

RESEARCH

Open Access

Network analysis of internal migration in Croatia



Dino Pitoski* , Thomas J. Lampoltshammer and Peter Parycek

*Correspondence:
dino.pitoski@donau-uni.ac.at
Department
for E-Governance
and Administration,
Danube University Krems,
Dr.-Karl-Dorrek-Straße 30,
3500 Krems an der Donau,
AT, Austria

Abstract

Migration, and urbanization as its consequence, is among the most intricate political and scientific topics, predicted to have huge effects on human lives in the near future. Thus being said, previous works have mainly focused on international migration, and the research on internal migration outside of the US is scarce, and in the case of Europe—the ubiquitous center of migration affairs—only in its infancy. Observing migration between settlements, especially using network analysis indicators and models, can help to explain and predict migration, as well as urbanization originating from internal migration. We therefore conducted a network analysis of internal migration in Croatia, providing insights into the size of internal migration in population, and relative sizes between intra-settlement migration, inter-settlement migration and population. Through centrality analysis, we provide insights into hierarchy of importance, especially, in terms of the overall flow and overall attractiveness of particular settlements in the network. The analysis of the network structure reveals high presence of reciprocity and thus the importance of internal migration to urbanization, as well as the systematic abandonment of large cities in the east of the country. The application of three different community detection algorithms provides insights for the policy domain in terms of the compatibility of the current country administrative subdivision schemes and the subdivision implied by migration patterns. For network scholars, the analysis at hand reveals the status quo in applied network analysis to migration, the works published, the measures used, and potential metrics outside those applied which may be used to better explain and predict the intricate phenomenon of human migration.

Keywords: Complex network, Human migration, Internal migration, Croatia

Introduction

Migration is a topic ever-present in political and media spheres, having effects to virtually all aspects of our daily living; economic, demographic, cultural, and generally societal. International migration, although a relatively small phenomenon as measured by the share of international migrants in total human population (about 3%, [1]), occupies a large share of scientific and political space. Virtually all discussions ongoing in media about migration are actually on its international component, which discussions spill over to the domain of politics, and which extend to the domain of science through various policy-ensured sponsorship and funds. Many migration scholars will agree that research has given international migration priority over

internal migration [2, 3]. Yet, internal migration in the World, based on the several credible estimates [4, 5], is about three times greater than international migration; people migrate mainly between settlements within their own country of residence.

Apart from the fact that internal migration is being largely ignored as consequential for designing global, regional and national policies, the topic of urbanization, which is a direct consequence of international and internal movements, is also discussed separately from these two. Urbanization is a matter that is given almost equal urgency as migration in political circles (although not so much in media), as the share of people living in urban areas is projected to reach 70% in just next two decades [6].

Both migration, and urbanization as its consequence, have up to date received very little interest from network scientists. Essentially a social network constituted on origin-destination re-locations, whose scale-free structure is almost obvious (the “migrated-get-more-migrated”), has only started being observed from a complex networks perspective at the beginning of the current decade. However, this interest, as measured by the number of studies observing human migration as a complex network, appears to be rapidly rising, with the number of studies dedicated to migration networks almost tripling by decade’s end (see section “[Related work](#)” for more details).

What is missing from these novel works are more analyses on internal migration, using reliable data on the exact micro (human settlement-to-human settlement) movements. Previous network studies, in line with the described imbalance of interest, were mainly performed on international (inter-country) migration, using data on migrant stocks from origin countries in destination countries (available from [7]). These data are heavily estimated using various approaches to fill in census gaps, and provide only a rough insight into worldwide migration (for the exact estimation methods and explanations on data inaccuracies see [8]). Only a few studies have been done on internal migration with exact inter-settlement movements, and these have covered almost exclusively the United States [9–11], along with a couple of studies investigating China inter-city migration [12, 13]. Only very recently, a study on the Austrian intra- and inter- settlement migration was added to the set of internal migration network analyses [14]. Prior to the latter, no European country, to our knowledge, has been investigated for migration patterns from a complex networks perspective, although accurate and fine-grained data on migration are often immediately accessible through countries’ statistics offices. Moreover, as the Europe approaches towards the unified migration measurement systems through various agenda [15, 16], comparing, or even merging the analyses on different countries to understand migration, mobility, and urbanization within this international union, seems highly feasible.

With this study, we extend the existing body of observations to another EU country—Croatia, and hopefully initialize a series of subsequent network analyses for the European space, which are based on true numbers for migration at the most fine-grained level in terms of locational units (the human settlement: cities, towns and villages). The analysis provides insights on the yet non-investigated territories as well as topics (migration and urbanization) from the perspective of network science. The analysis validates the network analysis methods in this specific application, at the same time offering valuable insights about the country’s migration to national policymakers.

In the next section, we address the previous network studies on migration. This is followed by the description of the deployed migration data and the abstraction and definition of the Croatian internal migration network from this data. The analysis is partitioned into application of centrality measures, measures of network-structure, and network (link weight) estimation models. We conclude by recapitulating the main results, outlining the benefits of our analysis especially for policymakers, and pointing towards paths opened for further research.

Related work

As we mentioned in the introduction, network studies on human migration were generally very rarely performed, especially those on internal migration, using fine-grained data of migration between human settlements. Dividing the previous research in terms of geographies, where migration have been investigated as a complex network, most commonly researched is migration at the global international level (the global inter-country migration), in the works by [17–26]. The next most frequently covered geography is the US (inter-county US migration), analysed in the works by [9–11]. Additional geographies investigated are China for its inter-city migration, [12, 13], EU for its inter-country migration [27], UK for its inter-district migration [28], and Mexico for its inter-city migration [29]. Inter-settlement migration has only recently been investigated using a network-science approach on Austrian internal migration [14].

The most intensely investigated network features across the aforementioned works have been countries' or cities' centralities, followed by their clustering and community formations. A thorough overview of the network measures applied across these related works, including a feasibility analysis in terms of their application to migration, is available in [30]. The feasible measures and models traced across these related works are deployed in this article.

The findings from the former studies reveal that for the international (inter-country) migration, there is a steady increase in the small-world effect, high network clustering, and a robust community structure [20, 21, 25]). Developed countries attract migration from an increasingly diverse set of origin countries, while migration inter-connectivity is increasing in general, both internationally and internally in observed countries [22, 23]. In the US, a steady diversity of destinations is noted, but also a steady increase in the average size of migration between “best-friend” connections [10]. Migration is found to be more intense within states, while a greater variation in destination counties has been noted for inter-state migration [9]. In China, the small-world effect and clustering have aggravated concerning urban agglomerations [12, 13]. For the rest of the geographies and more detailed insights into all of these findings, we refer to the before-cited literature. The existing literature has limited use in terms of comparisons with the investigated case at hand (the case of Croatia), primarily because up to only very recently, there was virtually no comparable inter-settlement analysis, especially concerning the European region. There are additional reasons which relate to methodological congruence of network metrics, where weighted and binary network abstractions are contested. These are also covered in the aforementioned methodology review [30]. In the analysis that follows, comparisons are made mainly with established facts from migration theory or the original works in which the applied network metrics or models have been established.

The following analysis also directly serves to compare the results with existing analyses of Austrian internal migration flows [14].

Network data and definitions

Data for internal migration in Croatia for 2018 were obtained from the Croatian Bureau of Statistics upon request, and are accessible in the Supplementary Material [31]. The data contain accounts of the official address changes from one city to the other or the same city/municipality within this period. Exact definitions of migration, and the administrative subdivision of Croatia, can be found in [32]. The primary abstraction of the Croatian internal migration network from data is a weighted loop di-graph $\mathcal{G} = (\mathcal{N}, \mathcal{L}, \mathcal{W})$, whose:

1. nodes $\mathcal{N} = \{n_1, n_2, \dots, n_N\}$ are the country's second-level administrative units (cities and municipalities, $N=556$),
2. link weights $\mathcal{W} = \{w_{ij}\}_{N \times N}$, $i, j = 1, \dots, N$, where i can be equal to j , are the counts of official changes of address of residence from city or municipality i to city or municipality j in the year,
3. links $\mathcal{L} = \{l_{ij}\}_{N \times N}$ is a binary projection of \mathcal{W} , such that $l_{ij} = 1$ if $w_{ij} > 0$, and $l_{ij} = 0$ if $w_{ij} = 0$.

In the primary abstraction, self-loops are taken into account ($w_{ii} \geq 0$). From \mathcal{G} , we further identify a subgraph $\mathcal{G}' = (\mathcal{N}, \mathcal{L}', \mathcal{W}')$, where $\mathcal{W}' = \mathcal{W} \setminus \{w_{ii}\}$ and \mathcal{L}' is the according binary projection of \mathcal{W}' .

It is important to note that the cities and municipalities (nodes), for which migration data are reported, are not the lowest-level hierarchical units in terms of actual spread of population. Croatia reportedly, as of 2008, has 6.749 human settlements [33]. In the second-level administrative subdivision (cities and municipalities), the smaller settlements in close geographical proximity are contracted to the level of cities or municipalities (see regulation on categorization in [34]), at which level migration, as well as population statistics, are aggregated.

An issue related to this aggregation of data, consequential particularly for the analysis of loop (intra-city or intra-municipal) migration, is that for cities and municipalities, which officially constitute of one (same named) settlement, the intra-settlement migrations have not been recorded. To clarify this by example: while inter-migration between seventy settlements fitted under the largest City of Zagreb is recorded, no migration is recorded for the third largest city in Croatia, Rijeka, as it does not officially consist of more than this one settlement (Rijeka). For these, and several other cities and municipalities, even though in reality these are constituted of smaller settlements with migrations certainly occurring between them, since being differently administratively divided, intra-city or intra-municipal migrations are omitted.

Due to these data gaps, in the subsequent analysis we will not be able to infer that the conclusions from an analysis of \mathcal{G}' can be generalised for the entire network \mathcal{G} , as proposed in [30]. We will nevertheless provide as many as possible facts on intra-city/intra-municipal movements, based on the official data from Croatian Bureau of Statistics, as well as some estimations on this component of internal migration. For simplicity of

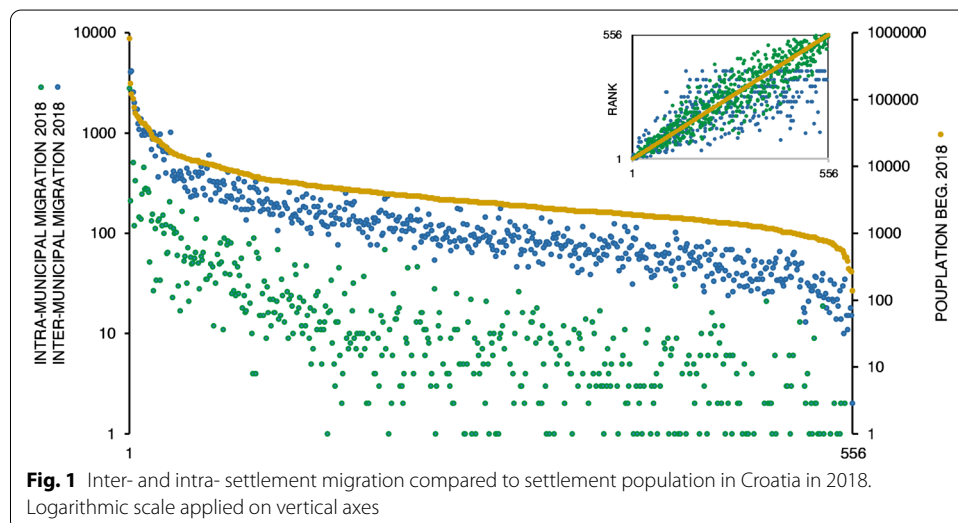
language, when discussing inter- or intra-locational movement, instead of the term city/municipality, henceforth we will use terms “settlement” (hence, not inter-city or inter-municipal, but inter-settlement migration).

Analysis

The size of the phenomenon of migration, reported by the Croatian Bureau of Statistics [35], calculated as the share of total migrant population within the total population in Croatia, is relatively small; about 1.75%. This share might, however, be substantially larger, given the previously addressed gaps in data on intra-settlement migration. In Fig. 1, we present the migration-population relation per settlement, using plots of the aggregated values for inter-settlement migration, intra-settlement migration, and the population for each settlement.

Aggregated per settlement, values for inter-settlement migration are essentially node strength (weighted degree) [36] values. The values for intra-settlement migration are missing for 117 settlements, among these some very large cities, such as Rijeka (3% of Croatia’s population), Pula, and Slavonski Brod (where each of these accounts for about 1.3% of Croatia’s population). Fitting the power-law curve of intra-settlement migrations, when ranked by the population (as in Fig. 1), we obtain very high values for the before-mentioned large cities; more than thousand migrations within Rijeka, and more than 300 migrations within Pula/Slavonski Brod. The total sum of fitted values is close to 19.000, substantially larger than the actually reported 14.238 intra-settlement migrations (19.85% of the reported total migration). The sum of such estimated intra-settlement migration and real inter-settlement migration returns a share of total migration within the population of about 1.86%, which is still much smaller than the global estimations on internal and international migration (about 11% and 3%, respectively, see [1, 4]). One possible explanation for this is the severe outflow of people from Croatia to other countries in the region [37], thus, a higher share of international in comparison to internal migration.

In the subsequent analysis, we proceed to analyse only the inter-settlement migration subnetwork (\mathcal{G}'), for which data are reliable and complete. To allow for easier reading,



we will refer to this subnetwork as, just, “network”. Due to gaps in data on intra-settlement migrations, it will remain uncertain, whether the findings from that application are generalizable to the whole network, but our suggestion after making the above-discussed estimations is that they are. One fact to support this suggestion is that the (Pearson) correlation between the actually reported (non-missing) values for intra-settlement migrations and node strength in \mathcal{G}' is 0.96. Also, both of these categories are strongly correlated with the size of population ($\min \rho = 0.97$).

Centralities

Node strength and degree

In Fig. 2, we depict Croatian inter-settlement migration by means of a network graph, in which we label most of the nodes incident to links with $w_i \geq 50$ migrations. Notable is the interaction with the capital. Node strength (s_i), essentially the total throughput of internal migrants per settlement, is represented by node size. For reference, note that $\max s_i' = s_{Zagreb}' = 15.799$, second largest $s_{Split}' = 4.077$, third largest $s_{Rijeka}' = 4.077$ and $\min s_i = 2$ people. The complete node strength ranking is provided in the Supplementary Material [31].

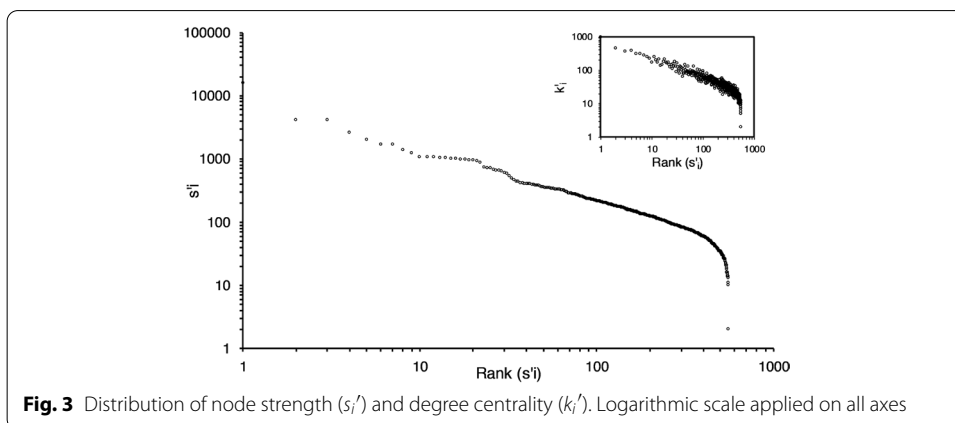
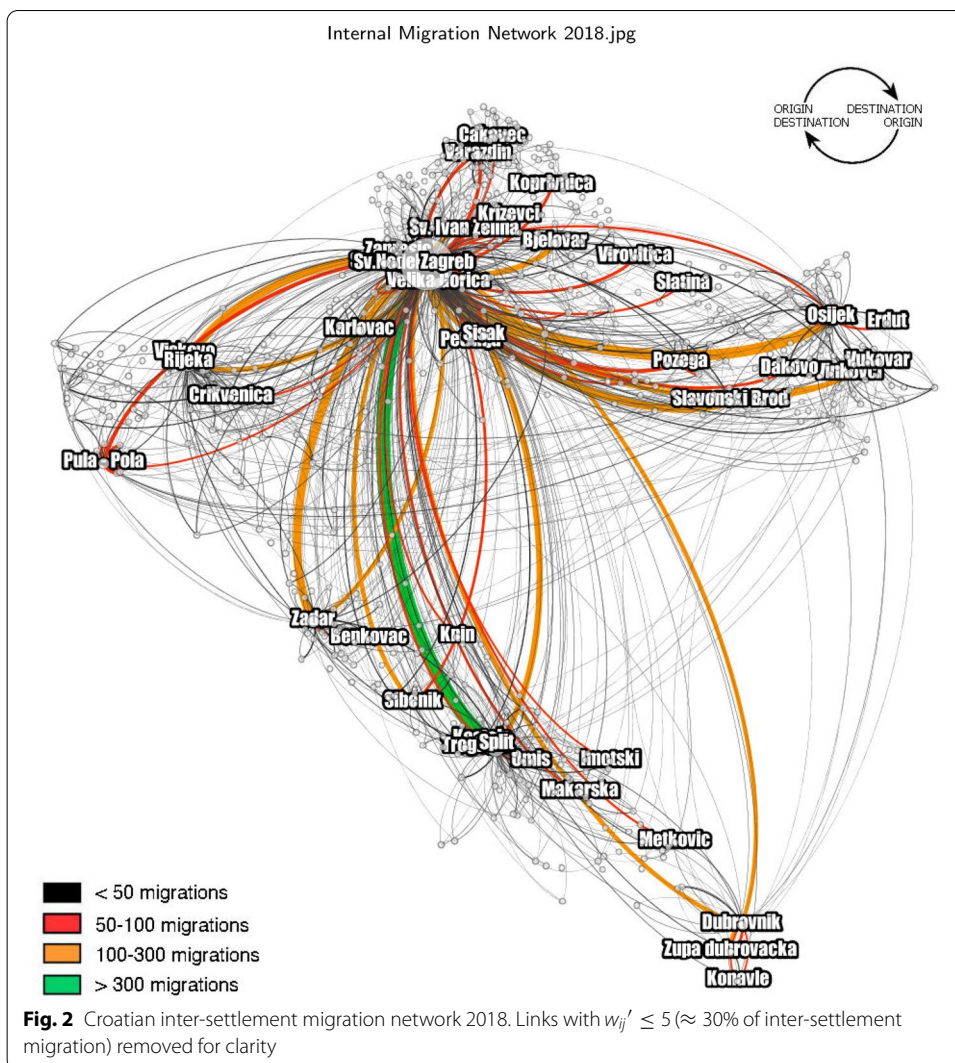
In Fig. 3, we report on the node strength and degree centrality (k_i') distribution. Both distributions fit a power law with exponents $\alpha_s = 2.29$ and $\alpha_k = 2.72$. Strength and degree values correlate, but not as strongly as may be expected; $\rho(s_i', k_i') \approx 0.89$. The size of out- and in-migration from/to a particular settlement is in line with the diversity of connecting locations, but not to a full extent. This manifestation is stronger for in-migration ($\rho(s_i^{in'}, k_i^{in}') \approx 0.85$) than for out-migration ($\rho(s_i^{out'}, k_i^{out}') \approx 0.90$).

By calculating the ratio of node strength and degree centrality ranks, simultaneously observing differentials in ranking of the two categories, one can identify the cities and municipalities that send or receive from a disproportional number of locations. Examples are the cities of Kastav and Sv. Nedelja (Zagreb County), which both send and receive from many different settlements, although, when ranked by the throughput of migrants (node strength), or by the size of population, these settlements do not result on top ranks. From this point of identification, policymakers and urban geographers can trace reasons why these settlements attract more people. In the case of the above-mentioned cities, the reason is quite evidently the large number of employment opportunities (companies) existing in these locations, also very close to the largest urban areas. The rest of exceptions can be evaluated using the Supplementary Material [31].

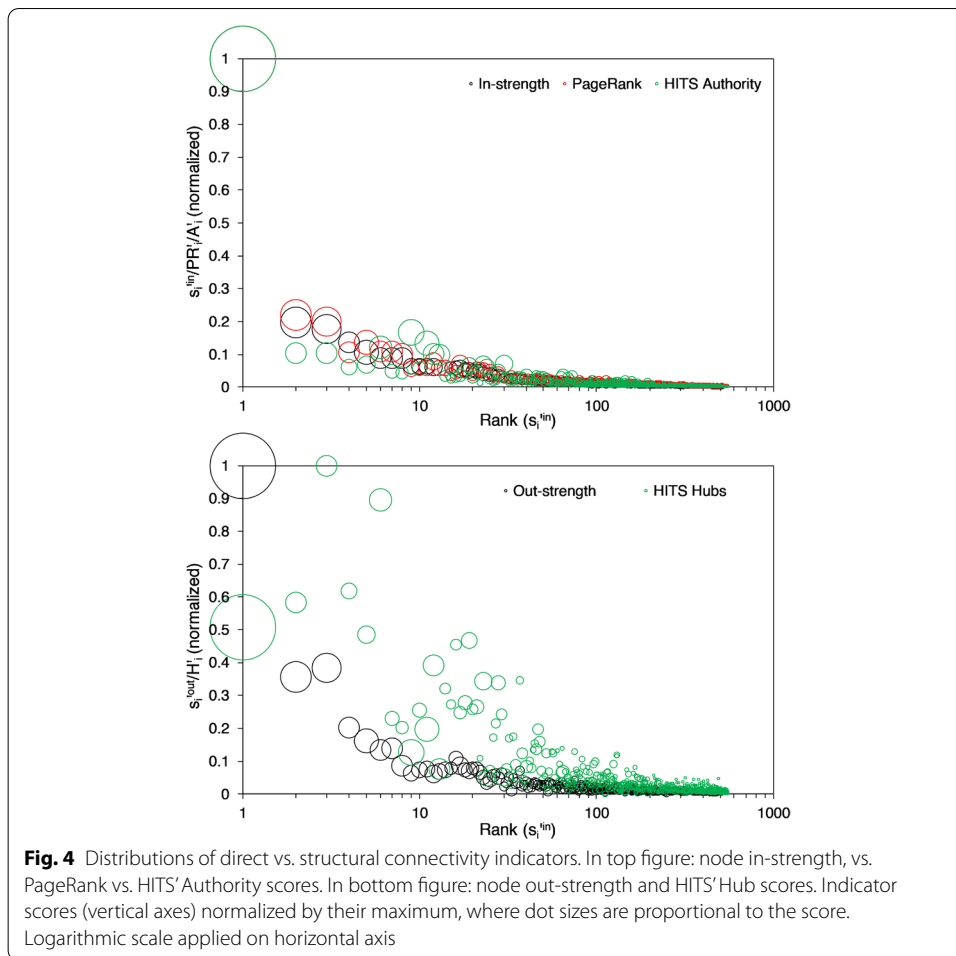
Outward and inward centrality scores are strongly correlated in both their weighted and binary versions ($\rho(s_i^{in'}, s_i^{out}') \approx 0.97$, $\rho(k_i^{in'}, k_i^{out}') \approx 0.94$), which, in addition to what can be observed from Fig. 2, points to a high reciprocity in the network. A thorough reciprocity analysis will follow as part of the analysis on network structure.

Hubs and authorities

The sensitivities entailing the application of eigencentality algorithms for the case of migration is elaborated in [30]. The core issue is that there is no firm structure (no travel constraints) in the network, while migration flow sinks in the directly connected destinations (in other words, there are no network paths). The calculation of eigencentality indicators performed on the Croatian migration network can serve



only as alternative node influence evaluation, which rewards settlements that send/receive to/from better connected settlements. Hence, the probabilistic interpretation, which is the basis for e.g. PageRank [38], where PageRank values can be interpreted as probability that a “random migrant” will land to a particular settlement as joining



the migration network at any moment in time, must be taken with reservations. In the Supplementary Material [31], we provide values for this ubiquitous algorithm, as well as for the alternative HITS [39]. All calculations are made on the weighted directed network (using weight-adjusted versions of the indicators), as suggested in [30].

In Fig. 4, we show the distributions for the comparable categories: on the one side node in-strength, PageRank, and HITS' authority score, and, on the other side, node out-strength and HITS' hub score. The first is used to provide intuition on authoritativeness, or attractiveness of particular settlements, in terms of exclusively direct migration (in-strength), as well as structural migration (PageRank, HITS). The second is a representation of repulsive potential of particular settlements, where HITS hub score addresses the tendency of each settlements to send migrants to authoritative locations. Distributions of authority settlements (top) are well aligned, with some differences in HITS' authority evaluations which are due to the dichotomous nature of the indicator.

What can be grasped from the above top chart is the strong correlation between all comparable authority scores ($\min \rho(PR_i', A_i') \approx 0.97$), which points to a clear hierarchy in terms of authoritativeness of different locations, which is in line with strength distributions. In the bottom chart (to be used for comparisons with the top only), the discussed high correlation of in- and out-strength is visualized, but also an inconclusive

relationship between authority and hub scores derived from the application of HITS algorithm ($\rho(A_i', H_i') \approx 0.46$). This is in line with reciprocity announced in the network graph visualization, where hubs return a lot of migration to authorities and vice versa. A thorough analysis of reciprocity and other structural metrics follows.

Network structure

Reciprocity, assortativity and transitivity

One of the clear features of Croatian inter-settlement migration network is reciprocity. Overall weighted reciprocity, as adopted from [40]:

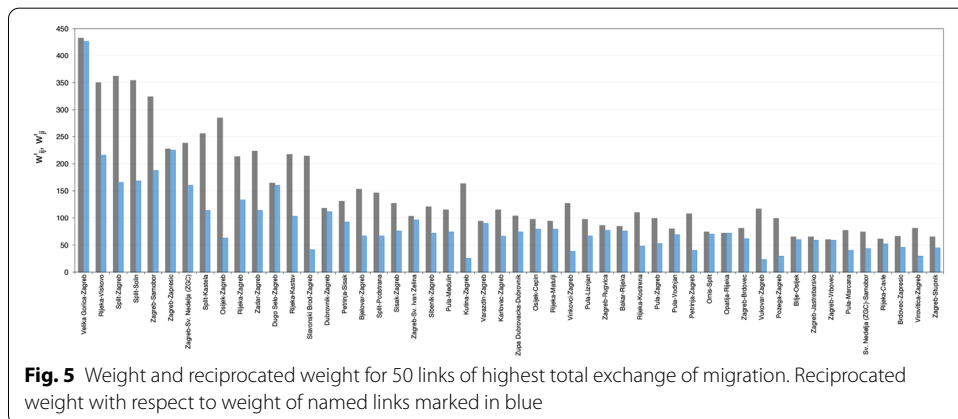
$$r = \frac{\sum_i \sum_{j \neq i} w_{ij}^{\leftrightarrow'}}{\sum_i \sum_{j \neq i} w_{ij}'} \tag{1}$$

where $w_{ij}^{\leftrightarrow'}$ marks the fully reciprocated weight component ($w_{ij}^{\leftrightarrow'} = w_{ji}^{\leftrightarrow'} = \min[w_{ij}', w_{ji}']$), is calculated at 0.49; structurally, about a half of inter-settlement migration in Croatia is actually the inter-settlement migration exchange.

Correlation between weights and reciprocal weights $\rho(w_{ij}', w_{ji}')$ is very strong, at about 0.91, which vouches for the consistency of reciprocation across the whole network. However, the average of the reciprocated component in total migration activity occurring between any two settlements is not particularly high (≈ 0.9 is the average share of reciprocated component in total per-link activity, or ≈ 0.16 , the average share of reciprocated component in the maximum one-way flow between any two settlements).

A more detailed inspection of the per-link reciprocity may provide the insight into those origin settlements that receive a much lower number of migrants then the number of migrants they send to the destination. Figure 5 shows the top 50 links with respect to the total migration exchange along these links, where one can spot the significantly lower return of people from the Capital to the large Croatian cities such as Split, Osijek, Slavonski Brod, Kutina, Vinkovci, and Vukovar. Network visualization in Fig. 2 may also be useful for perceiving such imbalances. The traditionally agriculture-oriented macro-region of Slavonia on the east, seems to be particularly inflicted by a low return of people to its major cities.

Importantly, the total differential of weight of migration in favour to the more populated areas, in terms of the end-2017 population as reported by the Croatian Bureau of



Statistics, is 15900 people (on average 2.55 persons), against 9771 people (on average 2.19 persons) migrating in favour to less populated areas, this excluding considerations on the relative sizes of the two exchanging settlements. Overall, the share of migration that went from less populated to more populated places in 2018 amounts to 55.3%, again, excluding the size relativity of migrating settlements. The numbers suggest, roughly, that internal migration's contribution to urbanization in Croatia, when looked at independently from other relevant components such as natural birth and international migration, is about 5%. This rough estimation needs to be updated for the actual adjacent settlements' population sizes, which will not be a part of this analysis, but the interested readers can inspect our reciprocity data, which includes population sizes, provided in the Supplementary Material [31].

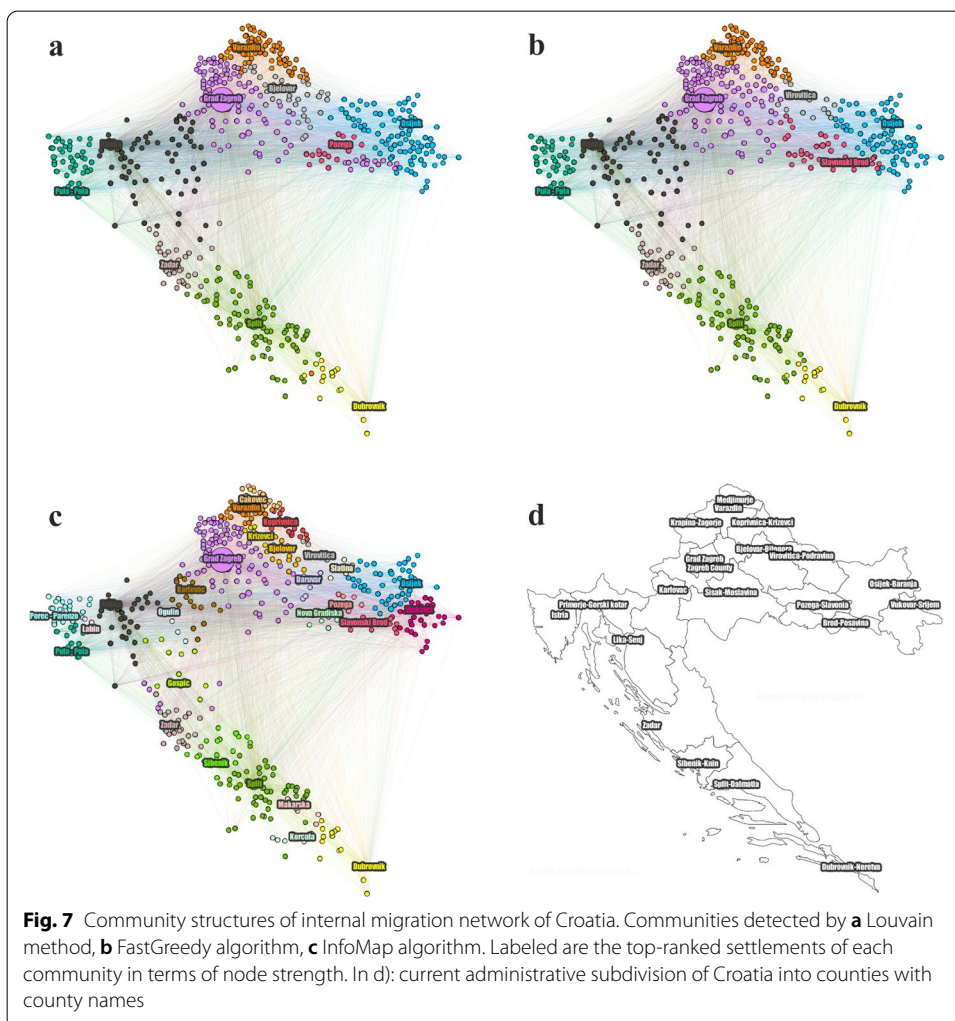
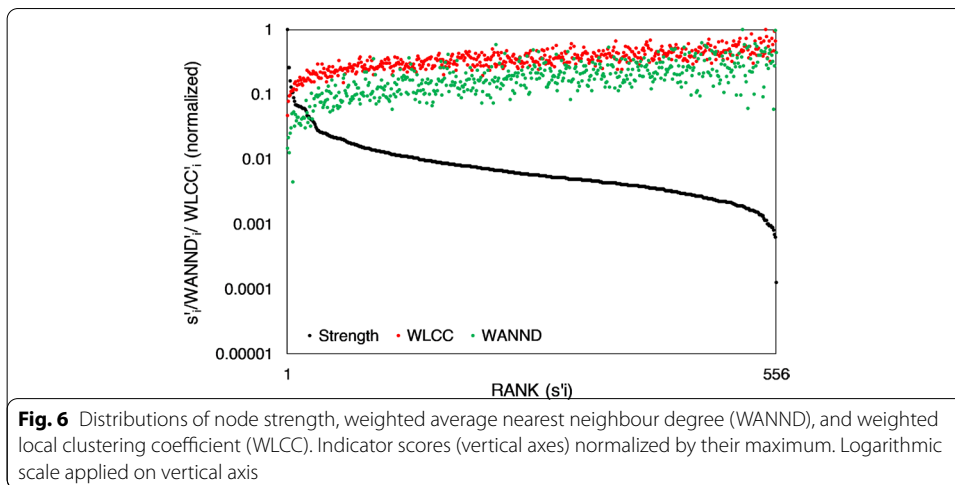
High reciprocity in the network has an effect to the calculations of assortativity and transitivity. In Fig. 6, we put in relation the scores for the weighted average nearest neighbour degree (WANND) and weighted local clustering coefficient (WLCC) [41], with node strength. Relation of WLCC and strength values suggests a substantial disassortative behaviour in the network, although the global weighted assortativity coefficient [42] calculated at -0.08 suggests only slight disassortativity. The interfering effect of reciprocity is also present when it comes to transitivity, the distribution of weighted clustering coefficients being almost uniform, while the global weighted transitivity is measured at 0.22. This later indicator has been assessed as theoretically of low interpretability in migration-as-network applications [30], due to, effectively, non-existence of network paths, but now we see an additional dimension that adds to this circumstance.

Modularity and random migrant clustering

For a detailed view into possible communities in the Croatian inter-settlement migration network, we applied three different algorithms; Louvain [43], FastGreedy [44] and Infomap [45]. We used the three variants to obtain the lowest bias possible, considering the ambiguity of the interpretation of community detection algorithms as addressed in [46].

The first two, based on modularity optimization [47], return very small differences in cluster membership allocations, and show a fairly modular structure ($Q \approx 0.32$ in both cases) of 10 communities. InfoMap, based on random walk optimization, produces a more fine-grained map of 27 communities. All three allocations are visualized in Fig. 7, and the exact membership allocations are provided in the Supplementary Material [31].

The community structure produced by these two algorithms is useful for comparing with the current administrative subdivision of the country (counties), and to understand what the aspect of migration implies when it comes to changing the administrative subdivision of Croatia. The possible changing of this subdivision into larger regions is one important issue, which is ongoing in political and public debates (see [48]). The information is, hence, a viable input for decision-making, especially considering the consistent output of the modularity-based algorithms that submit a more coarse grain picture of inter-migrating communities.



Network models

Forecasting of migration and urbanization is possible using various means. If data for several years in the past would be available, we could derive trend projections for the growth or decline of each individual settlements' node strength for the next year(s), and adjust for the same type of projection internal migration as a whole. More demanding forecasting models, which require a lot more additional data, may involve the here-covered structural algorithms, such as PageRank, amalgamated with statistical models based on migration driver variables (see [49] for the list of potential drivers). Here we provide two established models, where the second is the extension of the first, and which have been shown very efficient, while based only on two predictor variables - population and distance, for which predictors the data are, also, largely accessible.

Gravity law model

First, we test if the gravity law holds [50], with weights' estimation equation 2,

$$\hat{w}_{ij}' = k p_i^\alpha p_j^\beta d_{ij}^{-\gamma} \quad (2)$$

where p_i and p_j are the populations of origin and destination settlements, respectively, d_{ij} is the great circle distance from the origin to destination, and k , α , β and γ are adjustable parameters. Strongest correlation between real and estimated values has been obtained for $k \approx 0.23$, $\alpha \approx 0.35$ and $\gamma \approx 0.82$. In Fig. 8, we provide the plotted distributions of gravity model-estimated weights and real weights, and the exact values are available in the Supplementary material [31].

Radiation model

The second weight estimation model which we use is the extension to the gravity law model proposed by [51], and contained in equation 3

$$\hat{w}_{ij}' = M_i \frac{p_i p_j}{(p_i + s_{ij})(m_i + p_j + s_{ij})} \quad (3)$$

where

$$M_i = p_i \frac{\sum_j w_{ij}'}{\sum_i p_i} \quad (4)$$

Notations in the equations are the same as in gravity model (p for populations, d for distance), while s_{ij} denotes the total population in the circle of radius centred at i and tangent to j , excluding these origin and destination settlements' populations. Same as for the gravity model, we provide the plotted distributions of radiation model-estimated weights and real weights in Fig. 8, while the exact values we make available in the Supplementary material [31].

Comparing the two models, we can see that radiation model, applied to the case of Croatian inter-settlement migration, proves to be more efficient and balanced in its predictions, in line with suggestions put forward by [51] that were based on their analysis of internal US mobilities. The explanatory power increases substantially, from R^2 or ≈ 0.39 in the case of the gravity model to R^2 of ≈ 0.52 in the case of the

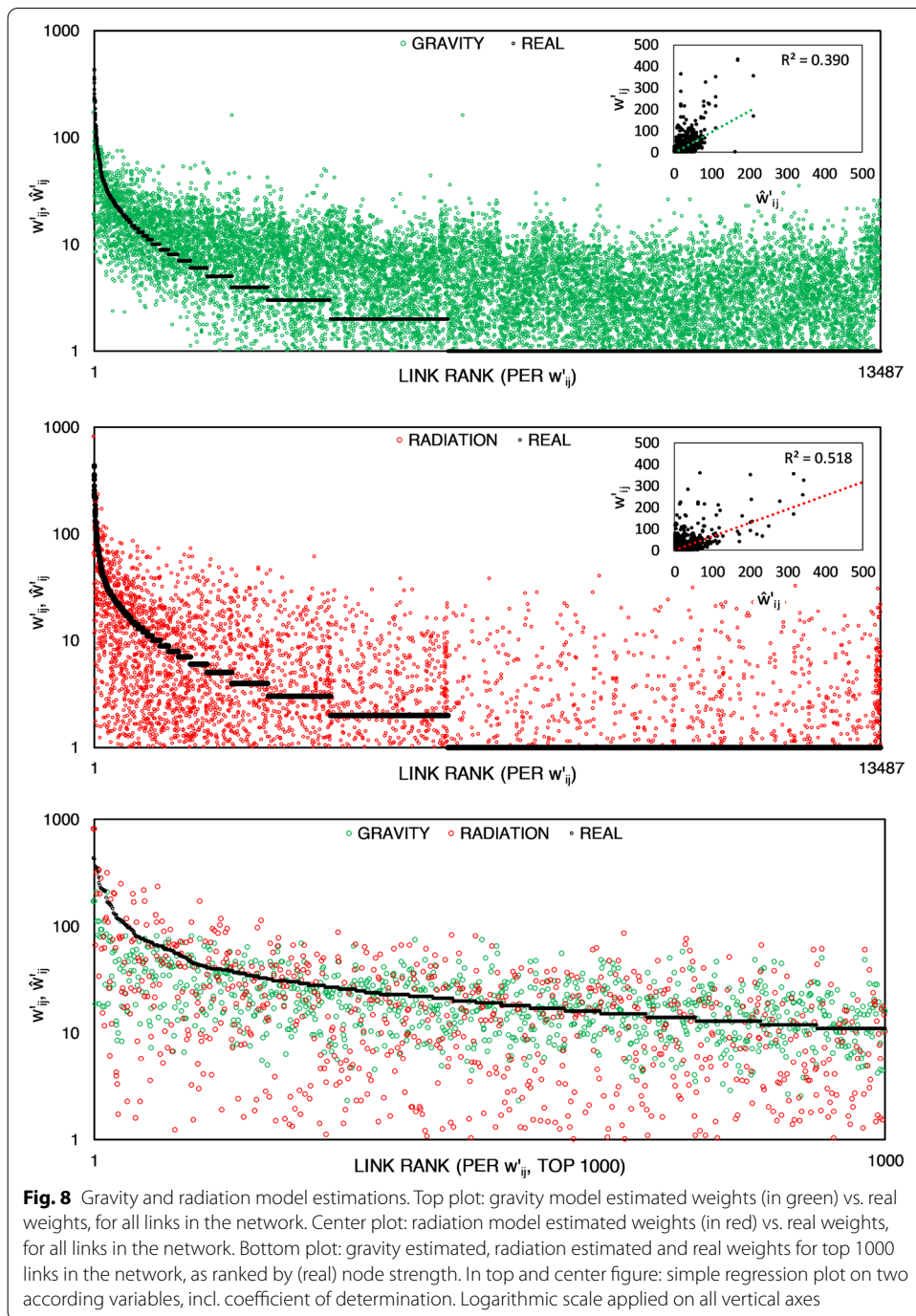


Fig. 8 Gravity and radiation model estimations. Top plot: gravity model estimated weights (in green) vs. real weights, for all links in the network. Center plot: radiation model estimated weights (in red) vs. real weights, for all links in the network. Bottom plot: gravity estimated, radiation estimated and real weights for top 1000 links in the network, as ranked by (real) node strength. In top and center figure: simple regression plot on two according variables, incl. coefficient of determination. Logarithmic scale applied on all vertical axes

radiation model. We also see results of the radiation model which are much more in line with the natural power law scaling. The gravity model, on the other hand, seems to structurally overestimate the actually low migrated links. However, in a close-up view of the estimation of the top links in the network (Fig. 8, bottom chart), we see that radiation model is quite more scattered than the gravity model.

Although overall, the radiation model outperforms the gravity model in terms of weight determination in this concrete case, our analysis suggests that the gravity

model still is a viable tool for migration forecasting. Moreover, our analysis shows that both models should be subject to improvement, especially by adding more predictors and exploring additional modelling concepts within the spatial accessibility domain, as, in their nascent form, these prevalently deployed models are able to explain only about a half of variance of real migration.

Findings and policy implications

This section briefly recapitulates the findings of our analyses, where we address, in particular, the implications for the migration social policy.

The size of internal migration in Croatia is relatively small (about 1.75%), but the data particularly for intra- settlement migrations (as complementary to inter-settlement migration), due to the current administrative subdivision, creates bias in this measurement. The policymaker should continue to work towards reconstructing the administrative subdivision, or particularize migration data collection as to include more precise human-settlement allocation.

The policymaker is herewith informed about the node centrality and network structure with different network indicators that are recommended for policy use as an upgrade to the current migration data capacities. Node centrality indicators expose the current and potential supreme migration attractors, repulsors, and outliers in terms of correlation with these settlement populations. These current and potential rankings are exposed in respectively the node strength and eigencentrality rankings alongside the settlement population rankings. Using these rankings, the policymaker can immediately spot the settlements with migration performance that lies in the extremes, and, with the assistance of, e.g. social geographers, outline the migration factors that lead to these extremes. From particular cases, migration policymakers can get closer to obtaining an overall picture on the causes of migration that should be affected, where outcomes on migration are available in this study and its supplementary material.

The node structural indicators also make easier conclusions on the migration factors working in and around the traced, strongly integrated, inter-settlement communities, as well as reveal the (non)accordance of the current administrative subdivision and the migration tendencies of Croatian citizens. The most important finding is on the high weighted reciprocity in the network, which has huge consequences on how migration is approached from the modelling/forecasting point of view. Reciprocity analysis enables the policymakers to understand migration's contribution to urbanization, where having specific flow and counterflow the policymaker can seek specific reasons for higher discrepancies.

Finally, the policymaker is informed about the effectiveness of current models for link weight estimation, essentially models for prediction of migration. In that regard, Radiation model manifests much better performance than the commonly used Gravity Law model and the policymaker is advised (how) to switch. The advances from Gravity Law to Radiation model, along with the exposure of the specific network structure characterized by high reciprocity, points towards the need of more sophisticated modelling approaches, the ones essentially based on the concept of spatial accessibility.

Conclusion

Migration, and urbanization as its consequence, are among the most intricate political and scientific topics, predicted to have huge effects on human lives in the near future. Although essentially a social phenomenon of evidently scale free structure, migration has been analysed in very few complex network studies. Networks of within-country migration received even lower interest, although internal migration is about three times more intense than international, and although data for internal migration are accessible, and show the exact re-locations between human settlements. Observing migration between settlements, especially using network analysis indicators and models, helps us to explain and predict migration, as well as urbanization coming from internal migration.

Our analysis of the network of internal migration in Croatia provides insights into the feasibility of network analysis in general, as well as into the value of these results as inputs for political decision-making. We provide insights into size of internal migration in population, and relative sizes between intra-settlement migration, inter-settlement migration and population. Through centrality analysis we provide insights into hierarchy of importance, especially, in terms of overall flow and overall attractiveness of particular settlements in the network. Although size and authority of settlements covaries with the size of population, centrality analysis identifies outliers, which is useful for the eventual analysis of particular migration drivers that may explain these anomalies. The analysis of network structure reveals high presence of reciprocity. Through reciprocity a valuable insight on internal migration's contribution to urbanization gets revealed, as well as the systematic abandonment of large cities in the country's east. High reciprocity is seen to have an impact on the application of other structural indicators, particularly assortativity and transitivity. The application of three different community detection algorithms provides insights for the policy domain in terms of the compatibility of the current country administrative subdivision schemes and the subdivision implied by migration patterns. Link weight prediction models show good approximation quality overall, and the effectiveness of radiation model over the gravity law model is clearly exposed.

These results offer both the policymakers and the scientists the insight into the structure of internal migration in this European country, which is a good starting point of subsequent addition of analyses of other countries in the region, and eventually understanding migration in the region as a whole. An analysis of one more European country, especially one which has substantial international exchange with the currently investigated (e.g. the recent analysis of Austria), may be used for a comparative analysis to mathematically code the regularities, and to furthermore test propositions on regularities occurring on the international scale. Furthermore, a temporal analysis is one crucial dimension that vouches that the results obtained here are valid long-term. Through the temporal analysis of especially the most prominent characteristic of reciprocity, the true insights on migration contribution to urbanization will be possible, as well as projections of migration and urbanization in the future. For network scholars, the analysis at hand reveals the status quo in applied network analysis to migration, the works published, the measures used, and potential metrics outside those applied which may be used to better explain and predict the intricate phenomenon of human migration.

Acknowledgements

Authors would like to express their special thanks to the Croatian Bureau of Statistics for providing the source data on Croatian internal migrations, required to perform this study. Special thanks also goes to the researchers at the Center for

Advanced Studies Southeast Europe (CAS SEE) at University of Rijeka, Croatia, for their valuable inputs over discussing our research.

Authors' contributions

DP and TJL conceptualised the paper and the methodology, DP carried out the analyses, DP and TJL drafted the initial manuscript. DP, TJL, PP revised the draft, PP was responsible for senior supervision and funding acquisition. All authors have read and agreed to the published version of the manuscript. Regarding the sequence of authors, the 'sequence-determines-credit' (SDC) approach was applied. All authors read and approved the final manuscript.

Funding

This study was co-financed by the Asylum, Migration and Integration Fund and the Austrian Federal Ministry of the Interior, Grant no. BMI-IA123004/0046-III/5/a/2016.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the Figshare repository, <https://doi.org/10.6084/m9.figshare.12497177>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 22 June 2020 Accepted: 21 February 2021

Published online: 04 March 2021

References

- De Haas H. Human migration: myths, hysteria and facts. Oxford: Wolfson College Oxford Plans and Prospects; 2014. p. 4–6.
- Otoi A, Titan E, Dumitrescu R. Internal and international migration: is a dichotomous approach justified? *Procedia Soc Behav Sci*. 2014;109:1011–5.
- Bell M, Muhidin S. Cross-National comparison of internal migration. *Human Development Research Paper (HDRP) Series*. 2009;30(2009).
- United Nations. Dept. of Economic and Social Affairs, Population Division: World Population Prospects: The 2002 Revision. Analytical report. United Nations; 2004.
- Bell M, Blake M, Boyle P, Duke-Williams O, Rees P, Stillwell J, Hugo G. Cross-national comparison of internal migration: issues and measures. *J R Stat Soc Ser A Stat Soc*. 2002;165(3):435–64.
- United Nations Department of Economic and Social Affairs: World urbanization prospects: The 2018 revision, key facts. United Nations; 2018.
- The World Bank: Migration and Remittances Data. <https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data>.
- Özden Ç, Parsons CR, Schiff M, Walmsley TL. Where on earth is everybody? The evolution of global bilateral migration 1960–2000. *World Bank Econ Rev*. 2011;25(1):12–56.
- Charyyev B, Gunes MH. Complex network of united states migration. *Comput Soc Netw*. 2019;6(1):28.
- Liu X, Hollister R, Andris C. Wealthy hubs and poor chains: constellations in the U.S. urban migration system. *Adv Geogr Inf Sci*. 2018;201509:73–86.
- Xu Z. The structure and dynamics of population migration among economic areas in the united states from 1990 to 2011. *Pap Reg Sci*. 2018;97(3):785–800.
- Li F, Feng Z, Li P, You Z. Measuring directional urban spatial interaction in china: a migration perspective. *PLoS ONE*. 2017;12(1):0171107.
- Sun Y, Pan K. Prediction of the intercity migration of Chinese graduates. *J Stat Mech*. 2014;2014(12):12022.
- Pitoski D, Lampoltshammer TJ, Parycek P. Network analysis of internal migration in Austria. *Digital Govern Res Pract*. 2021; <https://doi.org/10.1145/3447539>.
- Piper N. Migration and the sdgs. *Glob Soc Policy*. 2017;17(2):231–8.
- European Commission: Urban Agenda for the EU. <https://ec.europa.eu/futurium/en/urban-agenda>.
- Aleskerov FT, Meshcheryakova N, Rezyapova A, Shvydun S. Network analysis of international migration. *SSRN Electronic Journal*. 2016.
- Cerqueti R, Clemente GP, Grassi R. A network-based measure of the socio-economic roots of the migration flows. *Soc Indic Res*. 2019;146(1):187–204.
- Danchev V, Porter MA. Neither global nor local: heterogeneous connectivity in spatial network structures of world migration. *Social Netw*. 2018;53:4–19.
- Davis KF, D'Odorico P, Laio F, Ridolfi L. Global spatio-temporal patterns in human migration: a complex network perspective. *PLoS ONE*. 2013;8(1):53723.
- Fagiolo G, Mastroiello M. International migration network: topology and modeling. *Phys Rev E Stat Nonlinear Soft Matter Phys*. 2013;88(1):012812.
- Peres M, Xu H, Wu G. Community evolution in international migration top1 networks. *PLoS ONE*. 2016;11(2):0148615.
- Porat I, Penguigui L, Others: World migration degree. In: ERSA conference papers. No. Ers14p60 in regional development and globalisation: best practices; 2014. p. 1–19.
- Leone Scialabolazza V. A net of moving people: network analysis of international migration flows. *Networks of International Trade and Investment*. 2018;249–296.

25. Tranos E, Gheasi M, Nijkamp P. International migration: a global complex network. *Environ Plann B Plann Des*. 2015;42(1):4–22.
26. Windzio M. The network of global migration 1990–2013. *Soc Netw*. 2018;53:20–9.
27. Balaz V, Karasová K. Geographical patterns in the Intra-European migration before and after eastern enlargement: the connectivity approach. *Ekon Cas*. 2017;65(1):3.
28. Askar D, House T. Complex patterns of multiscale human mobility in united kingdom. University of Warwick Working Paper; 2010.
29. Caudillo-Cos C, Tapia-McClung R. Patterns of internal migration of mexican highly qualified population through network analysis. *Comput Sci Appl*. 2014;2014:169–84.
30. Pitoski D, Lampoltshammer TJ, Parycek P. Human migration as a complex network: appropriate abstraction and the feasibility of Network Science tools. In: Haber P, Lampoltshammer TJ, Mayr M, Plankensteiner K, editors. *Data science—analytics and applications*. Wiesbaden: Springer Vieweg; 2021. https://doi.org/10.1007/978-3-658-32182-6_17.
31. Pitoski D, Lampoltshammer TJ, Parycek P. Network analysis of internal migration in Croatia—supplementary material. figshare. Dataset; 2020. <https://doi.org/10.6084/m9.figshare.12497177>.
32. Croatian Bureau of Statistics: Population—Migration. Notes on methodology. <https://www.dzs.hr/Eng/DBHomepages/Population/Migrations/methodology.htm>.
33. Croatian Bureau of Statistics: Statistical information; 2008. https://www.dzs.hr/Hrv_Eng/StatInfo/pdf/StatInfo2008.pdf.
34. Official Gazette of the Republic of Croatia: Law on the Territories of Counties, Cities and Municipalities in the Republic of Croatia. https://narodne-novine.nn.hr/clanci/sluzbeni/1997_01_10_151.html.
35. Croatian Bureau of Statistics: First Release No. 7.1.2., Migration of population of Republic of Croatia, 2018. https://www.dzs.hr/Hrv_Eng/publication/2019/07-01-02_01_2019.htm.
36. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. *Proc Natl Acad Sci*. 2004;101(11):3747–52.
37. Croatian National Bank. Dynamics and determinants of migration—the case of Croatia and experience of new EU member states. <https://www.hnb.hr/documents/20182/2101832/24-dec-drazenovic-kunovac-pripuzic.pdf>.
38. Brin S, Page L. The anatomy of a large-scale hypertextual web search engine. *Comput Netw ISDN Syst*. 1998;30:107–17.
39. Kleinberg JM. Authoritative sources in a hyperlinked environment. *J ACM*. 2011;46:604–32.
40. Squartini T, Picciolo F, Ruzzenenti F, Garlaschelli D. Reciprocity of weighted networks. *Sci Rep*. 2013;3:2729.
41. Barrat A, Barthélemy M, Vespignani A. Weighted evolving networks: coupling topology and weight dynamics. *Phys Rev Lett*. 2004;92(22):228701.
42. Newman MEJ. Analysis of weighted networks. *Phys Rev E Stat Nonlinear Soft Matter Phys*. 2004;70(5 Pt 2):056131.
43. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech*. 2008;2008(10):10008.
44. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Phys Rev E Stat Nonlinear Soft Matter Phys*. 2004;70(6 Pt 2):066111.
45. Rosvall M, Axelsson D, Bergstrom CT. The map equation. *Eur Phys J Spec Top*. 2009;178(1):13–23.
46. Yang Z, Algesheimer R, Tessone CJ. A comparative analysis of community detection algorithms on artificial networks. *Sci Rep*. 2016;6:30750.
47. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlinear Soft Matter Phys*. 2004;69(2 Pt 2):026113.
48. Assembly of European Regions. Regionalisation in Croatia: established but uncertain. <https://aer.eu/regionalisation-croatia-established-uncertain/>.
49. Pitoski D, Lampoltshammer TJ, Parycek P. Drivers of human migration: a review of scientific evidence. *Soc Sci*. 2021;10(1):21.
50. Zipf GK. The P 1 P 2/d hypothesis: on the intercity movement of persons. *Am Sociol Rev*. 1946;11(6):677–86.
51. Simini F, González MC, Maritan A, Barabási A-L. A universal model for mobility and migration patterns. *Nature*. 2012;484(7392):96–100.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.