

Wealthy Hubs and Poor Chains: Constellations in the U.S. Urban Migration System

Xi Liu, Ransom Hollister, and Clio Andris

Abstract Flows of people connect cities into complex systems. Urban systems research focuses primarily on creating economic models that explain movement between cities (whether people, telecommunications, goods or money), and more recently, finding strongly and weakly-connected regions. However, geometrically graphing the dependency between cities within a large network may reveal the roles of small and peripheral city agents in the system to show which cities switch regions from year to year, which medium-sized cities serve as collectors for large cities, and how the network is configured when connected by wealthy or deprived agents.

We propose a network configuration method called ‘best friend’ networks, where a node attaches to one preferential node, so that edges = nodes = n . Our case study is 20 years of migrants, sourced from the U.S. Internal Revenue Service, traveling between U.S. cities. In our networks, an edge is created to link a city to its most popular migrant destination city for a given year. The resulting configurations reveal closely connected “constellations” of cities comprised of chains, trees, and hub-spoke structures that show how urban regions are configured. We also show routing behavior within these networks to reveal that high-income migrants tend to flock to hub cities, while low-income migrants form local city chains via nearby movements.

Keywords Migration • Urban hierarchy • Economic systems • Regional science • Spatial interaction • Complex systems

1 Introduction

In an urban hierarchy, larger cities are connected to smaller cities with medium size cities as intermediaries. Within this network, goods, information, capital, flights, migrants, commuters, etc. flow through planar and non-planar veins, providing cities with valuable resources. While cities are often studied in terms of demographics and

X. Liu • R. Hollister • C. Andris (✉)

Department of Geography, The Pennsylvania State University, University Park, PA, 16802, USA
e-mail: xiliu@psu.edu; maskedchicken@gmail.com; clio@psu.edu

© Springer International Publishing AG 2018

L. Perez et al. (eds.), *Agent-Based Models and Complexity Science in the Age of Geospatial Big Data*, Advances in Geographic Information Science,
DOI 10.1007/978-3-319-65993-0_6

production (a static representation), conceptualizing their position within an urban system such as the hierarchy (using a dynamic representation of in and out flow) allows researchers to examine the city within this larger corpus of transactions.

The urban hierarchy is comprised of groups of spatial regions where each is anchored by one very large city. This large anchor city (e.g. Chicago) exerts a gravitational pull on its surrounding cities unless another large city, perhaps Minneapolis, MN or St Louis, MO claims what would usually be Chicago's surrounding cities as part of their own functional regions. Cities that lie on a region's periphery or circumference are more likely to switch regions than those closer to the anchor city. In the past, regions were bound into geographically-cohesive areas in order to minimize the costly movement of natural resources and commodities [1]. When peripheral settlements send many flows to a city's central business district, this settlement is considered part of the larger city's functional region. In practice, flows such as migrants and commuters help the U.S. Office of Budget and Management define the spatial boundaries for Business Economic Areas (BEAs) and Metropolitan Statistical Areas (MSAs).

The traditional regional approach has been explained by the gravity model, which estimates interaction (i.e. flows, connection strength) between two places as the product of their respective populations divided by the square of the distance between the places [2]. The resulting estimate simulates the economic pull strength of cities, assuming that many people will choose to connect to a nearby place whose large size signifies many opportunities [3]. A gravity model using just population and distance has been shown to predict about 57% of U.S. inter-city migrant flows [4].

Regions are also delineated by areas of homogenous industry [5] or cohesive economic activity [6]. More recently, creative methods like dollar bill circulation [7], telephone calls [8], surname clustering [9] and maps of sports team popularity from Facebook likes [10] have been used to delineate regions around functional anchor cities.

Today, the regional hierarchy can be re-examined with a network approach. The economic transition from manufacturing to digital services and information technologies has allowed regions to form and function not just as a group of nearby cities, but as a network of connected cities that may or may not be proximal. This network is formed by "leapfrogging" (skipping over) nearby cities to create connections with distant cities that have economic benefit [11, 12]. These networked economies are not geometrically contiguous and thus, the connections are harder to predict in theory, but larger and more comprehensive data sets allow for the investigation of factors beyond traditional place-to-place connectivity (such as the gravity model) [13].

Here, we focus on descriptive properties of migration in the U.S. urban system. Migration choice has been explained by factors such as searching for the best job possible [14] seeking out a certain lifestyle [15], or capitalizing on social networks and interpersonal relationships [16]. Instead of building a model that attempts to correlate high migration volume with demographic or economic variables of different cities, we view the migration system as a network of cities connected

by volumes of migrants, as per the topic of migration systems theory (MST) [17]. Similar studies partition city systems into communities that are closely connected internally [18, 19], or search for network hub cities [20, 21]. These studies advance the use of network science in studies of the urban hierarchy, but do not address the extent to which migrants surpass near cities to connect with those further away—as they may when interpersonal relationships and institutions are involved.

We use county-to-county migrant flows sourced from the U.S. Internal Revenue Service (IRS) for 21 years (details in Sect. 2) to form a network of 917 cities (nodes) that connect to each other (edges) weighted by the number of migrants exchanged by cities (an undirected network). Because the complete network of migrants ties many cities together and we are interested in uncovering the urban hierarchy, we experiment with the following concepts:

1. *Best Friend*: Best friend networks are created by drawing an edge between an origin and the destination to which it most frequently sends migrants.
2. *Best High/Low Income Friends*: These networks differentiate high-income flows from low-income flows. The network is made from gathering each city's highest income outflow and connecting it with that destination. (i.e. an edge is made to the destination that attracts migrants with the highest average income). The same procedure is repeated for each city's lowest income destination.
3. *Constellations*: This method produces a collection of graphs (i.e. disconnected subgraphs of networks) of cities that due to the number of nodes involved in each graph, their configuration and their spatial genesis, resemble constellations which can be classified into motifs. We create a single 'galaxy' of constellations for each of 21 years.

Our results show that migration networks exhibit significant structural temporal persistence, and clear ensemble rules can be used to construct the networks. We find that some cities switch preferences to alternative large city anchors over time, and that some large city anchors become popular or decline in popularity. We also determine that low-income flows create different networks than high-income flows. We validate and contextualize these findings by comparing our model to the gravity model and radiation model. Our proposed networks can respond to the following questions: Which cities are popular for migrants? What regions (i.e. connected graph structures) arise? Which cities feed into larger cities? Which cities bypass closer and larger cities to connect directly to a more distant metropolis? Does a population hierarchy emerge? Are systems of cities closed or do they connect in larger chains? How do these patterns change for high- and low-income migrants?

In Sect. 2, we describe the migration dataset, network and analysis methods. In Sect. 3, we explore re-occurring constellations in the networks, compare our model to other prevailing models such as the gravity and radiation model, which reflect the structure of urban hierarchy. We conclude in Sect. 4.

2 Data and Methods

2.1 U.S. Migration Data and Population Data

We use data from the U.S. Internal Revenue Service (IRS) Statistics of Income Migration Data for years from 1992–1993 through 2012–2013 for this study. These data are free and available online. The original data were generated from the yearly change in address reported on individual tax returns from one year to the next, and aggregated at the county level to produce a network of county-to-county flows. Each flow contains three attributes: the number of returns, the number of exemptions, and the adjusted gross income (AGI), which is the sum of all income moving on the flow. Flows must contain at least ten returns to be reported in the dataset. We use number of exemptions to estimate the migrant population, as this value reflects the size of families, including children and jointly-filing spouses. Alternatively, using the number of filers would estimate the number of heads of households that migrate.

We aggregated the county-to-county flows into flows among Core Based Statistical Areas (CBSAs), formerly referred to as MSAs. CBSAs are defined as urban cores and peripheries with a population of at least 10,000 residents. Since CBSAs (henceforth, cities) follow county boundaries, aggregation required only flow summation. The aggregated data contains 917 cities reporting migration flows throughout the 21 years period. Each city is accompanied by a population count defined by the U.S. Census Bureau at the county level, as aggregated to the city level.

2.2 The Best Friend Configuration Model

The network is configured based on the *single allocation* [22] of edges to nodes (i.e. cities). In this configuration, a single city is only permitted to attach to the city to which it sends the highest proportion of its flows. For example, New York City is only attached to Miami because it sends more migrants to Miami than to any other city. In this model, each city is allowed only one outgoing connection (out-degree = 1), but the in-degree can be as large as the number of other nodes in the system ($n-1$). Thus, this network is referred to as the *best friend* network. Edges are assigned a weight (w) calculated as the proportion of migrants (m) city i sends to city j (Eq. 1):

$$w = \frac{m_{ij}}{\sum_1^k m_{ij}} \quad (1)$$

where k is the total number of cities to which city i is connected (i.e. its outgoing degree).

Over 21 years, the best friend model contains a total of 3.1% possible city-to-city edges, but accounts for 20.7% of total system-wide migrants. The average yearly migration flow magnitude ranges from 170–200 migrants and the average best friend magnitude ranges from 1100–1500 migrants. Over time, the total number of system-wide migrants grew from 5 million to 6.6 million and best friends accounted for 1 to 1.4 million migrants each year (hence, about 20% of total migration). Average AGI incomes range from \$120,000 to \$9000 per flow.

We also derive two special types of best friend models where edges are characterized by average income on the flow, calculated as the AGI of that flow divided by the number of returns on the flow. High and low-income migration networks are each created by selecting the best-high-income friend and best-low-income friend of a city, defined as the largest migration streams amongst the top 10% (high-income) and bottom 10% (low income) average income migration connections leaving city i . The top 10% is used rather than single the highest/lowest income flow to ensure a high number of migrants and thwart anomalies.

Temporally, each city in the income networks has an average of nine different best-high/low-income friends over the time period. On average, cities are connected to their best-high-income friends for 6.5 years and to their best-low-income friends for 5.4 years. Generally, wealthy best friend pairs are more stable over time.

2.3 Constellations

Constellations, or motifs [23], are basic graph structures that repeatedly appear in networks. In this study, constellation is a relaxed definition of motif that refers to families of basic structures that are widely seen in best friend networks, as compared to their probability of arising in a null models based on the gravity model. We detected five types of motifs (Fig. 2) in the best friend networks: pairs, chains, hubs, stars, and trees [24]. These graph structures are enumerated and analyzed using community detection methods within the R statistical computing environment's *igraph* package [25, 26]. Their definitions are as follows:

Pairs: A pair is formed by two cities that are each other's best friend. They are isolated from population hubs and may have strong dependency on each other.

Chains: A chain is a series of single directional connected cities where for $i = 1 \dots n$, city i points to city $i + 1$. Usually, city n connects to a local hub. Chains can reveal how a series of many migrants connect to nearby non-hub cities, possibly facilitated by a lack of social connections in large cities, poor mobility, or high levels of local social capital.

Hubs: A hub is a node with an in-degree larger than one with 'spoke' cities directly connected to it. The hub node may point to one of its spokes or to other hubs, creating stars and trees. Hubs are popular destinations for both chain and non-chain nodes.

Stars: A star is defined by hubs that point to one of their spokes. Small local hubs tend to form stars with proximal cities, and rely less on the influence of distant, larger hubs.

Trees: A tree is hub that connects to other hub nodes, and is typically attracted by higher-level hubs. Trees tend to connect small, medium and large cities.

2.4 Analytical Methods

Gravity model. The gravity model, as in [27], is a classical model for predicting flows based on population and distance, so that the magnitude of migrants T_{ij} between city i and j is estimated as:

$$T_{ij} = K \frac{P_i P_j}{d^\beta} \quad (2)$$

where P_i and P_j represents the population in city i and j , respectively, d is the distance between the two cities and K is a constant. β is a distance decay factor, often referred to as the coefficient of friction, and most commonly estimated with value of 2.

Radiation model. The radiation model [28] is used to predict flow volumes T_{ij} between city i (with population P_i) and j (with population P_j) as:

$$T_{ij} = T_i \cdot \frac{P_i P_j}{(P_i + s_{ij})(P_i + P_j + s_{ij})} \quad (3)$$

where T_i represents outflows from city i , and s_{ij} denotes the total population of alternative population centers within a given radius of the destination city.

Distance between two cities is calculated as the using Euclidean distance between each CBSA's geometric centroid.

3 Results

3.1 Best Friend Network

A series of best friend networks was created for each year (ex. Fig. 1). Most cities (60%) have no in-degree (degree = 1), 21% of cities have a degree of two (one outgoing flow, one incoming flow), 9% have a degree of three and 5% of cities have a degree of six or higher. These larger hubs include the U.S.'s ten largest cities, with Dallas consistently having the highest degree at over 20 best friend connections. The best friend network detects the regional importance of more geographically-

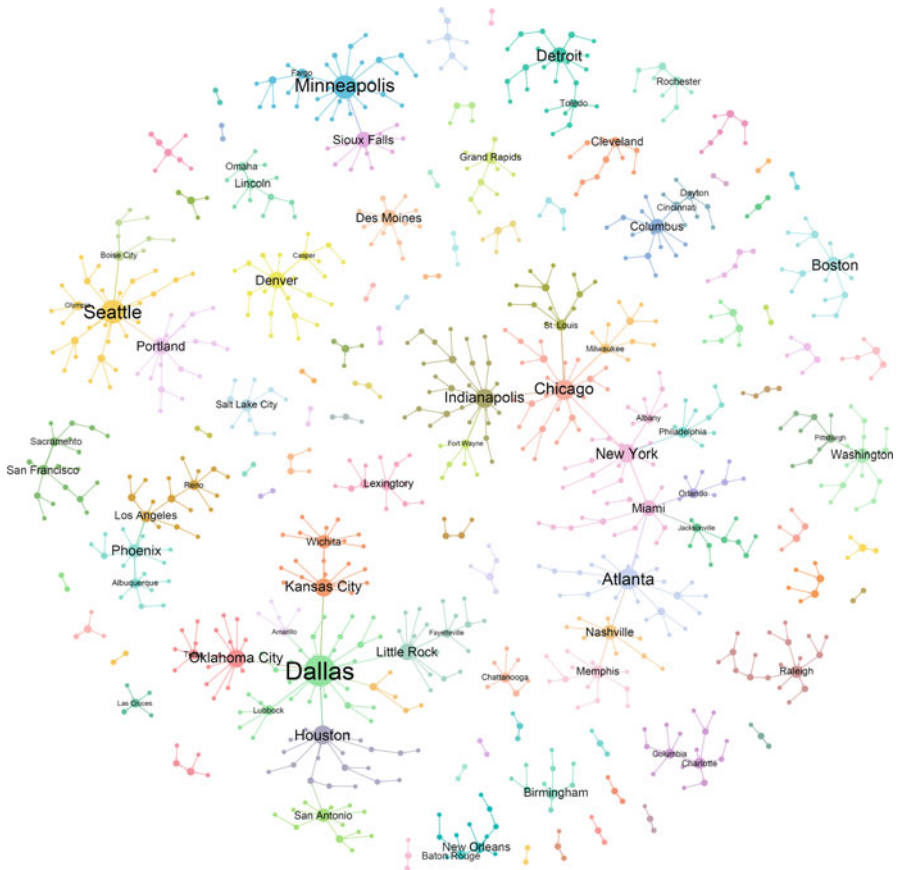


Fig. 1 Best friend network constellations for year 2012. In the network, the size of nodes corresponds to their degrees using Yifan-Hu's proportional method in Gephi [29] and each color represents a separate constellation. Some connected constellations are divided into different components due to their relatively weak connections, as determined by the community detection algorithm [25]

isolated cities such as Oklahoma City, Sioux Falls, Salt Lake City, Wichita, Des Moines, Memphis, Jackson, Grand Rapids, and Little Rock (Fig. 1) in their local hierarchical systems.

As migrant streams change each year, we can expect some fluctuation in the network. The average time spent with a best friend is 13.8 years. 504 of 917 cities (55%) have only one best friend for the entire period and 877 (96%) have at most three different best friends. On average, 110 cities change best friends each year, a turnover rate of 12% per year. Because the number of different best friends is low, this turnover rate does not compound at a high rate over longer time periods (e.g. 15% of cities have a different best friend in 2000 than in 2012). The strength of a best friendship, i.e. the percentage of migrants sent to a best friend

city (Eq. 1), ranges from 3.6% (Chicago to Los Angeles in 1992) to 100% for Mount Sterling, KY to Lexington, KY (1996–1999; 2001) and Big Stone Gap, VA to Kingsport-Bristol, TN-VA (2006). In general, cities Chicago, Columbus, OH and Atlanta have the smallest percentages of migrants sent to their best friend cities.

Crucially, we do not see an increase in the diversity of places to which a city sends its migrants. We had hypothesized that the rise of the Internet and mobile technologies in the late 1990s would promote more swirling/churn in the preferences of the migrants, given the new opportunities to research potential destinations. With more diverse information, migrants may have experienced other places, i.e. travelled more, and garnered friends in multiple locales. Yet, our analysis does not reveal a diversification of movement over multiple destinations at any point during this time. In fact, we see a steady increase in the average percentage of migrants a city sends to its best friend, starting at 0.35 in the early 1990s and rising to 0.37 in the 2010s.

The number of separate constellations and their size remains relatively stable over time. For each year, there was an average of 105 constellations, each containing from 2 to 66 cities, with an average size of 8.8 cities. The majority of constellations are small clusters, with 80% of the constellations comprised of fewer than 13 cities and 49% comprised of fewer than 5 cities (Fig. 1). Most constellations are geographically compact (averaging about 130 km), driven in part by small constellations, especially mutual best friend pairs which limits the average geographic spread. Notable exceptions include the strong New York City-Miami connection and any constellation connecting Alaska or Hawaii to the mainland, as well as some recurring connections between major cities. The New York City-Miami connection is driven in part by retirees from New York City choosing to move to a warmer climate (Fig. 1). These migrants are colloquially known as “snowbirds”.

We next categorize individual nodes based on the following observed motifs: pairs (both isolated pairs and those pairs within larger constellations), hubs (star and non-star hubs), spokes (nodes directly connect to hubs and with 0 in-degree), two categories of trees: {tree hubs (local hubs in a tree motif) and tree spokes (nodes directly connect to tree hubs and with 0 in-degree)}, and chains (all members in a chain category) (Fig. 2, Table 1). A node can only be placed into one category. The categorization is implemented with an algorithm that uses in/out degree, the in/out degree of their best friend, and the node type of their best friend as input parameters. We find that most constellations, especially large constellations, have single central nodes. Since large hubs have more resources, they may be less likely to rely on other hubs.

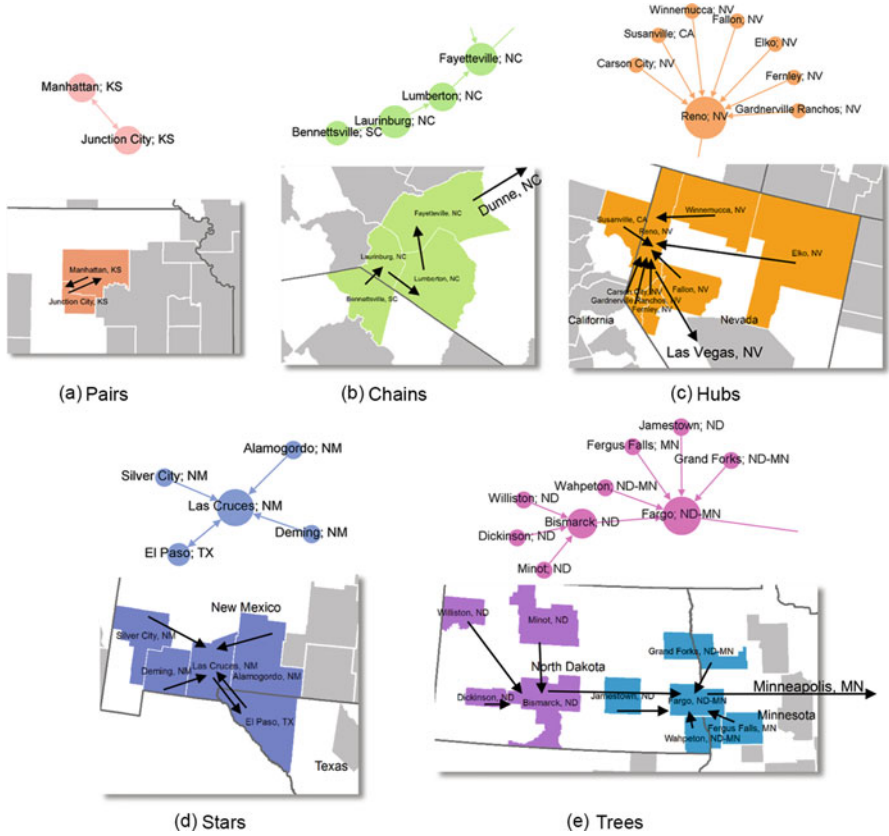


Fig. 2 Graph motifs with geographic examples. When cities are connected to their best friends, different network motifs arise, including pairs, chains, hubs, stars, and trees. These schematics illustrate differences between the roles of distinct cities within their regional systems, and what the regional systems look like as a network of flows

Table 1 Proportions of different types of nodes throughout 21 years in best friend networks

Node types percentage (%)	Hubs	Spokes	Pairs	Tree hubs	Tree spokes	Chains
Best friend	12.1	30.2	8.3	7.3	19.3	22.8
High-income	4.4	20.2	0.9	12.4	26.9	35.2
Low-income	6.5	13.4	2.7	13.9	26.4	37.2

3.2 Comparison to Prevailing Methods

The best friend method highlights the backbone structure of urban hierarchy. Since the structure is based on migration flows, we compare the best friend method with related prevailing models, such as the gravity model and radiation model,

to demonstrate that the urban hierarchy produces some patterns that are not well explained by population and distance.

Using the most recent data, we find that a parameterized gravity model with a pre-determined coefficient of friction (β) of 3 predicts about 60% of best friend ($n = 552$). The closest city is a city's best friend 28% of the time ($n = 259$). The radiation model predicts the best friend with 39% accuracy ($n = 357$). We visualized the best friend network in 2012 (Fig. 3a) based on a fitted gravity model for each city and the corresponding constellations. Instead of showing various motifs, the network is dominated by hub-spoke structure. If two cities do not connect in the data, we do not consider them as candidates for best friend cities using the gravity model. In other words, cities may have a clear choice given the gravity model, but if the city did not send any migrants to this attractive choice (or choices), they connected to their next best choices to which they actually sent migrants.

We also discover cities that “defy” gravity. We normalized the flow weight (w) by the interaction measured by a fitted gravity model to isolate flows that are large despite a small interaction estimate. These cities draw origin cities despite high travel cost (distance) and relatively low population. These networks contain more long-distance connections than best friend networks, a manifestation of “leapfrogging” in the hierarchy, and illustrate how migration connects labor forces to employment opportunities. For example, San Jose, in Silicon Valley, has an agglomeration of leading technology companies, and connects to faraway college towns throughout the years (Fig. 3b); Williston, North Dakota, becomes a hub in the network after oil resources were found in the early 2010s, and spurred jobs and economic growth.

3.3 *High- and Low-Income Routing*

The motifs of high and low income networks are collectively distinguished from best-friend networks. First, there are fewer pairs in the income networks, presumably because smaller towns that depend on each other in general take more preferential (and less mutual) routes when income is involved. Interestingly, there are very few pairs that exchange high-income migrants (0.9%), but three times as many will be best friends for low-income migrants (2.7%) (Table 1). There are more chains in the income networks although these chains are shorter. There are more hubs that are parts of trees in the income networks, indicating that hubs also connect to other hubs in these networks (Table 1).

There are also differences between low-income and high-income networks. High income networks have fewer hubs, but these hubs are quite large, as indicated by in-degree, and draw more distant connections (Fig. 4a), while hubs of low-income networks have fewer spokes. The high-income network has 39.38 unique constellations and the low-income network has 51.14 constellations, indicating more local regionalization and less overall connectivity in the low-income network. The

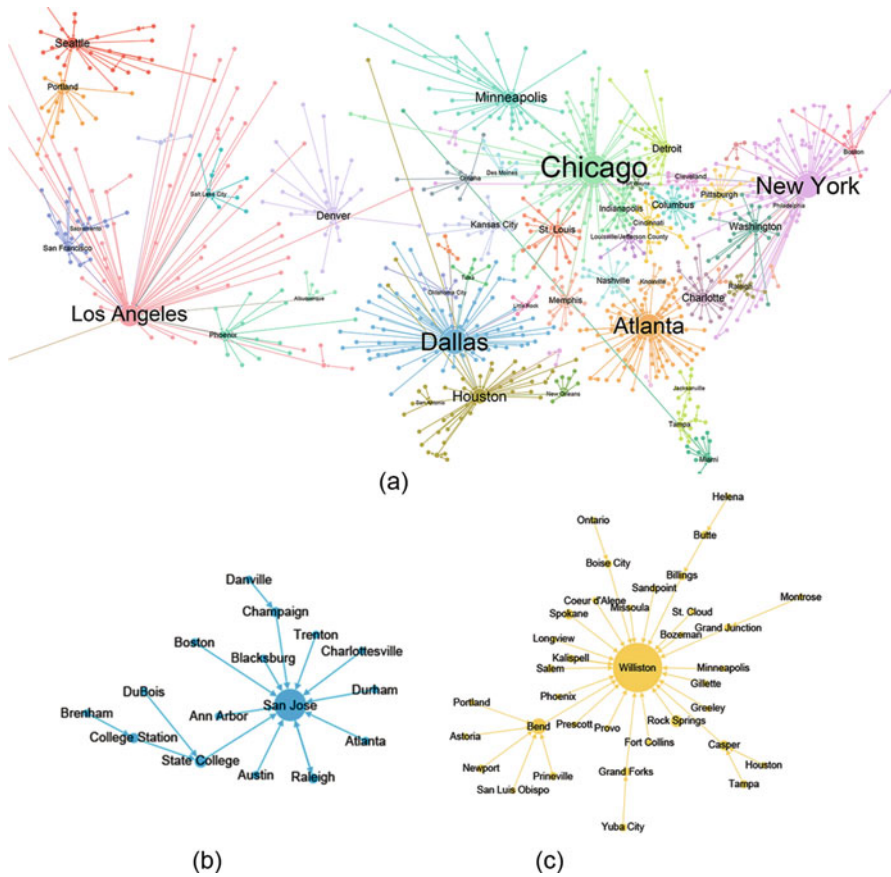


Fig. 3 (a) Best friend network in 2012 with best friend cities selected based on a fitted gravity model. The network is dominated by hub-spoke structure and real world motifs such as trees and chains are rarely seen. However, our data reveal that best friends do not always choose the destination predicted by the gravity model. For example, San Jose, California (b) is the hub for many college towns (this example is derived from the 2010 annual dataset), while Williston, North Dakota (c) starts attracting many cities after a boom in the energy industry (this example is derived from the 2012 annual data)

average constellation size is 23.49 nodes for the high income network and 18 nodes for the low-income network (Fig. 4b). The temporal change of constellation number and size does not have significant trends; the average distance of best low-income friends fluctuates while the average distance between high-income friends grows, suggesting greater mobility (Fig. 4a).

When mapping the constellations onto the geographic boundaries of their respective cities, we find that the low-income network depicts more local clustering, i.e. cities within the same constellation tend to be nearby, and also tends to follow state lines (Fig. 4c, d). Conversely, the high-income network constellations are not

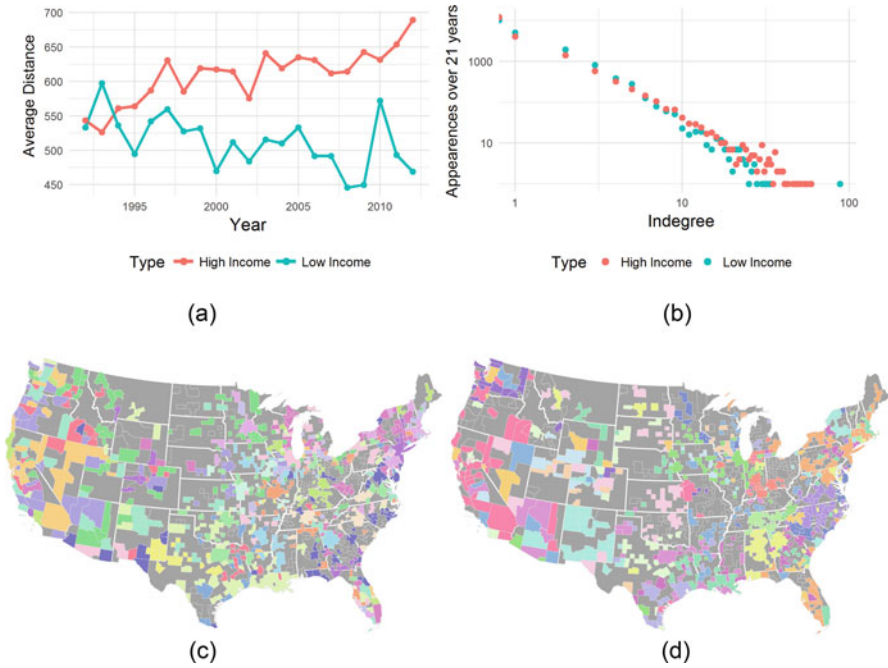


Fig. 4 High and low income network results. (a) Average edge distance in the high-income network grows over time, while distances between low-income cities have no clear trend. (b) The in-degree distribution of nodes in best-high/low-income-friend networks throughout the 21 years show that the high-income network formed hubs with high degrees, and the low-income network is marked with more cities with few incoming flows. The top 20 largest constellations using the 2012 data were identified and mapped to their corresponding city in high-income (c) and low-income (d) networks. Cities with the same colors belong to the same constellations

as contained geographically; cities in many different states often belong to the same cluster, indicating that neither boundaries nor distance appear to deter movement as significantly as in the low-income network (Fig. 4c, d).

4 Conclusions

In this study, we proposed a series of methods to study the U.S. urban hierarchy using 21 years of migration data from the U.S. IRS. We built a single allocation (best friend) network from all migrant flows, and similar networks highlighting only high- and low-income flows. Our results showed that the best friend network did not align well with the gravity and radiation models of urban interaction, and was distinguished by urban hubs, spokes and chains. Cities also tended to keep a maximum of three best friends over the time period. The income networks were marked with stronger hubs (with more spokes), that served in a system of connected hubs. The high-income network encouraged longer flows, more leapfrogging, and

exhibited more spokes that directly migrate to hubs that attract high earners nationwide, such as Cape Coral and Naples, Florida. The low-income networks contain a few more chains, which may represent that low-income migrants are more likely to move to nearby cities first and ‘climb up’ to the hubs gradually, which may result from limited mobility and limited social capital in hub cities.

The biggest limitation of our study is the variation in the meaning of a city’s “best friend”. For some cities, the best friend is a significant dependent tie as a city may send all of its migrants to this city, while for others (such as Chicago and Memphis), the best friend only absorbs about 5% of migrants. The edge that results from both of these scenarios is indistinguishable in the network. One potential remedy is to use analytical methods that account for edge weights.

The larger, eventual goals of testing the best friends method is to use it to (1) unearth ties that may not make sense economically, but may be the result of interpersonal relationships, and (2) come closer to understanding how flows affect the places to which they connect. When a city’s migrants are attracted to a city, we consider these cities to be in a similar functional region—as they are exchanging the same people between multiple cities. These functional regions are increasingly geographically disconnected, which should be accounted for in geographic partitioning exercises and in location-allocation models that use distance as an input parameter.

Acknowledgments Thanks to Prathamesh Aher for helping prepare data.

References

1. Greenwood MJ (1985) Human migration: theory, models, and empirical studies. *J Reg Sci* 25:521–544
2. Dodd SC (1950) A gravity model fitting physical masses and human groups. *Am Sociol Rev* 15:245–256
3. Pred A (1980) *Urban growth and city systems in the United States, 1840–1860*. Harvard University Press, Cambridge, MA
4. Andris C, Halverson S, Hardisty F (2011) Predicting migration system dynamics with conditional and posterior probabilities. In: *ICSDM 2011—Proceedings 2011 IEEE international conference on spatial data mining and geographical knowledge services*
5. Harris JR, Todaro MP (1970) Migration, unemployment and development: a two-sector analysis. *Am Econ Rev* 60:126–142
6. Green HL (1955) Hinterland boundaries of New York City and Boston in Southern New England. *Econ Geogr* 31:283–300
7. Brockmann D, Hufnagel L, Geisel T (2006) The scaling laws of human travel. *Nature* 439:462–465
8. Calabrese F, Dahlem D, Gerber A, Paul D, Chen X, Rowland J, Rath C, Ratti C (2011) The connected states of America: quantifying social radii of influence. In: *Privacy, security, risk and trust (PASSAT) and 2011 IEEE third international conference on social computing (socialcom)*, pp 223–230
9. Cheshire JA, Longley PA, Yano K, Nakaya T (2014) Japanese surname regions. *Pap Reg Sci* 93:539–555

10. Meyer R (2014) The Geography of NFL Fandom. The Atlantic. <https://www.theatlantic.com/technology/archive/2014/09/the-geography-of-nfl-fandom/379729/>
11. Kotkin J (2001) The new geography: how the digital revolution is reshaping the American landscape. Random House, New York
12. Heim CE (2001) Leapfrogging, urban sprawl, and growth management: Phoenix, 1950–2000. *Am J Econ Sociol* 60:245–283
13. Treyz GI, Rickman DS, Hunt GL, Greenwood MJ (1993) The dynamics of U.S. internal migration. *Rev Econ Stat* 75:209–214
14. Greenwood MJ, Sweetland D (1972) The determinants of migration between standard metropolitan statistical areas. *Demography* 9:665–681
15. Chen Y, Rosenthal SS (2008) Local amenities and life-cycle migration: do people move for jobs or fun? *J Urban Econ* 64:519–537
16. Andris C (2016) Integrating social network data into GISystems. *Int J Geogr Inf Sci* 30:2009–2031
17. Fawcett JT (1989) Networks, linkages, and migration systems. *Int Migr Rev* 23:671–680
18. Thiemann C, Theis F, Grady D, Brune R, Brockmann D (2010) The structure of borders in a small world. *PLoS One* 5:e15422
19. Manduca RA (2014) Domestic migration networks in the United States. Massachusetts Institute of Technology, Cambridge, Massachusetts
20. Guimera R, Mossa S, Turttschi A, Amaral LAN (2005) The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci, USA* 102:7794–7799
21. Davis KF, D'Odorico P, Laio F, Ridolfi L (2013) Global spatio-temporal patterns in human migration: a complex network perspective. *PLoS One* 8:e53723
22. O'Kelly ME (1998) A geographer's analysis of hub-and-spoke networks. *J Transp Geogr* 6:171–186
23. Dunne C, Shneiderman B (2013) Motif simplification: improving network visualization readability with fan, connector, and clique glyphs. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp 3247–3256
24. Jackson MO (2008) Social and economic networks. Princeton University Press, Princeton, N.J
25. Clauset A, Newman MEJ, Moore C (2004) Finding community structure in very large networks. *Phys Rev E* 70:66111
26. Csardi G, Nepusz T (2006) The igraph software package for complex network research. *Int J Complex Syst* 1695:1–9
27. Tarver JD, McLeod RD (1973) A test and modification of Zipf's hypothesis for predicting interstate migration. *Demography* 10:259–275
28. Simini F, González MC, Maritan A, Barabási AL (2012) A universal model for mobility and migration patterns. *Nature* 484:96–100
29. Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. In: Third International AAAI Conference on Weblogs and Social Media. pp 361–362