SOCIAL SCIENCE

Origin–destination (OD) of the interprovincial floating population of China

Tinghua Ai, Mengjie Zhou, Jing Tian and Ning Ye

School of Resource and Environment Sciences, Wuhan University, Wuhan, People’s Republic of China; Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan, People’s Republic of China; Key Laboratory of Digital Mapping and Land Information Application Engineering, National Administration of Surveying, Mapping and Geoinformation, Wuhan University, Wuhan, People’s Republic of China

ABSTRACT

China’s large floating population catalyzes economic development, redistributes the population, and promotes social changes to the social structure, lifestyle patterns, and cultural values. The floating population is a unique group faced with numerous problems stemming from cultural and lifestyle differences. Understanding and visualizing the distributional characteristics and patterns of the floating population is crucial for developing effective social policies. In this study, the origin–destination (OD) map and its adjustment destination–origin (DO) map are used to represent the interprovincial floating population. The OD or DO map can be regarded as a two-level spatial treemap representing the floating population recorded by pairs of locations. In accordance with the different hierarchy of the two-level spatial treemap, the OD and DO are distinguished. The migration volume of the floating population data is represented by the OD and DO maps to find the population’s direction of movement, magnitude, and hot-spots. The migration indicators including migration effectiveness, migration preference indexes, and sex ratios are represented by the OD and DO maps, which assist in reviewing and studying the deep patterns of floating population.

ARTICLE HISTORY

Received 12 May 2015 Revised 3 August 2016 Accepted 8 August 2016

KEYWORDS

Floating population; OD map; China

1. Introduction and background

Over the last three decades a new demographic phenomenon, known as the floating population of China, has attracted increasing attention from the government, scientific institutions, and society (Liang & Ma, 2004). Prior to the 1980s population mobility in China was very low. However, since the mid-1980s, economic reforms and the relaxation of migration controls have brought about sharp increases in population mobility (Cai & Wang, 2003; Liang, 2001). As of 2010, the floating population has reached 261 million (Cai & Wang, 2003; Liang, 2001). China uses a two-track migration tracking system which includes both permanent migration and temporary migration (Chan, 1994; Gu, 1992; Sun & Fan, 2011). The concepts of permanent migration and temporary migration are rooted in the Chinese institution of household registration, known as ‘hukou.’ Temporary migrants are individuals whose place of residence differs from their place of registration. Permanent migrants, in contrast, are migrants who have changed their registration to their current place of residence. It is the place where an individual is registered, rather than the duration of the individual’s stay that determines whether someone is considered a permanent migrant or temporary migrant (Goldstein & Goldstein, 1991). The term ‘floating population’ is used to describe these temporary migrants (Cai & Wang, 2003).

China has been an attractive area of study in the last few decades because it has been experiencing notable changes in its economic and social spheres; the extensive floating population is one of the most important aspects of these changes. The occurrence and existence of the floating population has not only promoted rapid economic development, it also has given rise to profound social changes, including population redistribution from rural to urban areas, and changes in social structures, lifestyles, and personal and moral values (Chan, 2013). Understanding the spatial distribution of the floating population allows for the development of strategies to promote the living conditions of migrants, solve problems that arise from the phenomenon of floating population, and promote social progress.

Geographical movement can be better understood through cartographic visualization. Abundant work has been made in visual analytics of geographical movement, focusing on (1) looking at trajectories, (2) looking inside trajectories, (3) bird’s-eye view on movement, and (4) investigating movement in context. There are two major groups of analysis within bird’s-
eye view on movement. These are investigating the presence of moving objects and investigating migration flows (Andrienko & Andrienko, 2012). Nationwide floating population can be regarded as ‘flows’ when movement is recorded by pairs of locations: individuals’ moves from location A to B. Visualizing spatial patterns of the floating population is an essential yet challenging task, which demands that it can provide an overview at the national scale and detail at the provincial scale, offer the ability to distinguish migration origins from destinations, preserve spatial configuration, and distinguish in-migration and out-migration, while avoiding overlaps and occlusions of migration flows effectively.

Both flow maps and forms of origin–destination (OD) matrices can be used to visualize migration flows. Flow maps project a geographic space onto a plane and plot each migration track as a line from origin to destination (Tobler, 1987). This method is only effective in visualizing small datasets because the intersections and occlusions of flow symbols. Methods have been developed to solve these problems; however, those methods are limited. Coping with the problem by filtering the flows (Ellis & Dix, 2006; Rae, 2009) does not provide an overview of migration, and much information is lost in the process. Bundling flow edges (Cui, Zhou, Qu, Wong, & Li, 2008; Holten, 2006; Holten & van Wijk, 2009; Phan, Xiao, Yeh, & Hanrahan, 2005), aggregating flow locations (Guo, 2009), and cluster flows (Zhu & Guo, 2014) may facilitate the loss or distortion of knowledge, while some overlaps or intersections of flow edges may remain.

Unlike flow maps, there are some forms of the OD matrix that do not cause intersections and occlusions of flows. The idea to represent flow in a matrix originated from Voorhees (1955), who used a matrix grid to depict the origin and destination through correlating rows and columns. Although the original OD matrix has obvious benefits for detecting clusters, it lacks appropriate spatial content. Projecting a two-dimensional space into one-dimensional sequences (Marble, Gou, Liu, & Saunders, 1997), or linking the matrix view and the geographic view (Guo, 2007; Guo & Gahegan, 2006) to address the problem still creates difficulty in identifying the spatial characteristics of the matrix cells. Therefore these methods do not meet our demands. The OD map (Wood, Dykes, & Slingsby, 2010) can solve the problem. The OD map is a form of a two-level spatial treemap (Wood & Dykes, 2008) that arranges the OD matrices according to the geographic positions of the places. It has been used in visualizing London’s bicycle-rental scheme (Wood, Slingsby, & Dykes, 2011) and historical internal migration in Ireland (Slingsby, Kelly, Dykes, & Wood, 2012). In this study, the OD map is used to visualize the spatial distribution of China’s interprovincial floating population.

### 2. Data and methods

#### 2.1. Chinese migration data

The major mapping data used in this study was obtained from the sixth national population census of the People’s Republic of China (the 2010 Chinese Census), which defines the floating population as individuals who have resided at their place of destination for at least six months without local household registration status at the sub-county level. The floating population can be divided into two categories: the interprovincial floating population and the intraprovincial floating population. The former refers to the movement between provincial-level units (the highest-level Chinese administrative divisions; there are 34 such divisions, classified as 23 provinces, 4 municipalities, 5 autonomous regions, and 2 Special Administrative Regions). The 2010 population census of the People’s Republic of China recorded the female and male floating population volumes in pairs of provinces. Hong Kong, Macau, and Taiwan were not represented in the census.

#### 2.2. Methods

The OD map was used because of its ability to show spatial patterns of the floating population and its suitability for visualizing statistical data.

The OD map transforms graphic space so that each two-dimensional vector \( \mathbf{p}_o \rightarrow \mathbf{p}_d \) is represented by a cell \( p_{od} \) in a two-dimensional matrix, and the cell reflects the spatial relationships between origins (Wood et al., 2010). There are three steps to visualize the floating population using an OD map:

1. Order the cells (first level cells) to reflect their two-dimensional geographic location. This is achieved using a spatial treemap (Wood & Dykes, 2008). As shown in Figure 1(a), the provinces of China were tessellated into a grid using a spatial treemap where each province was represented by a grid cell. ‘Dummy’ cells are inserted into the grid to maintain the relative position of each province and the shape of the landmass as much as possible.

2. Nest the second level cells within each first level cell. Given one province, the corresponding first level cell is decomposed into self-similar second level cells. Each trajectory is referenced by two cell locations: the grid cell (first level cells) in which the trajectory’s origin lay, and the grid cell (second level cells) in which its destination lay, as shown in Figure 1(b). As shown in Figure 1, the red migration flow from Gansu to Zhejiang is represented by the red OD map cell and the green migration flow from Zhejiang to Gansu is
represented by the green OD map cell. The arrangement is a spatial treemap, denoted as sHier(/,$origin,$destination); sLayout(/,SP,SP); sSize(/,FIXED,FIXED) using the hierarchical visualization expression language (HiVE) to describe the representation (Slingsby, Dykes, & Wood, 2009). sHier describes the structure of the treemap; the first level cells (the ‘larger’ cells) represent the origins and the second level cells (the ‘smaller’ cells) represent the destinations. sLayout describes
the layout algorithm of the treemap, and a spatial layout is used in the first and second level cells. sSize describes the size of treemap cells. The sizes of treemap cells at the first and second level are fixed. In order to recognize the meaningful OD cells easily, the ‘dummy’ cells are removed from the second level, as shown in Maps 1, 2, 3, and 4.

(3) Render the second level cells according to an attribute of the migration flow. The second level cells are rendered according to migration volume, and processed data, including migration effectiveness, migration preference index, and sex ratio. When the first level and second level cells represent the same province, the second level cell is rendered gray, as shown in the Main Map.

2.2.1. Migration volume

The migration volume refers to the amount of movement, corresponding to the size of the floating population. Given a province, we concern where the floating population moves to and comes from. Map 1 shows the pattern of migration volume. The proportional point symbol is used to represent migration volume. Two different views of the migration origins and destinations are provided in order to represent the respective spatial patterns correctly. The first, denoted as sHier(/,$origin,$destination); sLayout(/,SP,SP); sSize(/,FIXED,FIXED), illustrates where the floating population moves to, which is shown in Map 1(a). The second, denoted as sHier(/,$destination,$origin); sLayout(/,SP,SP); sSize(/,FIXED,FIXED), as shown in Map 2. Red indicates that there are more in-migrants than out-migrants, and blue indicates that there are more out-migrants than in-migrants. The darker the color, the larger the gap is.

2.2.2. Migration effectiveness

The migration effectiveness indicates whether the movement of the floating population between a pair of provinces is unidirectional or bidirectional. Migration effectiveness for special pairs of provinces is defined as follows (Plane, 1984):

$$e_{ij} = \frac{n_{ij}}{t_{ij}} \times 100,$$

where

$$n_{ij} = m_{ij} - m_{ji},$$

$$t_{ij} = m_{ij} + m_{ji},$$

$m_{ij}$ represents the migration from the $i$th province to the $j$th province and $m_{ji}$ represents the migration from the $j$th province to the $i$th province.

Migration effectiveness values range from $-100$ to 100. A value of $-100$ indicates that all the migrants are in-migrants. A value of 100 means that all the migrants are out-migrants. Zero indicates a roughly equal flow between the provinces. The OD map illustrates the migration effectiveness in 2010, which is denoted as sHier (/,$origin,$destination); sLayout (/,SP,SP); sSize (/,FIXED,FIXED), as shown in Map 2. Red indicates that there are more in-migrants than out-migrants, and blue indicates that there are more out-migrants than in-migrants. The darker the color, the larger the gap is.

2.2.3. Migration preference index

The migration preference index reflects the different level of attractiveness that various provinces might have to the floating population. Taking the population volumes of origins and destinations into consideration, the migration preference index can be defined as follows (Shryock, Siegel, & Larmon, 1980):

$$I_{ij} = \frac{m_{ij}}{(p_i/p_j)((p_i - p_j)\sum_{k}m_{kj})}$$

where $m_{ij}$ represents the migration from the $i$th province to the $j$th province, $p_i$ is the population of the $i$th province, $p_j$ is the total population of the $j$th province, and $k$ is a constant, usually 100.

The denominator is the expected value of migration from $i$th province to $j$th province, considering the average influence of origin and destination population size. A migration preference index above 100 indicates that the trend of migration from the $i$th province to the $j$th province is stronger than the national average. A migration preference index below 100 indicates that the trend of migration from the $i$th province to the $j$th province is weaker than the national average. Map 3(a), denoted as sHier(/,$origin,$destination); sLayout (/,SP,SP); sSize(/,FIXED,FIXED), is also an OD map, and reflects the migration trend, while DO Map 3(b), denoted as sHier(/,$destination,$origin); sLayout(/,SP,SP); sSize(/,FIXED,FIXED), reflects the migration’s attractiveness.

2.2.4. Sex ratio

The sex ratio is another aspect of the floating population. The sex ratio is defined as follows:

$$g_{ij} = \frac{x_{ij}}{y_{ij}},$$

where $x_{ij}$ represents the male migration from the $i$th province to the $j$th province and $y_{ij}$ represents the female migration from the $i$th province to the $j$th province.

Map 4 illustrates the sex ratio of the floating population. Map 4(a), an OD map, is denoted as sHier(/,$origin,$destination); sLayout(/,SP,SP); sSize(/,
3. Discussion

The OD and DO maps demonstrate a basic pattern where the floating population moved from the west to the east, which is reflective of the economic disparity between the western and the eastern parts of China. In Map 1(b), the red squares in the first level cells representing the eastern parts of China are much larger than the squares in the cells representing the western part of China. This indicates that there was a greater percentage of the floating migration that migrated to the easternmost region of China. As depicted in Map 1(a), the blue squares in the first level cells found in the center of the map are significantly larger than those in the cells at the edges of the map. This indicates that a considerable percentage of the floating population originated from the central regions of the country.

Map 2 shows divergent provincial migration effectiveness. At the provincial level, in-migration and out-migration were extremely imbalanced. Provinces in China can be divided into three province groups, based on whether the movement of the floating population between a pair of provinces is unidirectional or bidirectional: (1) provinces of unidirectional in-migration, represented by Beijing, Shanghai, Tianjin, and Guangdong (the second level cells within the provinces are nearly all in dark red); (2) provinces of unidirectional out-migration, represented by provinces in central China such as Anhui and Sichuan (the second level cells within these provinces are nearly all in dark blue); and (3) provinces of bidirectional migration, represented by Shanxi and Yunnan (the second level cells within these provinces are in blue or red). Among provinces of unidirectional in-migration, Beijing and Shanghai were the strongest population gainers – they even gained population from other nearly unidirectional in-migration provinces, such as Tianjin and Guangdong.

The floating population of China shows a spatial focus, which refers to the ‘inequality that exists in the relative volumes of a set of origin and destination specific migration flows’ (Plane & Mulligan, 1997). The most striking destinations of the floating population in 2010 included Shanghai, Beijing, Guangdong, Zhejiang, Tianjin, and Fujian. These provinces drew a very strong attraction from other provinces, as demonstrated in Maps 1(b) and 3(b). The most striking origins of the floating population in 2010 were in central China, including Sichuan, Anhui, and Henan, as demonstrated in Maps 1(a) and 3(a). The floating population from Anhui showed the strongest migration trend to Shanghai.

The interprovincial floating population was characterized by two patterns of migration distance. The first pattern is that people tend to migrate to neighboring provinces and more developed provinces that are within a short distance. As illustrated in Map 3(b), the second level cells with relatively dark color or high color saturation are commonly around gray cells, and show a gradual color change pattern radiating out from the gray cells, reflecting the concept that migration is negatively related to distance to some extent (Fan, 2005). Flows from Hebei to Beijing and Tianjin, and from Guangxi to Guangdong are characterized by relatively short distance, as shown in Map 3(a). The second pattern is that people tend to migrate to coastal or more developed provinces even when they are far away. As illustrated in Map 3(a), the floating population from populous places of origin usually migrated to provinces that were relatively far away, concentrating on populous destinations. An example of this is the flows from Sichuan to Guangdong, Zhejiang, Shanghai, and Jiangsu, as shown in Map 3(a).

The conventional wisdom about sex balance in migration is that the migrant sex ratio (male/female) is higher than the sex ratio for the general population as men may have greater migration propensity than women (Sun & Fan, 2011). Indeed, most regions of China conformed to this expectation. However, there are some variances and outliers among the regions. In Map 4(a), the second level cells in the western part of China, Tianjin, and Inner Mongolia are a darker blue than other provinces, indicating a relatively higher male proportion of the floating population from these regions than other regions. And the Particularly, the floating populations from Tibet to Tianjin, from Inner Mongolia to Jiangsu, from Xinjiang to Guangdong, and from Qinghai to Guizhou had an extremely high sex ratio – the corresponding dark blue grid cells are striking. In Map 4(b), the second level cells in the western part of China are in very light blue or even in red, indicating a much lower sex ratio in the floating population migrating to these regions. It is interesting to note that the floating population migrating to Tibet had a high proportion of females – the second level cells in Tibet are nearly all red. Specifically, the floating population from Gansu to Tibet had an extremely low sex ratio, as the corresponding grid cell is notable.

4. Conclusion

The floating population is a unique group in China, which has affected many aspects of society and is attracting increasing attention from the government, scientific institutions, and society. It is necessary to investigate the patterns and spatial and temporal
changes of the floating population in order to better understand all aspects of this phenomenon.

In this paper, an OD map and its adjusted DO map were used to illustrate the interprovincial floating population of China. The OD map transforms graphic space whereby each migration flow from origin to destination is presented by a cell in a two-dimensional matrix. The cells reflect the spatial relationships between origins or destinations. The OD and DO maps of China effectively illustrate the distribution patterns of the interprovincial floating population, as they maintain the relative positions of the provinces and also show details and an overview of this special population. In the future, an OD map may be used to illustrate the structure of the floating population on other characteristics, such as age and education level through a reasonable representation.

Software

The data calculation was performed using Microsoft Excel. The OD and DO maps were generated using C++. The final maps were produced with Coreldraw x6.

Notes

1. http://www.gov.cn/test/2012-04/20/content_2118413.htm.
2. Sub-county level units include streets (jiedao), towns (xiang), and townships (zhen).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work was supported by the Special Program for Basic Science of China [grant number 2013FY112800] and the National Natural Science Foundation of China [Grant number 41531180].

ORCID

Tinghua Ai  http://orcid.org/0000-0002-6581-9872

References

Andrienko, N., & Andrienko, G. (2012). Visual analytics of movement: An overview of methods, tools and procedures. Information Visualization 12(1), 3–24, doi:10.1177/1473871612457601

Cai, F., & Wang, D. (2003). Migration as marketization: What can we learn from China’s 2000 census data? The China Review, 3(2), 73–93.

Chan, K. W. (1994). Cities with invisible walls: Reinterpreting urbanization in post-1949 China. Hong Kong: Oxford University Press.

Chan, K. W. (2013). China: Internal migration. In I. Ness & P. Bellwood (Eds.), The encyclopedia of global human migration. Oxford: Blackwell Publishing.

Cai, W., Zhou, H., Qu, H., Wong, P. C., & Li, X. (2008). Geometry-based edge clustering for graph visualization. IEEE Transactions on Visualization and Computer Graphics, 14(6), 1277–1284.

Ellis, G., & Dix, A. (2006). Enabling automatic clutter reduction in parallel coordinate plots. IEEE Transactions on Visualization and Computer Graphics, 12(5), 717–724.

Fan, C. C. (2005). Intermigration population, migration redistribution, and regional development in China: 1990 and 2000 census comparisons. The Professional Geographer, 57(2), 295–311.

Goldstein, S., & Goldstein, A. (1991). Permanent and temporary migration differentials in China. Honolulu: East-West Population Institute.

Gu, S. (1992). Two types of population migration in China: A comparative study. Chinese Journal of Population Science, 4(1), 75–84.

Guo, D. (2009). Flow mapping and multivariate visualization of large spatial interaction data. IEEE Transactions on Visualization and Computer Graphics, 15, 1041–1048.

Guo, D., & Gahegan, M. (2006). Spatial ordering and encoding for geographic data mining and visualization. Journal of Intelligent Information Systems, 27, 243–266.

Holten, D. (2006). Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. IEEE Transactions on Visualization and Computer Graphics, 12(5), 741–748.

Holten, D., & van Wijk, J. (2009). Force-directed edge bundling for graph visualization. Computer Graphics Forum, 28, 983–990.

Liang, Z. (2001). The age of migration in China. Population and Development Review, 27(3), 499–528.

Liang, Z., & Ma, Z. (2004). China’s floating population: New evidence from the 2000 census. Population and Development Review, 30(3), 467–488.

Marble, D. F., Gou, Z., Liu, L., & Saunders, J. (1997). Recent advances in the exploratory analysis of interregional flows in space and time. In Z. Kemp (Ed.), Innovations in GIS (Vol. 4, pp. 75–88). London: Taylor & Francis.

Phan, D., Xiao, L., Yeh, R., & Hanrahan P. (2005). Flow map layout. In J. Stasko & M. Ward (Eds.), IEEE symposium on information visualization (InfoVis 05), (pp. 219–224). Washington, DC: IEEE.

Plane, D. A. (1984). A systemic demographic efficiency analysis of U.S. interstate population exchange, 1935–1980. Economic Geography, 60, 294–312.

Plane, D. A. & Mulligan, G. F. (1997). Measuring spatial focusing in a migration system. Demography, 34(2), 251–262.

Rae, A. (2009). From spatial interaction data to spatial interaction information? Geovisualisation and spatial structures of migration from the 2001 UK census. Computers, Environment and Urban Systems, 33(3), 161–178.

Shryock, H. S., Siegel, J. S., & Larmon, E. A. (1980). The methods and materials of demography (Vol. 2). Washington, DC: Department of Commerce, Bureau of the Census.

Slingsby, A., Dykes, J., & Wood, J. (2009). Configuring hierarchical layouts to address research questions. IEEE Transactions on Visualization and Computer Graphics, 15, 977–984.
Slingsby, A., Kelly, M., Dykes, J., & Wood, J. (2012). OD maps for studying historical internal migration in Ireland. In IEEE conference on Information Visualization (InfoVis), Seattle, WA.

Sun, M., & Fan, C. C. (2011). China’s permanent and temporary migrants: Differentials and changes, 1990–2000. The Professional Geographer, 63(1), 92–112.

Tobler, W. (1987). Experiments in migration mapping by computer. The American Cartographer, 14, 155–163.

Voorhees, A. (1955). A general theory of traffic movement. In Institute of traffic engineers past presidents’ award paper (pp. 46–56). New Haven, CT: Institute of Traffic Engineers.

Wood, J., & Dykes, J. (2008). Spatially ordered treemaps. IEEE Transactions on Visualization and Computer Graphics, 14(6), 1348–1355.

Wood, J., Dykes, J., & Slingsby, A. (2010). Visualisation of origins, destinations and flows with OD maps. The Cartographic Journal, 47(2), 117–129.

Wood, J., Slingsby, A., & Dykes, J. (2011). Visualizing the dynamics of London’s bicycle-hire scheme. Cartographica: The International Journal for Geographic Information and Geovisualization, 46(4), 239–251.

Zhu, X., & Guo, D. (2014). Mapping large spatial flow data with hierarchical clustering. Transactions in GIS, 18(3), 421–435.