székesfehérvár

Fig. 10: Anomalies of the Growth Ratios (2007–2024) of Node-Level indicators for the Top Twelve Incoming Student Micro-Regions

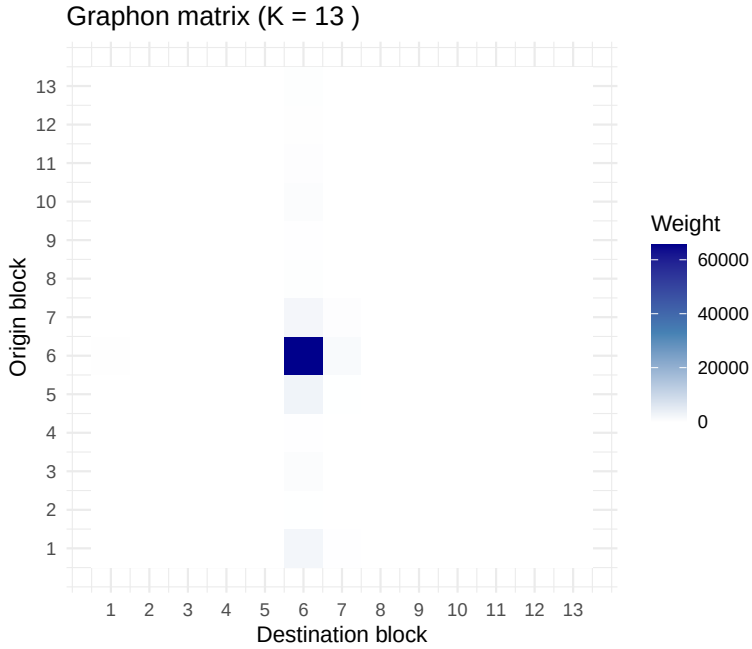
Across all micro-regions, the influence and leadership group (Group 5: LC, VR, LOB) and the positional and structural role group (Group 6: ECC, COR, BURT, CRIT) show the most frequent anomalies. Kecskemét (59) displays the fewest, Székesfehérvár (111) displays the greatest anomalies overall, while eastern micro-regions (Debrecen, Miskolc, Eger, Nyíregyháza) show more volatile patterns than middle micro-regions (Kecskemét, Budapest) or western ones (Győr, Veszprém).

4.2.4 Comparing Graphon-Based and Module-Based Community Detection

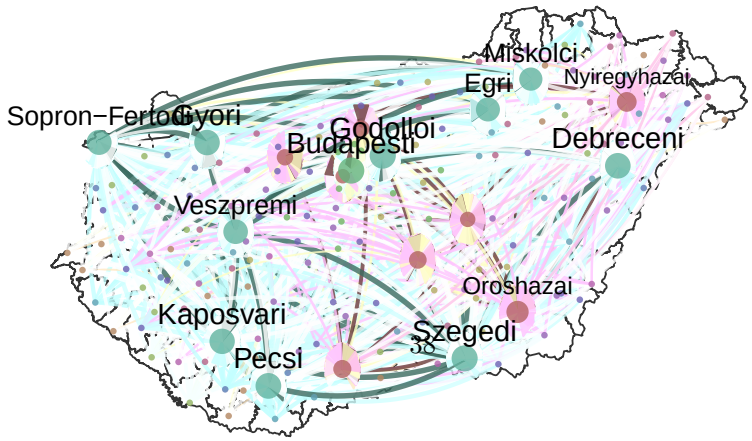
Figure 11 compares Leiden-based and graphon-based community detection for the 2015 application data. The names of the 13 micro-regions hosting the most students are labelled. Figure (11a) shows the Leiden communities, which identify groups of micro-regions with denser internal connections. Figure (11b) shows the graphon block matrix. Figure (11c) shows graphon-based communities together with inter-community flow pathways, where edges of different colours identify different flow routes.



(a) Leiden's Modularity



(b) Graphon Matrix

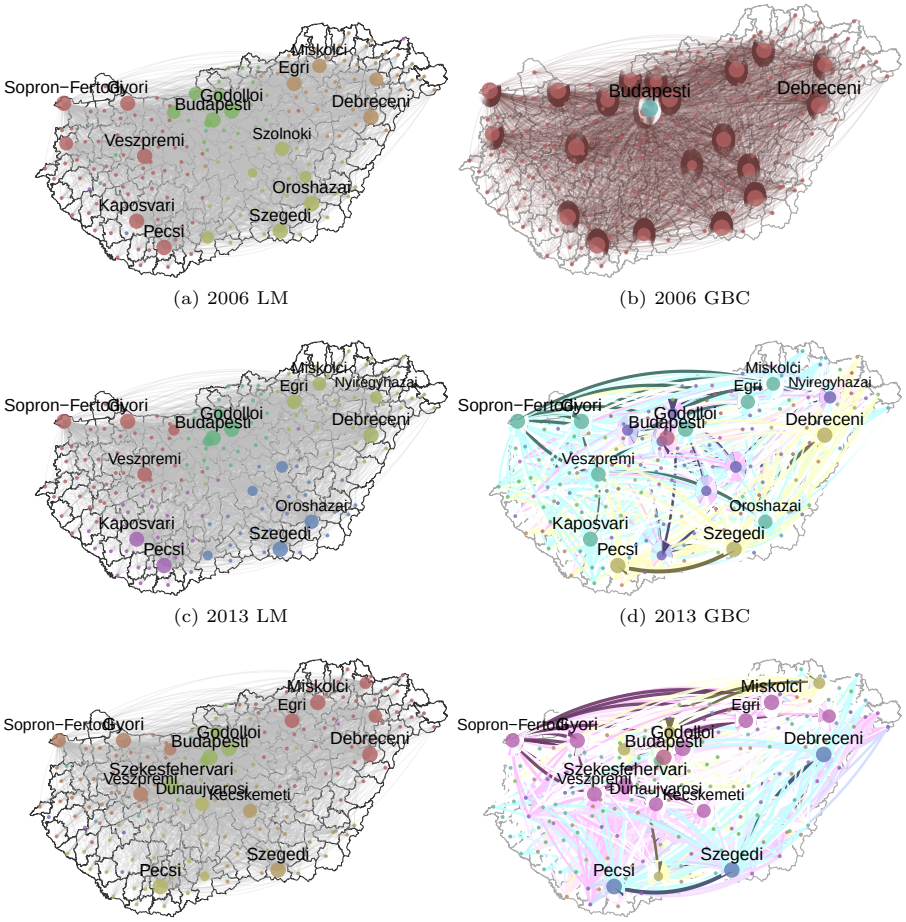


(c) Graphon-Based Communities

Fig. 11: Leiden's Modules (LMs), Graphon Matrix (GM), Graphon-Based Communities (GBCs) (2015)

he Leiden method identified 5 communities, roughly corresponding to geographic regions. The graphon-based method identified 13 communities with a clear hierarchy: Budapest alone forms block˜6; major university cities (Debrecen, Szeged, P'ecs, Győr, Miskolc, Eger, Veszpr'em, Kaposv'ar, Sopron) share block˜7.

Figure 12 compares the two detection methods across 2006, 2013, and 2024. Leiden communities (12a, 12c, 12e) are broadly similar across years. In contrast, graphon-based detection identified only 2 blocks in 2006 but 13 communities in both 2013 and 2024, reflecting increasing structural differentiation in the mobility network.



(a) 2006 LM (b) 2006 GBC

(c) 2013 LM (d) 2013 GBC

(e) 2024 LM (f) 2024 GBC

Fig. 12: Comparison of Module-Based (Leiden) and Graphon-Based Community Detection (2006, 2013, 2024)

5 Discussion

The section 4 documented a wide range of structural changes in Hungarian higher education mobility networks over nearly two decades. In this section, we interpret these findings through the lens of our three research questions, connecting the empirical patterns to policy events, international trends, and methodological insights. Rather than discussing each figure in isolation, we weave the evidence into a coherent narrative that addresses what has changed, why, and what the graphon-based approach uniquely reveals.

5.1 Structural Evolution and Policy Events (RQ_1)}

Our first research question asked how the structural patterns of Hungarian application networks evolved from 2006 to 2024 and which policy events are associated with detectable structural shifts.

The Bologna Transition and the Financial Crisis (2006–2010).

The period immediately following the Bologna transition was characterized by a sharp decline in application volume (Figure 1: from 148,216 in 2006 to 107,897 in 2008), followed by recovery. The graphon distance analysis (Figure 6) reveals that the JS and TV metrics reached their maximum (1.0) beginning in 2009, indicating that the probabilistic structure of the network—the distribution of flow intensities—underwent a fundamental shift. This is consistent with the introduction of Master's programmes in 2008, which created new pathways and redistributed application flows. Traditional network metrics showed relatively few anomalies in this period (Figure 3), with exceptions being KIGC (Kirchhoff index) in 2007 and NRI (resilience) in 2007–2008, suggesting that the network's robustness was temporarily affected.

The STEM Funding Reform (2011–2013).

The 2011 reform, which dramatically reduced state-funded places in Business, Law, and Social Sciences while raising minimum scores, caused the second major application dip (Figure 1: 109,720 in 2013). However, a striking finding is that the field-level anomaly patterns (Figures 5 and 9) show remarkable similarity between Engineering and Business—the structural changes were not field-specific. This suggests that while the reform successfully redirected student volumes (the share of Business declined from 24% to 22%), it did not differentially alter the spatial structure of mobility networks across fields. The reform moved students between fields but did not change where they applied from or to.

COVID-19 and the Foundation-Model Transformation (2019–2024).

The years 2019–2020 saw the highest concentration of network-level anomalies (16 indicators in 2019, 13 in 2020; Figure 3). All graphon distance metrics peaked in 2019 (Figure 6), indicating a shift in dominant structural modes. The graphon peak analysis shows 2021 as another key year, with four metrics simultaneously peaking (Figure 7). The IM (Infomap modularity) became persistently anomalous from 2021 onward

(Figure 3), suggesting a sustained community structure reorganisation coinciding with the foundation-model university transformation.

These findings extend earlier longitudinal studies of domestic mobility, such as Kosztyán et al (2021), by providing the first graphon-based evidence that policy events leave detectable structural signatures not only in traditional metrics but in the global geometric structure of mobility networks.

5.2 Sensitivity of Metrics to Structural Change (RQ_2)}

Our second research question asked which metrics are most sensitive to different types of policy-induced structural changes.

Traditional Network Metrics.

The growth-rate anomaly analysis (Figure 3) reveals a clear hierarchy of sensitivity. Resilience and robustness indicators (RRes: 8 anomalous years; NRI: 7) are the most reactive, consistent with the theoretical expectation that network robustness is sensitive to structural perturbations (Albert et al, 2000; Schneider et al, 2011). Connectivity and efficiency metrics (GLE, Arcs, Density: 3–4 anomalies each) form a second tier. Conversely, basic centralization measures (SZA, PRZ), assortativities, and fractal dimensions (FDIMALL, FDIMGC) show zero anomalies, indicating that the network's fundamental degree distribution and scale-free properties are structurally stable even as finer features change.

The correlation analysis (Figure 2) identified two growth-rate clusters: Cluster 1 captures volumetric changes (degree, strength, connectivity), and prestigious centralizations, such as EZ, AZ, PRZ, Cluster 2 captures resilience and network efficiency indicators, while Cluster 3 captures complexity and heterogeneity indicators, such as fractal dimensions, lacunarity and Kirchoff index, implying that volume changes and structural reorganisation are distinct phenomena that may respond to different policy drivers.

Graphon Distance Metrics.

The six graphon distances show complementary sensitivity patterns (Figure 6; Table 5). Cut and Frobenius distances are most sensitive to sudden volume-driven restructuring (peaking in 2019-2020, coinciding with structural change within the pandemic era). JS and TV distances respond to distributional changes (elevated in 2009–2013 during the Bologna/financial crisis period and again in 2021). Spectral distance is most sensitive to changes in dominant structural modes (peaking in 2019). Wasserstein distance, which quantifies the magnitude of structural rearrangement needed to transform one year's mobility pattern into another's, responds to both types of change.

This pattern directly addresses RQ_2 and provides practical guidance: researchers studying sudden volume shocks should prioritise Cut/Frobenius metrics, while those interested in gradual distributional shifts should focus on JS/TV metrics.

Node-Level Sensitivity.

At the node level (Figure 10), influence and leadership indicators (Group 5) and positional indicators (Group 6) are most frequently anomalous. This is consistent with

the theoretical argument that brokerage and structural hole positions (Burt, 2003) are more volatile than basic degree centrality in evolving networks.

The middle micro-regions, such as Kecskemét and Budapest displayed the fewest node-level anomalies, confirming its structural stability as the dominant hub. Eastern micro-regions (Debrecen, Miskolc, Eger) were more volatile than western ones (Győr, Veszprém), which may reflect the stronger impact of demographic decline and more aggressive institutional restructuring in eastern Hungary. The government's ambition to strengthen regional higher education institutions is partially visible: Szeged's out-strength uniquely increased (Table 7), suggesting enhanced regional connectivity, while Pécs experienced dramatic decline (out-strength: −78%).

5.3 Graphon-Based vs. Traditional Analysis (RQ_3)

Our third research question asked to what extent graphon-based metrics reveal structural changes invisible to traditional metrics and how the two approaches complement each other.

What Graphon Distances Add.

The comparison reveals several instances where graphon-based analysis identifies changes that traditional metrics miss or underrepresent:

1. **The 2019-2020 structural shift before the pandemic era.** In 2019, local maximums of all graph distance metrics can be observed. After 2019, the 2020 mobility network changed less structurally, which shows that COVID halted changes for a while, and then in 2021 the mobility network underwent another structural transformation.
2. **The 2021 structural shift.** Four graphon metrics simultaneously peaked in 2021 (Figure 7), yet only NRI and RRes showed anomalies in the traditional analysis (Figure 3). The graphon distances thus detected a global structural reorganisation—likely related to the foundation-model transformation—that was not fully captured by individual network statistics.
3. **The 2009–2013 distributional change.** JS and TV distances remained at or near 1.0 throughout this period (Figure 6), signalling a persistent probabilistic restructuring. Traditional metrics showed only scattered anomalies (KIGC in 2007 and 2012, BEZ in 2012), failing to capture the sustained nature of this change.
4. **Community differentiation over time.** The graphon-based community detection (Figure 12) identified only 2 blocks in 2006 but 13 in 2013 and 2024, revealing a dramatic increase in structural differentiation. The Leiden method found similar numbers of communities across all three years, missing this evolution entirely.

Where Traditional Metrics Excel.

Traditional metrics provide node-level specificity that global graphon distances cannot: identifying exactly which micro-regions changed (Figure 10) and which aspects of their role shifted. They also identify the specific growth-rate clustering (Figure 2), which informs interpretation of what type of change occurred.

Complementarity.

The two approaches are complementary rather than competing. Traditional metrics answer `what changed and where,'' while graphon distances answer`how different is the entire network from its previous state.'' Graphon-based community detection adds a third layer: "how has the internal grouping of the network evolved.'' This integrative framework (C_4) provides a richer picture than either approach alone.

Alignment Between Network Metrics and Graphon Distances.

Comparing the anomaly patterns of traditional metrics (Figure 3) with graphon distance peaks (Figure 7), we find partial overlap: both identify 2019 and 2021 as critical years. However, the overlap is incomplete, confirming that the two approaches capture partially different aspects of structural change. Among the traditional metric groups, connectivity and efficiency metrics (Group 5) and complexity metrics (Group 9) align most closely with Frobenius and Cut graphon distances, consistent with the theoretical connections outlined in Table 5. Modularity and community metrics (Group 4) align with Spectral distance, as expected. Resilience metrics show temporal patterns distinct from any single graphon metric, suggesting they capture complementary structural information.

5.4 Local vs. Global Drivers of Structural Change

Many of the structural shifts we identified can be attributed to domestic policy interventions (the Bologna transition, the STEM reform, the foundation-model transformation, COVID-induced adaptations). However, some patterns likely reflect global trends. The sustained growth of IT applications (+115% from 2006 to 2018; Figure 1) mirrors the worldwide STEM appreciation trend documented by Choudaha (2017). The decline of Social Sciences and Human Studies parallels patterns observed in Erasmus mobility networks (Breznik and Gologranc, 2017). The increasing hierarchical centralization (AZ: +2,878%; Table 6) echoes findings from international student flow analyses, where dominant hubs increasingly attract disproportionate flows (Shields, 2013; Barnett et al, 2016).

The co-movement of anomaly patterns across fields (Figure 5) is a novel finding: to our knowledge, no previous study has documented that field-specific mobility networks undergo structurally similar—rather than divergent—changes in response to field-targeted policy reforms. This suggests that the underlying spatial structure of mobility is more strongly determined by geography and institutional prestige than by field-specific policy incentives.

An important limitation is that our data do not capture applications to foreign institutions. Estimates suggest that 5–10% of Hungarian applicants also consider international options (Kosztyán et al, 2021). This invisible outflow primarily affects first-preference networks and may partially explain the decline in domestic first-preference volumes.

6 Summary and Conclusion

This study examined the structural evolution of Hungarian higher education mobility networks at the micro-region level from 2006 to 2024. We analysed directed, weighted networks of 175 micro-regions using 37 network-level indicators, 37 node-level indicators, and 6 graphon distance metrics, disaggregated by preference rank and field of study.

Our main findings are as follows. The network became sparser, more hierarchical, and more strongly modularised over the study period, with three major structural breakpoints: the Bologna transition (2006–2008), the 2019 structural change, and the foundation-model transformation period (2021–2024). Resilience metrics proved most sensitive to short-term shocks, while graphon distances—particularly JS and TV—captured sustained distributional changes invisible to traditional metrics. Graphon-based community detection revealed a hierarchical layering (Budapest–major universities–smaller institutions) that Leiden modularity failed to identify.

6.1 Implications for Scholars

This study makes four methodological and empirical contributions:

- $\mathbf{C}_1$ **First application of graphon distance metrics to domestic mobility networks.** We demonstrated that graphon-based analysis captures global structural shifts—such as the 2021 reorganisation and the 2009–2013 distributional change—that traditional metrics miss or underrepresent (Figures 6, 7).
- $\mathbf{C}_2$ **Systematic comparison of six graphon distance metrics.** We showed that Cut/Frobenius distances are most sensitive to volume-driven restructuring, JS/TV to distributional shifts, and Spectral distance to dominant structural mode changes (Figure 6, Table 5). This provides practical guidance for metric selection in future studies.
- $\mathbf{C}_3$ **Comprehensive longitudinal analysis of sub-national mobility networks.** Our 19-year, preference-disaggregated, field-disaggregated analysis is among the most detailed longitudinal studies of domestic mobility. A novel finding is that field-specific networks undergo structurally similar changes despite field-targeted policies (Figures 5, 9).
- $\mathbf{C}_4$ **Integrative framework linking traditional metrics to graphon-based analysis.** We demonstrated that the two approaches are complementary: traditional metrics specify what changed and where, graphon distances quantify how different the entire network became, and graphon-based community detection reveals hierarchical layering invisible to modularity-based methods (Figures 11, 12).

The graphon-based community analysis represents a particularly promising avenue for future research. Its ability to identify not only communities but also inter-community flow hierarchies (Figure 11) opens new possibilities for understanding the layered structure of educational mobility.

6.2 Implications for Decision-Makers

Our findings carry several practical implications for higher education policy:

1. **Geographic distance is not the only driver.** The graphon-based community detection revealed significant inter-regional flows between university micro-regions that are not geographically adjacent (Figure 11). This suggests that institutional prestige and programme offerings matter alongside distance, which should inform recruitment and enrolment strategies.
2. **Funding reforms redirect volumes but not structures.** The 2011 STEM reform successfully shifted student volumes between fields (Figure 1) but did not differentially alter the spatial structure of field-specific mobility networks (Figure 5). Policymakers should recognise that volume-targeting instruments may not achieve spatial restructuring goals.
3. **Regional strengthening is uneven.** The government's ambition to strengthen regional institutions is only partially reflected in the data: Szeged's out-strength uniquely increased (+16%; Table 7), while Pécs experienced dramatic decline (−78% out-strength). Eastern micro-regions (Debrecen, Miskolc, Eger) show more volatile structural indicators (Figure 10), reflecting active but uneven restructuring.
4. **A hierarchical application structure exists.** The graphon community analysis consistently identified a three-tier hierarchy—Budapest, major regional universities, and smaller institutions—suggesting that enrolment planning should account for this layered structure rather than treating all non-Budapest regions uniformly.

Declarations

- **Funding:** N/A
- **Competing interests:** The authors declare no competing interests.
- **Data availability:** All data are available in the Github ...
- **Code availability:** All code are available in the Github ...

References

Airoldi EM, Costa TB, Chan SH (2013) Stochastic blockmodel approximation of a graphon: Theory and consistent estimation. Advances in Neural Information Processing Systems 26:692–700

Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. Nature 406:378–382. https://doi.org/10.1038/35019019

Anand K, Bianconi G (2009) Entropy measures for networks: Toward an information theory of complex topologies. Physical Review E 80(4):045102. https://doi.org/10.1103/PhysRevE.80.045102

Avola M, Giuga R (2025) Social origin, geographical inequalities and international student mobility: the italian case. International Journal of Sociology and Social Policy 45(5-6):546–566. https://doi.org/https://doi.org/10.1108/IJSSP-12-2024-0630

Barabási AL (2016) Network Science. Cambridge University Press, Cambridge

Barnett GA, Lee M, Jiang K, et al (2016) The flow of international students from a macro perspective: A network analysis. Compare: A Journal of Comparative and International Education 46(4):533–559. https://doi.org/10.1080/03057925.2015.1015551

Beine M, Noël R, Ragot L (2014) Determinants of the international mobility of students. Economics of Education Review 41:40–54. https://doi.org/10.1016/j.econedurev.2014.03.003

Blázquez-García A, Conde A, Mori U, et al (2021) A review on outlier/anomaly detection in time series data. ACM Computing Surveys 54(3):1–33. https://doi.org/10.1145/3444690

Bonacich P (1987) Power and centrality: A family of measures. American Journal of Sociology 92(5):1170–1182. https://doi.org/10.1086/228631

Borgs C, Chayes JT, Lovász L, et al (2008) Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. Advances in Mathematics 219(6):1801–1851. https://doi.org/10.1016/j.aim.2008.07.008

Borgs C, Chayes JT, Lovász L, et al (2012) Convergent sequences of dense graphs II: Multiway cuts and statistical physics. Annals of Mathematics 176(1):151–219. https://doi.org/10.4007/annals.2012.176.1.2

Borgs C, Chayes JT, Cohn H, et al (2019) An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. Transactions of the American Mathematical Society 372(5):3019–3062. https://doi.org/10.1090/tran/7543

Breznik K, Gologranc G (2017) Erasmus student mobility flows. European Journal of Education 52(1):105–117. https://doi.org/10.1111/ejed.12193

Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30(1-7):107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Burt RS (2003) The social structure of competition, vol 13. Oxford University Press New York, NY

Carlson S, França T, Lomer S, et al (2025) Social inequalities in international student mobility: reviewing conceptual approaches. British Journal of sociology of Education 46(6):834–853. https://doi.org/https://doi.org/10.1080/01425692.2025.2520288

Chankseliani M, Kwak J (2025) The ripple effect: Understanding the societal implications of international student mobility. International Journal of Educational

Research 129:102520. https://doi.org/https://doi.org/10.1016/j.ijer.2024.102520

Chatterjee S (2015) Matrix estimation by universal singular value thresholding. Annals of Statistics 43(1):177–214. https://doi.org/10.1214/14-AOS1272

Chen C, Liu LM (1993) Joint estimation of model parameters and outlier effects in time series. Journal of the American Statistical Association 88(421):284–297. https://doi.org/10.1080/01621459.1993.10594321

Chen D, Lü L, Shang MS, et al (2012) Identifying influential nodes in complex networks. Physica A: Statistical Mechanics and its Applications 391(4):1777–1787. https://doi.org/10.1016/j.physa.2011.09.017

Choudaha R (2017) Three waves of international student mobility (1999–2020). Studies in Higher Education 42(5):825–832. https://doi.org/10.1080/03075079.2017.1293872

Chung FRK (1997) Spectral Graph Theory, CBMS Regional Conference Series in Mathematics, vol 92. American Mathematical Society, https://doi.org/10.1090/cbms/092

Civera A, Meoli M, Paleari S (2023) International student mobility: Onset for a future career or an experiential opportunity? International Studies in Sociology of Education 32(2):559–578. https://doi.org/https://doi.org/10.1080/09620214.2021.2007417

Colizza V, Flammini A, Serrano MÁ, et al (2006) Detecting rich-club ordering in complex networks. Nature Physics 2:110–115. https://doi.org/10.1038/nphys209

Derzsy N, Néda Z, Santos M (2012) Topology of the Erasmus student mobility network. Physica A: Statistical Mechanics and its Applications 391(23):6190–6196. https://doi.org/10.1016/j.physa.2012.07.038

Di Pietro G, Perez-Encinas A (2025) The effect of covid-19 on international student credit mobility: a gravity model approach. Education Economics 33(1):76–92. https://doi.org/https://doi.org/10.1080/09645292.2023.2297151

Diaconis P, Janson S (2008) Graph limits and exchangeable random graphs. Rendiconti di Matematica e delle sue Applicazioni 28:33–61. https://doi.org/https://doi.org/10.48550/arXiv.0712.2749

Donnat C, Holmes S (2018) Tracking network dynamics: A survey using graph distances. Annals of Applied Statistics 12(2):971–1012. https://doi.org/10.1214/18-AOAS1176

Findlay AM, King R, Smith FM, et al (2012) World class? an investigation of globalisation, difference and international student mobility. Transactions of the Institute

of British Geographers 37(1):118–131. https://doi.org/10.1111/j.1475-5661.2011.00454.x

Freeman LC (1979) Centrality in social networks: Conceptual clarification. Social Networks 1(3):215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Gao C, Lu Y, Zhou HH (2015) Rate-optimal graphon estimation. Annals of Statistics 43(6):2624–2652. https://doi.org/10.1214/15-AOS1354

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99(12):7821–7826. https://doi.org/10.1073/pnas.122653799

González CR, Mesanza RB, Mariel P (2011) The determinants of international student mobility flows: An empirical study on the Erasmus programme. Higher Education 62(4):413–430. https://doi.org/10.1007/s10734-010-9396-5

Joyce KE, Laurienti PJ, Burdette JH, et al (2010) A new measure of centrality for brain networks. PLoS ONE 5(8):e12200. https://doi.org/10.1371/journal.pone.0012200

Katz L (1953) A new status index derived from sociometric analysis. Psychometrika 18(1):39–43. https://doi.org/10.1007/BF02289026

Klein DJ, Randić M (1993) Resistance distance. Journal of Mathematical Chemistry 12:81–95. https://doi.org/10.1007/BF01164627

Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. Journal of the ACM 46(5):604–632. https://doi.org/10.1145/324133.324140

Klopp O, Tsybakov AB, Verzelen N (2017) Oracle inequalities for network models and sparse graphon estimation. Annals of Statistics 45(1):316–354. https://doi.org/10.1214/16-AOS1443

Kondakci Y, Bedenlier S, Zawacki-Richter O (2018) Social network analysis of international student mobility: Uncovering the rise of regional hubs. Higher Education 75(3):517–535. https://doi.org/10.1007/s10734-017-0154-9

Kosztyán ZT, Csányi VV, Telcs A (2020) Scores and clusters of hungarian universities. Acta Oeconomica 70(2):229–249. https://doi.org/https://doi.org/10.1556/032.2020.00012

Kosztyán ZT, Csányi VV, Banász Z, et al (2021) The role of higher education in spatial mobility. Applied Network Science 6(1):88. https://doi.org/https://doi.org/10.1007/s41109-021-00428-w

Kropp K, Larsen AG (2025) Students in unequal exchange: The european social space of erasmus student mobility. Social Science Information 64(1):60–83. https://doi.

org/https://doi.org/10.1177/05390184251324888

Kunszenti-Kovács D, Lovász L, Szegedy B (2022) Multigraph limits and exchangeability. Memoirs of the American Mathematical Society 279. https://doi.org/10.1090/memo/1392

Latora V, Marchiori M (2001) Efficient behavior of small-world networks. Physical Review Letters 87(19):198701. https://doi.org/10.1103/PhysRevLett.87.198701

Litmeyer ML, Gareis P, Hennemann S (2023) Comparing student mobility pattern models. European Journal of Geography 14(1):21–34. https://doi.org/https://doi.org/10.48088/ejg.m.lit.14.1.21.34

Lovász L (2012) Large Networks and Graph Limits, Colloquium Publications, vol 60. American Mathematical Society, Providence, RI, https://doi.org/10.1090/coll/060

Lovász L, Szegedy B (2006) Limits of dense graph sequences. Journal of Combinatorial Theory, Series B 96(6):933–957. https://doi.org/10.1016/j.jctb.2006.05.002

von Luxburg U (2007) A tutorial on spectral clustering. Statistics and Computing 17(4):395–416. https://doi.org/10.1007/s11222-007-9033-z

Milo R, Shen-Orr S, Itzkovitz S, et al (2002) Network motifs: Simple building blocks of complex networks. Science 298(5594):824–827. https://doi.org/10.1126/science.298.5594.824

Morone F, Makse HA (2015) Influence maximization in complex networks through optimal percolation. Nature 524:65–68. https://doi.org/10.1038/nature14604

Newman MEJ (2018) Networks, 2nd edn. Oxford University Press, Oxford, https://doi.org/10.1093/oso/9780198805090.001.0001

Pérez AM, Cortés R, Carrier A, et al (2025) Contributions of educational investment to international student mobility: A multivariate and neutrosophic analysis of educational service quality in oecd data. Neutrosophic Sets and Systems 89(1):29. URL https://digitalrepository.unm.edu/nss_journal/vol89/iss1/29

Perkins R, Neumayer E (2015) Geographies of educational mobilities: Exploring unevenness, difference and changes in international student flows. The Geographical Journal 180(3):246–259. https://doi.org/10.1111/geoj.12070

Peyré G, Cuturi M (2019) Computational optimal transport: With applications to data science. Foundations and Trends in Machine Learning 11(5–6):355–607. https://doi.org/10.1561/2200000073

Qi X, Fuller E, Wu Q, et al (2012) Laplacian centrality: A new centrality measure for weighted networks. Information Sciences 194:240–253. https://doi.org/10.1016/j.ins.2011.12.027

Reichardt J, Bornholdt S (2006) Statistical mechanics of community detection. Physical Review E 74(1):016110. https://doi.org/10.1103/PhysRevE.74.016110

Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences 105(4):1118–1123. https://doi.org/10.1073/pnas.0706851105

Sá C, Florax RJGM, Rietveld P (2004) Determinants of the regional demand for higher education in The Netherlands: A gravity model approach. Regional Studies 38(4):375–392. https://doi.org/10.1080/03434002000213905

Schneider CM, Moreira AA, Andrade Jr. JS, et al (2011) Mitigation of malicious attacks on networks. Proceedings of the National Academy of Sciences 108(10):3838–3841. https://doi.org/10.1073/pnas.1009440108

Shields R (2013) Globalization and international student mobility: A network analysis. Comparative Education Review 57(4):609–636. https://doi.org/10.1086/671752

Song C, Havlin S, Makse HA (2005) Self-similarity of complex networks. Nature 433:392–395. https://doi.org/10.1038/nature03248

Tan Y, Lo L, Li W, et al (2024) Intellectual capital and student mobility. In: Intellectual Migration. Routledge, p 65–85, https://doi.org/https://doi.org/10.4324/9781003532934

Tantardini M, Ieva F, Tajoli L, et al (2019) Comparing methods for comparing networks. Scientific Reports 9:17557. https://doi.org/10.1038/s41598-019-53708-y

Telcs A, Kosztyan ZT, Toeroek A (2016) Unbiased one-dimensional university ranking–application-based preference ordering. Journal of Applied Statistics 43(1):212–228. https://doi.org/https://doi.org/10.1080/02664763.2014.998180

Traag VA, Waltman L, van Eck NJ (2019) From Louvain to Leiden: Guaranteeing well-connected communities. Scientific Reports 9:5233. https://doi.org/10.1038/s41598-019-41695-z

Van Mol C, Timmerman C (2014) Beyond the nation-state: Cultural capital and the shift of student mobility patterns in the European Higher Education Area. Population, Space and Place 20(2):123–140. https://doi.org/10.1002/psp.1772

Vickers D, Rees P (2017) Ground-truthing the geography of student migration. Environment and Planning A 49(12):2697–2715. https://doi.org/10.1177/0308518X17729880

Villani C (2009) Optimal Transport: Old and New. Springer, Berlin, https://doi.org/10.1007/978-3-540-71050-9

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:440–442. https://doi.org/10.1038/30918

Weber T, Van Mol C (2023) The student migration transition: an empirical investigation into the nexus between development and international student migration. Comparative Migration Studies 11(1):5. https://doi.org/https://doi.org/10.1186/s40878-023-00329-0

Wills P, Meyer FG (2020) Metrics for graph comparison: A practitioner's guide. PLoS ONE 15(2):e0228728. https://doi.org/10.1371/journal.pone.0228728

Zhang JL, Chen XS, Lu JA, et al (2016) Voterank: A new centrality measure for identifying the most influential spreaders in complex networks. Scientific Reports 6:27823. https://doi.org/10.1038/srep27823

Zhou S, Mondragón RJ (2004) The rich-club phenomenon in the Internet topology. IEEE Communications Letters 8(3):180–182. https://doi.org/10.1109/LCOMM.2004.823426

A

A.1 Node-Level Indicators: Detailed Definitions

In the following, let $G = (V, E, w)$ be a directed, weighted graph with $n = |V|$ nodes. For nodes $i, j \in V$: $w(i, j)$ denotes the edge weight (0 if no edge); $a_{ij} = \mathbf{1}[w(i, j) > 0]$ is the binary adjacency indicator; $N^{\text{in}}(i) = \{j : a_{ji} > 0\}$ and $N^{\text{out}}(i) = \{j : a_{ij} > 0\}$ are the in- and out-neighbourhoods; $N(i) = N^{\text{in}}(i) \cup N^{\text{out}}(i)$ is the undirected neighbourhood; $d^{\text{in}}(i) = |N^{\text{in}}(i)|$, $d^{\text{out}}(i) = |N^{\text{out}}(i)|$, $d(i) = |N(i)|$ are the in-, out-, and total degrees; $d(i, j)$ denotes the shortest path distance from i to j (in terms of hop count or weighted distance); σ_{st} is the number of shortest paths from s to t; $\sigma_{st}(v)$ is the number of shortest paths from s to t passing through v; and $\lambda_1 \geq \lambda_2 \geq \cdots$ are the eigenvalues of the adjacency matrix A (or weight matrix W); c_v is the local clustering coefficient of node v (the fraction of pairs among v's neighbours that are connected); $C_{\max}(v)$ is the largest connected component in the subgraph induced by $N(v)$, and $E(S)$ denotes the edge set of the subgraph induced by node set S; $R(v) = \{u \in V : d(v, u) < \infty\}$ is the reachable set from v; $\partial B(v, \ell) = \{u : d(v, u) = \ell\}$ is the boundary of the ball of radius ℓ around v; H_s denotes the s-core of the graph (the maximal subgraph in which every node has weighted degree at least s); $s^w(u) = \sum_{v \in H_s} w(u, v)$ is the weighted degree of u within H_s; $\text{GE}(G) = E_{\text{glob}}(G)$ is the global efficiency of graph G; $\text{MFPT}(s, v)$ is the mean first-passage time from node s to v in a random walk on G; $S(G)$ is the number of nodes in the largest (strongly) connected component of G, and $\Delta S(v) = S(G) - S(G \setminus \{v\})$; $E_L(G) = \sum_i \lambda_i^2(L)$ is the Laplacian energy, where $L = D - A$ is the Laplacian matrix and D is the diagonal degree matrix; $M(v, r)$ is the number of nodes within distance r from v (the "mass'' at scale r).

Group 1: Degree & Strength Centrality

Abbr.	Name	Formula	Description
SCI	In-strength Centrality	$s^{\text{in}}(i) = \sum_{j \in V} w(j, i)$	Sum of incoming edge weights; total number of applications received by micro-region i.
SCO	Out-strength Centrality	$s^{\text{out}}(i) = \sum_{j \in V} w(i, j)$	Sum of outgoing edge weights; total number of applications sent from micro-region i.
SCA	Strength Centrality	$s(i) = s^{\text{in}}(i) + s^{\text{out}}(i)$	Total weighted degree.
DCI	In-degree Centrality	$d^{\text{in}}(i) = \sum_{j \in V} a_{ji}$	Number of incoming edges.
DCO	Out-degree Centrality	$d^{\text{out}}(i) = \sum_{j \in V} a_{ij}$	Number of outgoing edges.
DCA	Degree Centrality	$d(i) = d^{\text{in}}(i) + d^{\text{out}}(i)$	Total (unweighted) degree.

Group 2: Path-based Centrality
Group 3: Prestige Centrality
Group 4: Neighbourhood & Local Structure
Group 5: Influence & Leadership
Group 6: Positional & Structural Role
Group 7: Lacunarity and Node-Level Resilience

Abbr.	Name	Formula	Description
BC	Betweenness Centrality	$B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$	Fraction of all shortest paths through v; measures the micro-region's bridging role in the application flow network.
CC	Closeness Centrality	$C_C(v) = \frac{n-1}{\sum_{u \neq v} d(v,u)}$	Inverse of average distance to all other nodes; measures reachability.
HC	Harmonic Centrality	$H(v) = \frac{1}{n-1} \sum_{u \neq v} \frac{1}{d(v,u)}$	Average inverse distance; handles disconnected graphs.
GEOK	Geodesic K-path Centrality	$G_K(v) = \sum_{u \neq v} \frac{1}{d(v,u)} \cdot \mathbf{1}[d(v,u) \leq K]$	Reachability within K hops, weighted by inverse distance.
LINC	Lin Centrality	$L(v) = \frac{\lvert R(v)\rvert^2}{\sum_{u \in R(v)} d(v,u)}$	Balances reachable set size against average distance; $R(v) = \{u : d(v,u) < \infty\}$.

Abbr.	Name	Formula	Description
EC	Eigenvector Centrality	$x_i = \frac{1}{\lambda_1} \sum_j a_{ij} x_j$	Component of the dominant eigenvector of A.
AUT	Authority Centrality	$a_i^{(k+1)} = \sum_j h_j^{(k)} \cdot a_{ji}$	HITS authority score: high if pointed to by good hubs (Kleinberg, 1999).
HBS	Hub Centrality	$h_i^{(k+1)} = \sum_j a_j^{(k)} \cdot a_{ij}$	HITS hub score: high if pointing to good authorities.
PRC	PageRank Centrality	$\mathrm{PR}(v) = \frac{1-\alpha}{n} + \alpha \sum_{u \in N^{\mathrm{in}}(v)} \frac{\mathrm{PR}(u)}{d^{\mathrm{out}}(u)}$	Damped random-walk importance; α is the damping factor (Brin and Page, 1998).
AC	Alpha Centrality	$\mathbf{x} = (\mathbf{I} - \alpha \mathbf{A}^T)^{-1} \mathbf{e}$	Generalised eigenvector centrality with exogenous input $\mathbf{e}$ (Katz, 1953; Bonacich, 1987).
PC	Power Centrality	$c_i(\beta) = \sum_j (\alpha + \beta c_j) a_{ij}$	Bonacich power centrality; β controls neighbour influence direction (Bonacich, 1987).

Abbr.	Name	Formula	Description
KNN	K-nearest Neighbour Degree	$k_{\mathrm{nn}}(v) = \frac{1}{d(v)} \sum_{u \in N(v)} d(u)$	Average degree of neighbours; indicates degree–degree correlation.
SCR	S-core Value	Maximal s s.t. $v \in H_s$, $\forall u \in H_s : s^w(u) \geq s$	Weighted k-core membership; s^w is the weighted degree within H_s.
CLR	ClusterRank	$\mathrm{CR}(v) = f(c_v) \cdot (d^{\mathrm{out}}(v) + 1)$	Combines local clustering c_v with out-degree; $f(\cdot)$ is a nonlinear transform (Chen et al, 2012).
DMNC	Density of Max. Neighbourhood Component	$\frac{\lvert E(C_{\max}(v))\rvert}{\lvert C_{\max}(v)\rvert \cdot (\lvert C_{\max}(v)\rvert - 1)}$	Edge density of the largest connected component in $N(v)$.
LCC	Local Community Centrality	$\mathrm{LCC}(v) = \sum_{u \in N(v)} \lvert\{w \in N(v) : (u,w) \in E\}\rvert$	Counts internal community connections around v.
SLC	Semi-local Centrality	$\mathrm{SL}(v) = \sum_{u \in N(v)} \sum_{w \in N(u)} d(w)$	Two-hop degree sum; captures extended local influence.
LE	Local Efficiency	$E_{\mathrm{loc}}(v) = \frac{1}{\lvert N(v)\rvert(\lvert N(v)\rvert - 1)} \sum_{\substack{j,k \in N(v) \\ j \neq k}} \frac{1}{d_{jk}^{N(v)}}$	Efficiency of communication among v's neighbours (excluding v).

Abbr.	Name	Formula	Description
LC	Leverage Centrality	$l(v) = \frac{1}{d(v)} \sum_{u \in N(v)} \frac{d(v)-d(u)}{d(v)+d(u)}$	Positive if v has higher degree than its neighbours on average (Joyce et al, 2010).
LR	LeaderRank	PageRank variant with a "ground node" connected to all	Adds a bidirectional ground node to handle sinks; converges without damping.
VR	VoteRank	Iterative voting: neighbours vote for v; top voter removed each round	Identifies spreading influencers by iterative elimination (Zhang et al, 2016).
CI	Collective Influence	$\mathrm{CI}_\ell(v) = (d(v)-1) \cdot \sum_{u \in \partial B(v,\ell)} (d(u)-1)$	Optimal percolation influence at distance ℓ; identifies micro-regions whose removal would most disrupt network connectivity (Morone and Makse, 2015).
LOB	Lobby Index	Largest l s.t. v has $\geq l$ neighbours with degree $\geq l$	Analogue of the h-index for node neighbourhood.

Abbr.	Name	Formula	Description
ECC	Eccentricity	$\mathrm{ecc}(v) = \max_{u \in V} d(v,u)$	Maximum distance to any node; nodes with minimum eccentricity form the centre.
COR	Coreness	Max k s.t. v belongs to the k-core	Nested shell structure; higher coreness means deeper embedding.
BURT	Burt's Constraint	$C_i = \sum_j \left(p_{ij} + \sum_{q \neq i,j} p_{iq}\, p_{qj} \right)^2$	Measures structural constraint—how much a micro-region's contacts are interconnected. Low values indicate brokerage potential (bridging otherwise disconnected parts of the mobility network) (Burt, 2003). $p_{ij} = w_{ij} / \sum_k w_{ik}$.
CRIT	Criticality	$\Delta\mathrm{GE}(v) = \mathrm{GE}(G) - \mathrm{GE}(G \setminus \{v\})$	Drop in global efficiency upon removal of v.
MC	Markov Centrality	$m(v) = \frac{1}{\frac{1}{n-1} \sum_{s \neq v} \mathrm{MFPT}(s,v)}$	Inverse of mean first-passage time to v in a random walk.

Abbr.	Name	Formula	Description
LAC	Local Lacunarity	$\Lambda(v,r) = \frac{\mathrm{Var}[M(v,r)]}{(\mathbb{E}[M(v,r)])^2} + 1$	Gap heterogeneity in the r-neighbourhood of v; $M(v,r)$ is the mass at scale r.
RC	Resilience Centrality	$\mathrm{RC}(v) = \frac{\Delta S(v)}{S(G)}$	Fractional change in the size of the giant component upon removal of v.
LPC	Laplacian Centrality	$C_L(v) = \frac{E_L(G) - E_L(G \setminus \{v\})}{E_L(G)}$	Fractional change in Laplacian energy $E_L = \sum_i \lambda_i^2(L)$ upon removal of v (Qi et al, 2012).

A.2 Network-Level Indicators: Detailed Definitions

The notation follows Appendix A.1. centralization indices are computed using the Freeman centralization formula (Freeman, 1979): for a node-level centrality X,

$$C_X(G) = \frac{\sum_{i=1}^{n}[X_{\max} - X(i)]}{\max_G \sum_{i=1}^{n}[X_{\max} - X(i)]}, \tag{9}$$

where $X_{\max} = \max_{i \in V} X(i)$ and the denominator is the theoretical maximum (achieved in a star graph for most centrality measures).

Group 1: Degree & Strength Centralization

Abbr.	Name	Formula	Description
SZI	In-strength centralization	Eq. (9) with $X = s^{\text{in}}$	centralization of incoming flow.
SZO	Out-strength centralization	Eq. (9) with $X = s^{\text{out}}$	centralization of outgoing flow.
SZA	Strength centralization	Eq. (9) with $X = s$	centralization of total flow.
DZI	In-degree centralization	Eq. (9) with $X = d^{\text{in}}$	centralization of in-degree.
DZO	Out-degree centralization	Eq. (9) with $X = d^{\text{out}}$	centralization of out-degree.
DZ	Degree centralization	Eq. (9) with $X = d$	centralization of total degree.

Group 2: Path & Distance-based Centralization

Abbr.	Name	Formula	Description
BZ	Betweenness centralization	Eq. (9) with $X = B$	Concentration of bridging importance.
CZ	Closeness centralization	Eq. (9) with $X = C_C$	Concentration of reachability.
HZ	Harmonic centralization	Eq. (9) with $X = H$	Concentration of harmonic centrality.

Group 3: Prestige Centralization

Abbr.	Name	Formula	Description
EZ	Eigenvector centralization	Eq. (9) with $X = \text{EC}$	Concentration of eigenvector importance.
PRZ	PageRank centralization	Eq. (9) with $X = \text{PR}$	Concentration of PageRank.
AZ	Alpha centralization	Eq. (9) with $X = \text{AC}$	Concentration of alpha centrality.
PZ	Power centralization	Eq. (9) with $X = \text{PC}$	Concentration of Bonacich power centrality.

Additional notation used below: $m = \frac{1}{2}\sum_{ij} w_{ij}$ is the total edge weight (used in modularity definitions); c_i denotes the community label of node i in a given partition;

$\delta(\cdot,\cdot)$ is the Kronecker delta; $S(q/n)$ is the fraction of nodes in the largest connected component after q nodes have been removed; $E_{>k}$ and $N_{>k}$ denote, respectively, the number of edges among and the number of nodes with degree greater than k; $\langle d\rangle = \frac{1}{n}\sum_i d(i)$ and $\langle d^2\rangle = \frac{1}{n}\sum_i d(i)^2$ are the first and second moments of the degree distribution.

Group 4: Community & Modularity

Abbr.	Name	Formula	Description
LeM	Leiden Modularity	$Q = \frac{1}{2m}\sum_{ij}\left[w_{ij} - \frac{s_i s_j}{2m}\right]\delta(c_i, c_j)$	Modularity under Leiden community partition (Traag et al, 2019). $m = \frac{1}{2}\sum_{ij} w_{ij}$; c_i is the community label of node i.
IM	Infomap Modularity	$L(\mathcal{M}) = q_{\curvearrowright} H(\mathcal{Q}) + \sum_{i=1}^{m} p^i_{\circlearrowright} H(\mathcal{P}^i)$	Map equation modularity: minimum description length of a random walk (Rosvall and Bergstrom, 2008).
SGM	Spin-Glass Modularity	$\mathcal{H}(\sigma) = -\sum_{ij} J_{ij}\delta(\sigma_i, \sigma_j)$	Spin-glass Hamiltonian formulation of modularity (Reichardt and Bornholdt, 2006).

Group 5: Connectivity & Efficiency

Abbr.	Name	Formula	Description
Arcs	Number of Arcs	$\lvert E\rvert$	Total directed edges with positive weight.
Dens	Density	$\frac{\lvert E\rvert}{n(n-1)}$	Fraction of possible directed edges present.
Diam	Diameter	$\max_{u,v} d(u,v)$	Longest shortest path in the network.
AVPL	Avg. Path Length	$\bar{d} = \frac{1}{n(n-1)}\sum_{u\neq v} d(u,v)$	Mean shortest-path distance.
ALE	Avg. Local Efficiency	$\bar{E}_{\text{loc}} = \frac{1}{n}\sum_v E_{\text{loc}}(v)$	Mean of local efficiencies across all nodes (Latora and Marchiori, 2001).
GLE	Global Efficiency	$E_{\text{glob}} = \frac{1}{n(n-1)}\sum_{u\neq v}\frac{1}{d(u,v)}$	Inverse-distance global efficiency (Latora and Marchiori, 2001).

Group 6: Edge & Higher-order Structure

Abbr.	Name	Formula	Description
BEZ	Edge Betweenness centralization	Eq. (9) with $X = B_e$	centralization of edge betweenness $B_e(i,j) = \sum_{s\neq t}\frac{\sigma_{st}(e)}{\sigma_{st}}$ (Girvan and Newman, 2002).
Mot	Motifs	$\sum_k n_k(G)$	Count of directed triadic motifs (Milo et al, 2002).
Tra	Transitivity	$\frac{3\times\text{triangles}}{\text{connected triples}}$	Global clustering coefficient (Watts and Strogatz, 1998).

Group 7: Resilience & Robustness

Abbr.	Name	Formula	Description
RRes	Random Attack Resilience	$R_r = \frac{1}{n}\sum_{q=1}^{n} S(q/n)$	Area under the curve of giant-component size S as nodes are removed randomly (Schneider et al, 2011).
SRes	Systematic Attack Resilience	$R_s = \frac{1}{n}\sum_{q=1}^{n} S(q/n)$	Same as RRes but nodes removed in decreasing order of degree.
NRI	normalized Resilience Index	$\mathrm{NRI} = \frac{R_r - R_s}{R_r}$	Relative difference; high values indicate vulnerability to targeted attacks (Albert et al, 2000).

Group 8: Asymmetry & Hierarchy

Abbr.	Name	Formula	Description
VAs	Mean Vertex Asymmetry	$\bar{A} = \frac{1}{n}\sum_i \frac{s^{\mathrm{in}}(i) - s^{\mathrm{out}}(i)}{s^{\mathrm{in}}(i) + s^{\mathrm{out}}(i)}$	Average directional imbalance across all nodes.
RCC	Rich Club Coefficient	$\phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k}-1)}$	Edge density among nodes with degree $> k$ (Zhou and Mondragón, 2004; Colizza et al, 2006).

Group 9: Complexity, Fractal & Information-theoretic

Abbr.	Name	Formula	Description
FDIMALL	Fractal Dim. (full graph)	$d_B = -\lim_{r\to 0} \frac{\ln N_B(r)}{\ln r}$	Box-counting fractal dimension of the entire graph (Song et al, 2005). $N_B(r)$ is the minimum number of boxes of radius r.
FDIMGC	Fractal Dim. (giant comp.)	Same as FDIMALL, restricted to the giant component.	Fractal dimension of the largest connected component.
KIGC	Kirchhoff Index (GC)	$\mathrm{Kf}(G) = n\sum_{i=2}^{n} \frac{1}{\mu_i}$	Sum of effective resistances; μ_i are non-zero Laplacian eigenvalues (Klein and Randić, 1993).
NE	Network Entropy	$H(G) = -\sum_i p_i \ln p_i$	Shannon entropy of the normalized degree (or strength) sequence; $p_i = d(i)/\sum_j d(j)$ (Anand and Bianconi, 2009).
PCTGC	Percolation Threshold (GC)	$p_c = \frac{\langle d\rangle}{\langle d^2\rangle - \langle d\rangle}$	Critical occupation probability for the emergence of a giant component (Molloy–Reed criterion).
NLAC	Network Lacunarity	$\Lambda = \frac{\mathrm{Var}[M(r)]}{(\mathbb{E}[M(r)])^2} + 1$	Gap heterogeneity at scale r; high values indicate heterogeneous spatial distribution of connections.

A.3 Graphon Distance Metrics: Detailed Definitions

Let $W_1, W_2 : [0,1]^2 \to [0,1]$ be two graphon functions (not necessarily symmetric, to accommodate directed graphs). Let $\mathcal{S}$ denote the set of all measure-preserving bijections $\phi : [0,1] \to [0,1]$. For a graphon W and $\phi \in \mathcal{S}$, define the relabelled graphon $W^\phi(x,y) = W(\phi(x), \phi(y))$.

1. Cut distance ($\delta_\square$)

The cut norm of a function $U : [0,1]^2 \to \mathbb{R}$ is:

$$\|U\|_\square = \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} U(x,y)\, dx\, dy \right|, \tag{10}$$

where the supremum is over all measurable subsets S, T. The cut distance between W_1 and W_2 is:

$$\delta_\square(W_1, W_2) = \inf_{\phi \in \mathcal{S}} \|W_1 - W_2^\phi\|_\square. \tag{11}$$

Interpretation: $\delta_\square$ measures the maximum discrepancy in total flow between any two subsets of nodes, optimised over all possible relabellings. It captures global, large-scale structural differences and is insensitive to small, localised perturbations. It is the fundamental distance in the Lovász–Szegedy theory: convergence in cut distance is equivalent to convergence of all subgraph densities (Lovász and Szegedy, 2006). In a mobility context, a large cut distance indicates that the two networks differ in how they partition flow across macro-regions.

2. Spectral distance (δ_{spec})

Let $\{\lambda_k(W)\}_{k=1}^\infty$ be the eigenvalues of the integral operator $T_W : L^2[0,1] \to L^2[0,1]$ defined by $(T_W f)(x) = \int_0^1 W(x,y) f(y)\, dy$, ordered by decreasing absolute value. The spectral distance is:

$$\delta_{\text{spec}}(W_1, W_2) = \left(\sum_{k=1}^{\infty} (\lambda_k(W_1) - \lambda_k(W_2))^2 \right)^{1/2}. \tag{12}$$

In practice, the sum is truncated at rank K:

$$\hat{\delta}_{\text{spec}}(W_1, W_2) = \left(\sum_{k=1}^{K} (\hat{\lambda}_k(W_1) - \hat{\lambda}_k(W_2))^2 \right)^{1/2}. \tag{13}$$

Interpretation: The spectral distance compares the dominant structural modes (eigenvalues) of the two graphons. Since leading eigenvalues correspond to community structure, assortative mixing, and core–periphery patterns, δ_{spec} is particularly sensitive to changes in the rank structure and community organisation of the mobility network. It is computationally efficient and well-suited for networks with clear modular structure (Chung, 1997; von Luxburg, 2007).

3. Frobenius (L^2) distance (δ_F)

$$\delta_F(W_1, W_2) = \inf_{\phi \in \mathcal{S}} \left(\int_0^1 \int_0^1 (W_1(x,y) - W_2^{\phi}(x,y))^2 \, dx \, dy \right)^{1/2}. \tag{14}$$

Interpretation: The Frobenius distance measures the root-mean-square element-wise difference between the graphon functions, optimised over relabellings. Unlike the cut distance, which tests only against rectangular subsets, the Frobenius distance penalises all local differences equally. This makes it more sensitive to fine-grained weight changes across the entire network. In a mobility context, δ_F captures both global rearrangements and localised shifts in flow intensity.

4. Wasserstein distance (δ_W)

Let μ_W be the measure on $[0,1]$ induced by the degree function $d_W(x) = \int_0^1 W(x,y)\, dy$, and let $\Gamma(\mu_{W_1}, \mu_{W_2})$ be the set of couplings between μ_{W_1} and μ_{W_2}. The p-Wasserstein distance (we use $p = 2$) is:

$$\delta_W(W_1, W_2) = \left(\inf_{\gamma \in \Gamma} \int_{[0,1]^2} \|W_1(x,\cdot) - W_2(y,\cdot)\|_2^2 \, d\gamma(x,y) \right)^{1/2}. \tag{15}$$

Interpretation: The Wasserstein distance quantifies the minimum structural rearrangement needed to transform one graphon into another (Villani, 2009; Peyré and Cuturi, 2019). It respects the geometry of the graphon space and captures both changes in the shape of the degree distribution and rearrangements of connectivity patterns. Unlike the Frobenius distance, which treats all element-wise differences equally, the Wasserstein distance accounts for the *spatial extent* of rearrangements: moving flow between similar node types costs less than moving flow between dissimilar ones. In a mobility context, a large Wasserstein distance indicates that substantial reconfiguration of the application flow pattern would be needed to convert one year's network into another's—with changes spanning structurally distant parts of the mobility system.

5. Jensen–Shannon divergence (δ_{JS})

normalize graphons to obtain probability distributions: $\tilde{W}(x,y) = W(x,y) / \iint W$. Define the mixture $M = \frac{1}{2}(\tilde{W}_1 + \tilde{W}_2)$. The Jensen–Shannon divergence is:

$$\delta_{JS}(W_1, W_2) = \frac{1}{2}\mathrm{KL}(\tilde{W}_1 \| M) + \frac{1}{2}\mathrm{KL}(\tilde{W}_2 \| M), \tag{16}$$

where

$$\mathrm{KL}(P\|Q) = \int_0^1 \int_0^1 P(x,y) \ln \frac{P(x,y)}{Q(x,y)} \, dx \, dy. \tag{17}$$

Interpretation: The Jensen–Shannon divergence is a symmetrised, bounded information-theoretic measure of distributional dissimilarity. It is sensitive to the *shape* of the weight distribution rather than its absolute values. In a mobility context, δ_{JS} detects changes in the informational structure of the network—e.g., whether the distribution of flows has become more uniform or more concentrated—even when the total volume of applications has not changed.

6. Total Variation distance (δ_{TV})

Using the same normalized graphons:

$$\delta_{TV}(W_1, W_2) = \frac{1}{2} \int_0^1 \int_0^1 |\tilde{W}_1(x, y) - \tilde{W}_2(x, y)| \, dx \, dy. \tag{18}$$

Interpretation: The total variation distance measures the maximum absolute difference in probability mass that the two graphon distributions assign to any measurable event. It is the strongest standard distributional distance and is sensitive to structural outliers—isolated large changes in specific edges. In a mobility context, δ_{TV} is particularly useful for detecting sharp, localised structural shifts (e.g., the sudden emergence or disappearance of a major flow route) that might be diluted in the Frobenius or Jensen–Shannon distances.

A.4 Graphon-Based Community Detection

This subsection describes the algorithmic steps for deriving communities and inter-community flow pathways from the estimated graphon.

Input.

A step-function graphon $\hat{W}_G : [0, 1]^2 \to \mathbb{R}_{\geq 0}$ estimated from the directed, weighted mobility network $G = (V, E, w)$ with $n = |V|$ nodes (micro-regions), as described in Section 3.6.

Step 1: Block identification.

The step-function graphon partitions the unit interval $[0, 1]$ into K blocks $B_1, B_2, \ldots, B_K$. Each block B_k corresponds to a set of nodes (micro-regions) that share similar connectivity patterns. The number of blocks K is determined during the graphon estimation procedure (e.g., via the stochastic block model approximation or spectral clustering of the graphon matrix). Let $\mathcal{B} = \{B_1, \ldots, B_K\}$ denote the partition.

Step 2: Node-to-block assignment.

Each node $i \in V$ is assigned to the block B_k whose interval contains the node's position in the ordering π:

$$i \in B_k \quad \Leftrightarrow \quad \frac{\pi(i) - 1}{n} \in I_k, \tag{19}$$

where $I_k \subset [0, 1]$ is the interval corresponding to block B_k. The resulting partition $\{B_1, \ldots, B_K\}$ defines the *graphon-based communities*: groups of micro-regions that play structurally equivalent roles in the mobility network.

Step 3: Block interaction matrix.

The graphon values between blocks define the block interaction matrix $\mathbf{P} \in \mathbb{R}_{\geq 0}^{K \times K}$:

$$P_{kl} = \hat{W}_G(x, y), \quad x \in I_k, \; y \in I_l. \tag{20}$$

The diagonal entries P_{kk} represent the intensity of within-community application flows; the off-diagonal entries P_{kl} $(k \neq l)$ represent inter-community flow intensities.

Step 4: Flow pathway identification.

An inter-community *flow pathway group* is identified for each block pair (B_k, B_l) where $P_{kl} > 0$. This means that micro-regions in block B_k systematically send applications to micro-regions in block B_l (or vice versa, since the network is directed and $P_{kl} \neq P_{lk}$ in general). The set of flow pathway groups is:

$$\mathcal{F} = \{(B_k, B_l) : P_{kl} > 0,\ k, l \in \{1, \ldots, K\}\}. \tag{21}$$

Step 5: Hierarchical interpretation.

The block interaction matrix $\mathbf{P}$ reveals the hierarchical organisation of the mobility network. Blocks with high P_{kk} values represent self-contained regional clusters. Blocks that receive flows from many other blocks (high column sums in $\mathbf{P}$) correspond to major destination hubs. The directed nature of $\mathbf{P}$ distinguishes sender blocks from receiver blocks, enabling the identification of the mobility system's hierarchical tiers.

Distinction from modularity-based methods.

Modularity-based algorithms (Leiden, Infomap) group nodes that are *densely interconnected*. Graphon-based community detection groups nodes that are *structurally equivalent*—i.e., they have similar connection profiles across the entire network. Consequently:

- Two micro-regions can belong to the same graphon block without being directly connected, provided they connect to the same types of regions (e.g., two university cities that both attract applicants from similar rural micro-regions).
- Inter-block flow pathways are explicitly identified, providing information about *how* communities interact—not just *which* nodes belong together.
- The number of blocks may differ substantially from the number of modularity-based communities.

Glossary

AC Alpha Centrality. 15
ALE Average Local Efficiency. 11
Arcs Number of Arcs. 11
AUT Authority Centrality. 15
AVPL Average Path Length. 11
AZ Alpha Centralization. 11, 24, 41, 43

BC Betweenness Centrality. 14
BEZ Edge Betweenness Centralization. 11, 24
BURT Burt's Constraint. 17, 37
BZ Betweenness Centralization. 11

CC Closeness Centrality. 14

CI Collective Influence. 17
CLR Cluster Rank Value. 16
COR Coreness. 17, 37
CRIT Criticality Value. 17, 37
CZ Closeness Centralization. 11

DCI Indegree Centrality. 14
DCO Outdegree Centrality. 14
Dens Density. 11
Diam Diameter. 11
DMNC Density of Maximum Neighborhood Component. 16
DZI Indegree Centralization. 28

EC Eigenvector Centrality. 15
ECC Eccentricity Value. 17, 37
EZ Eigenvector Centralization. 11, 41

FDIMALL Network Box Counting Fractal Dimension for the Entire Graph. 12, 28, 41
FDIMGC FDIMGC Network Box Counting Fractal Dimension for the Giant Component. 12, 25, 28, 41

GLE Global Efficiency. 11, 41

HBS Hubness Centrality. 15
HC Harmonic Centrality. 14
HZ Harmonic Centralization. 11

IM Infomap's Modularity Value. 11, 28

KIGC Kirchoff Index for the Giant Component. 12, 40
KNN K-th Nearest Neighbourhood. 15

LAC Local Lacunarity. 18
LC Leverage Centrality. 16, 37
LCC Local Community Centrality. 16
LE Local Efficiency. 16
LeM Leiden's Graph Modularity Value. 11, 25
LINC Lin Centrality. 14
LOB Lobby Index. 17, 37
LPC Laplacian Centrality. 18
LR Leader Rank in a Directed Graph. 16

MC Markov Centrality. 17
Mot Motifs. 11

NE Network Entropy. 12, 28
NLAC Networks' Lacunarity. 12
NRI Normalized Resilience Index. 11, 28, 40, 41

PC Power Centrality. 15
PCTGC Percolation Treshold for the Giant Component. 12, 28
PRC PageRank Centrality. 15
PRZ PageRank Centralization. 11, 28, 41
PZ Power Centralization. 11, 24

RC Resilience Centrality. 18
RCC Rich Club Coefficient. 12, 28
RRes Resilience for Random Attacks. 11, 28, 41

SCA Strength Centraliy. 12
SCI Instrength Centrality. 12, 25
SCO Outstrength Centrality. 12, 25
SCR S-core Value. 16
SGM Spin-Glass Modularity Value. 11
SLC Semi Local Centrality. 16
SRes Resilience for Systematic Attacks. 11, 28
SZA Strength Centralization. 9, 28, 41
SZI Instrength Centralization. 9
SZO Outstrength Centralization. 9

Tra Transitivity. 11

VAs Mean Value of Vertex Asymmetries. 12
VR Vote Rank. 16, 17, 37