Beijing City Lab

SOCIAL-SPATIAL STRUCTURE OF BEIJING: A SPATIAL-TEMPORAL ANALYSIS

Xiaobai Yao
Assoc Prof, Department of Geography, University of Georgia
xyao@uga.edu
http://geography.uga.edu/directory/profile/yao-xiaobai/

Abstract
In the past couple of post-reform decades of rapid urban growth of Chinese cities, Beijing experienced dramatic spatial and socio-economic transformation. This study aims to examine the evolution of Beijing’s social-spatial structure in the transforming process over the past 20 years. We first collect and process Beijing’s socio-demographic variables at the level of sub-district in three years between 1990 and 2009. Then a two-step spatial clustering approach is taken for spatial data mining of the spatial structures of social areas in the three time periods. Results are visualized in GIS. Finally, the interpretation of the analysis result is provided, followed by a discussion of policy implications for the development of a socially sustainable municipality.

INDEX WORDS: social-spatial structure, Beijing, spatial-temporal, cluster analysis
1 Introduction

Since China’s open policy in the late 1990s, China has experienced dramatic urban expansion. In the process, spatial and socio-economic reorganization and restructuring have taken place in most Chinese cities (Lai and Yat-Ming 2006; Feng, Zhou et al. 2007). Beijing, as the capital of China and one of the largest cities in China and in the world, has gone through an unprecedented process of urbanization in this time period (He, Okada et al. 2006). According to the latest census in 2010, Beijing’s population of permanent residents has reached 19.6 million, growing from the 10.8 million in 1990, indicating an increase rate of over 80 percent over the past two decades. Beijing has attracted the attention of many researchers from a range of perspectives such as land use and landscape changes, public transportation, migrating population and so on. However, to the best of our knowledge, no previous study has yet examined the identification of social areas in Beijing and the spatio-temporal changes of the social areas during the urban transformation processes. In this study, we use the term social areas to refer to the continuous urban areas that contain residents of similar socioeconomic characteristics, demographic status, or lifestyle.

After nearly two decades of urban reform, the once relatively homogeneous social areas in Beijing have become increasingly differentiated. Several reasons may account for it. First, the fast economic transformation has resulted in differentiation of socioeconomic status among residents in the urban space. Secondly, as industrialization and urbanization released large numbers of farmers from the rural land, the influx of rural migrants greatly reshaped urban structures of the area (Feng, Zhou et al. 2007). Thirdly, transformation of urban social structures has also been influenced by changes in education and employment systems, in land-use and housing policy, as well as in other neighborhood dynamics. Finally, after being selected as the
host city for the 2008 Olympic Games in 2001, the restructuring of Beijing’s urban space was accelerated through a series of Olympic-related construction, relocation and environment amelioration activities.

This study attempts to investigate the spatial structure of Beijing’s social areas and the change of the structure during the transformative period of the city. We collected and processed Beijing’s demographic and socio-economic GIS data for three years, 1990, 2000, and 2009, spanning the past two decades. Then we performed spatial data mining to identify social areas in each of the three years. Changes of the spatial structure over time are identified and discussed. The paper is organized as follows. The next section reviews related prior research on Beijing’s urban transformation. Section 3 introduces data and methods of our study. Section 4 concludes the paper with discussions of major findings and future research needs.

2 Literature Review

Beijing’s rapid urban growth is considered a unique and representative case of China’s rapid urbanization in recent years. Many researchers studied this process from various angles. Many studies investigated landscape and land use patterns (Qi, Henderson et al. 2004; Wu, Hu et al. 2006), land use changes (Chen, Gong et al. 2003; Wu, Li et al. 2006), housing issues (Li 2000; Huang 2004; Huang and Jiang 2009), the influx of rural migrants and its implication on urban forms (Gu, Chan et al. 2006; Barabantseva 2008; Zheng, Long et al. 2009), and transportation, job accessibility and residential relocation (Shin 2009; Zhao and Lu 2009; Zhao et al. 2009). More recently, the impact of the Olympic Games on urban planning and restructuring of Beijing’s social space has attracted considerable research attention (Smith and Himmelfarb 2007; Zhang and Zhao 2009). In many of such studies, geosimulation techniques were applied to
estimate and evaluate urban growth (Chen, Gong et al. 2002; He, Okada et al. 2006; He, Okada et al. 2008).

In the world, a large collection of literature can be found on social spatial structure related to social polarization and segregation in, for instance, American cities. Complex models were developed in the latter half of the twentieth century reflecting the increasing spatially complex urban society. Factorial ecology method has been one of the most popular techniques for studying intra-urban social spatial structures (Murdie 1969; Cadwallader 1996; Gu et al. 2006). The social-spatial structural studies have been considered particularly interesting because they can establish aggregated mosaics combining social-area and physical space. Nevertheless, the current body of literature is short of studies of the changing social structure in cities of developing countries. Studies of socialist cities are even fewer, given the fact that most researchers have been informed by theories and empirical studies pursued from a capitalist perspective (Sykora 1999; Wu 2005).

The process of reconfiguration of social areas in Chinese cities is expected to be quite different than those in most other countries for several reasons. First of all, Chinese cities are under the influence of mixed economy, reflecting characteristics of both market economy and planned economy. While continuing to adhere to socialist principles, Chinese economic reforms have fostered a great reliance on the free market, which not only accelerated economic prosperity but also led to growing socioeconomic divisions between residents (Hu and Kaplan 2001). Secondly, China has a unique Hukou registration system which confines people’s job-related moves. Thirdly, the family planning policy has also resulted in special family structure in China. All of the above promise direct impacts on social-spatial structures (Feng, Zhou et al. 2007). Despite of the considerable size of literature on Beijing’s dramatic urban growth, few studies
focus on structural change of Beijing’s social areas. Social-spatial areas are studied in other Chinese cities such as Guangzhou, Shanghai, and Hongkong (Lo 1994; Lo 2005; Wu 2005). Sit (1996) studied Beijing’s social areas in the early post-reform years. Highlighting changes in housing ownership patterns, the study considered centrally-planned land-use and the relationship between residence and work-unit as a function of social-spatial structure transformation. Gu et al (2005) analyzed the social-spatial structure of Beijing’s more developed area and concluded that social differentiation began to occur after a decade of urban reform, especially with regard to social-economic status. Feng et al (2007) investigated Beijing’s social landscape during 1982 and 2000, considering the relative strength of social differentiation processes from the standpoint of an ecological approach. However, as the urbanization continues intensively in the 21st century in the transitioning economy, how the social areas evolve in the process remains to be explored. Our study aims to fill the gap by studying with up-to-date data in Beijing and by reviewing the full extent of the changing landscape of social areas in Beijing in the two decades of post-reform years. We examine the spatial patterns of social areas in three characteristics years, 1990, 2000, and 2009.

3 Study Area and Methods

As the capital of China, Beijing is one of the largest metropolitan areas in China. It has a total area of 16.4 thousand square kilometer, of which 1.368 thousand is urban area and 1.289 thousand is built-up area. As discussed earlier, its population has increased over 80 percent in the past two post-reform decades.
Following Chinese cities’ tradition of self-sustainment, Beijing encompasses a mix of urban and rural land uses. Before merging two urban districts in 2010, it had 18 administrative districts (more urban land) or counties (more rural), which can be further categorized into three regions (Feng, Zhou et al. 2007): central city, inner suburb, and outer suburb. With a topographic background, Figure 1 shows the administrative boundary of the metropolitan area. The lighter areas, which are around the center of the satellite image, suggest built-up areas of Beijing.

3.1 Data collection and processing

This study collects and analyzes time-series population and socio-demographic data at sub-district (neighborhood) level, which is the finest spatial granularity in China’s census data. Data have been collected for the years of 1990, 2000 and 2009. The spatial boundaries of these spatial units are digitized from official publications including atlases, yearbooks, and others. Attributes of the spatial units are obtained from China’s 4th (1990) and 5th (2000) national censuses, while that of 2009 were processed based on Beijing’s statistical yearbooks of 2009 and remote sensing images in the same time period. For the year 2009 data, the statistical yearbook includes similar information as that of the national population census, except that the available data are aggregated at the district level. To produce data at finer spatial units that is consistent with the
other two time periods, we interpolate the district level statistics to the sub-district level based on dasymetric mapping, a population (or other variable) interpolation method based on ancillary information such as remote sensing images. In the interpolation process, we choose the sub-district boundaries in 2000 as the spatial footprint.

After digitizing, compiling, and data pre-processing, we have census or census-like data at sub-district level in GIS for the three time periods. The attribute data include 16 variables for year 1990, 65 for year 2000, and 35 for the processed 2009 data. These variables reflect demographic and socioeconomic characteristics in several categories. Table 1 lists the types of variables in 1990, 2000, 2010 respectively.

Table 1. Types of population variables in the datasets for the three time periods

<table>
<thead>
<tr>
<th>1990 (16 variables)</th>
<th>2000 (65 variables)</th>
<th>2009 (35 variables)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Population</td>
<td>Population</td>
</tr>
<tr>
<td>Households</td>
<td>Households</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nationality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population Movement</td>
<td></td>
</tr>
<tr>
<td>Hukou Status</td>
<td>Hukou Status</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age Structure</td>
<td></td>
</tr>
<tr>
<td>Education Level</td>
<td>Educational Level</td>
<td></td>
</tr>
<tr>
<td>Employed by types of occupation</td>
<td>Employed by types of occupation</td>
<td></td>
</tr>
<tr>
<td>Employed by Sector</td>
<td>Employed by Sector</td>
<td></td>
</tr>
<tr>
<td>Outside of Labor Force</td>
<td>Outside of Labor Force</td>
<td></td>
</tr>
<tr>
<td>Living Space</td>
<td>Living Space</td>
<td></td>
</tr>
<tr>
<td>Ownership or rental of houses by household</td>
<td>Ownership or rental of houses by household</td>
<td></td>
</tr>
</tbody>
</table>

3.2 Analysis

From Table 1, it is clear that we are dealing with high dimensional attributes. In addition, the types of variables for the three time periods are not exactly the same. Therefore we apply spatial data mining to tease out general social patterns from the varied datasets. As a type of spatial data mining, spatial clustering analysis is chosen for this purpose. We develop a two-fold
approach, integrating self-organizing maps (SOM) and principal component analysis (PCA) for spatial clustering. Combined with geographic information systems, SOM proves to be capable of filtering complex demographic situations and it is considered to have the advantage of retaining fundamental social fingerprints (Yao 2007; Spielman and Thill 2008). However, the method shares limitations with many other clustering techniques in that the interpretation is somewhat subjective and thus needs to be considered with caution. Thus we propose to integrate it with PCA to gain more objective evidences for the interpretation of clusters. Figure 2 is the proposed method for our analysis.

![Figure 2. Research Design](image)

**Spatial Clustering Analysis with Self-Organizing Maps**

Spatial clustering is the process of grouping a set of objects into classes or clusters based on similarity. It has been widely applied in various data-driven research (Steenberghen, Aerts et al.; Hallencreutz and Lundequist 2003; Wang, Luo et al. 2005; Wang, Leung et al. 2006; Baum, Haynes et al. 2007; Yee, Tung et al. 2007; Mu and Wang 2008; Steenberghen, Aerts et al. 2009). In this study, we want to apply spatial clustering techniques to group sub-districts in consideration of their population make-ups. Self-Organizing-Maps (SOM) is an artificial neural
network algorithm developed in the early 1980s by Kohonen to identify clusters from complex raw data. The method projects high-dimensional input data onto a regular, low-dimensional array of nodes (Kohonen 1982; Kohonen 2001). SOM has demonstrated to be more effective than traditional methods in many studies due to its superior features such as unsupervised learning, ability to handle large data sets, and its visualization of output for easy identification of clusters (Tayman, Smith et al. 2007).

The dataset prepared for the SOM analysis is in the form of a matrix, with each row representing a sub-district and each column representing an attribute of the sub-district. Each observation (a sub-district) has a vector of the chosen variables, denoted as \(<x_1, x_2, \ldots, x_n>\). The output of SOM is a 2-dimensional grid. Each cell of the grid is called an output node. Each output node has a vector (also called weights) with the same dimensionality as that of the input, the attribute vector of an observation. Therefore, input and output vectors are comparable and distances between each pair can be calculated. The observations are presented one by one to train the output nodes according to the SOM algorithm. An observation is assigned to the output node that is closest in the attribute space, and this node is called the best-matching node for the observation. The weights of the best-matching node and other neighboring nodes are slightly modified so that they become closer to the input observation in the attribute space. In the training process, neighborhood function is an influential parameter contributing to the training result. We choose the popular Gaussian function for this. That is, the adjustment of weights will be maximized for the best-matching unit itself, and decrease with the increase of distance between a neighborhood node to the best-matching node (Pragya Agarwal 2008). Finally, the output map will have all the nodes (weights) trained, with adjacent nodes more likely to be similar to each other. Clusters of output nodes are visually identifiable on the 2D output grid map.
Because each sub-district is assigned to the best-matching output node, the sub-district thus belongs to the cluster of the associated output node. To provide more measurable summary of a SOM result, constructing unified distance matrix (U-matrix) is one of the most commonly used methods. U-matrix visualizes the distances between neighboring SOM output nodes, and thus shed lights on cluster structure in a SOM map. High values in the U-matrix suggests borders of clusters, while patches of low values indicate clusters themselves (Vesanto et al 1999).

**Principal Component Analysis**

However, what remains a problem is the interpretation of each cluster, which can be used to understand and to label the clusters. The U-matrix provides little insight on the meaning of the observed structure. Spielman and Thill (2008) used a so-called backward approach to interpreting SOM results. They identify selected census tracts with known characteristics and pinpoint their locations on the SOM output map. Locations that have been assigned to the same SOM output node are then considered to share similar social characteristics. This approach implies that one must be highly familiar with the study area in order to select representative tracts. In addition, important patterns may also be neglected due to subjective choices of representatives.

To provide a more objective insight into the characteristics of identified SOM clusters, we propose a forward approach by incorporating principle component analysis (PCA) method. We apply PCA on the dataset to gain understanding of the major components in the attribute space. Once the clusters were identified using SOM, the interpretation of clusters involves the following two steps: 1) PCA is applied to extract major components in seeking to explain most variance of original data sets; and 2) for each identified cluster, a U-Matrix is computed to visualize the cluster structure for each PCA component. Figure 3 in the next section shows an example of the
four component planes for each identified cluster. Looking at the first cluster, for instance, the first component is shown to have the most homogenous pattern of low values suggesting little differences between neighboring nodes. Therefore, the first PCA component is regarded most representative of that cluster. A further look into component planes of variables within the first components is then needed to help label the cluster.

3.3 Results and Discussions

Spatial Pattern of Social Areas in 1990

For the 1990 dataset, SOM identified four clusters. Further PCA analysis found three major components, namely the education level, Hukou status, and household structure. Figure 3 shows the stratified U-matrices and Figure 4 shows the spatial pattern of the four clusters. By comparing the SOM U-Matrix generated for each PCA component, we can have more objective interpretation of clusters which further helped us label them. This is done for all the three time periods. In consideration of space limit of the paper, we will not list the U-matrices for the next two time periods here.
Cluster 1. *Area of Agricultural Population*. This category includes 245 contiguous sub-districts, mainly in outer suburb areas that consist of a primarily agricultural population. This area is characterized by low population density and low educational level.

Cluster 2. *Lower Density, Inner Suburb Area*. This category consists of 23 sub-districts primarily located between urban and outer suburb areas. These sub-districts form a belt-like shape surrounding the urban area and are characterized by lower population density in comparison with the urban core, and a higher education level compared to the outer suburb area.

Cluster 3. *Migrant Concentrated Area*. This area consists of 13 sub-districts dispersed in the outer suburb area, characterized by high proportion of migrants.

![Figure 3. U-Matrices show the degree of similarity between neighboring SOM nodes for each class by PCA components. The component 1 through 3 are identified as education level, Hukou status, and household structure respectively](image-url)
Cluster 4. *High Density, Urban Area*. This area consists of 114 sub-districts which are clustered in old urban core areas and inner suburb areas. These sub-districts share similar characteristics, such as high population density, higher educational level and a higher proportion of residents with registered “*Hukou*” status.

**Spatial Pattern of Social Areas in 2000**

Six clusters are identified from the SOM algorithm and six components are identified from the PCA analysis for the year 2000. The six components are professional workers, education level, agricultural population, living conditions of households, population with labor-intensive works,
and household structure. Figure 5 illustrates the spatial distribution of the six clusters. The identified clusters are labeled and interpreted below.

Cluster 1. *Inner Suburb Residential*. This type of social areas consists of 32 sub-districts. It is characterized by relatively high population density and employment ratio. They were mainly found between inner suburbs and outer suburbs. It is noteworthy that a small number of sub-districts in the central city also belong to this type of social areas.

Cluster 2. *High Density, Intellectual Area*. Major characteristics of this type of social areas include high employment ratio, high education level, and high population density. There are 56
sub-districts of them, primarily located within inner suburbs and the central city, indicating an expansion of this type of social areas from urban core to the suburbs. Many colleges, universities and professional jobs located in this area, which attracts a high concentration of intellectuals in this area.

Cluster 3. *Migrant-concentrated Area*. 26 sub-districts were categorized into this type of social areas, distinguishing themselves by high proportions of migrant population. These sub-districts distribute primarily in the fringe of inner suburbs, probably due to relatively lower housing price and reasonable job opportunities.

Cluster 4. *High Density, Commercial Area*. This type of social areas features high population density and high proportions of people engaging in commercial-related activities. Fifty-six sub-districts of this type dispersed in the central city and its periphery area.

Cluster 5. *Area of Agricultural Population*. With lower population densities and larger average living spaces, this type of social areas have high proportions of workers in agricultural related-fields and low education levels. These social areas concentrated in the outer suburb of Beijing.

Cluster 6. *Outer Suburb Area with Urban Residents*. Patterns within the PCA component of professional workers suggest that this cluster has reasonable share of professional workers besides the agricultural population residing in these regions. Fifty-nine sub-districts of this type concentrated in the southwestern outskirts of the city, while some are scattered in the northern outer suburb.

**Spatial Pattern of Social Areas in 2009**

Following the same process that was used for the two previous years, five clusters were found with SOM algorithm and five PCA components were identified for the year 2009 data. The five
PCA components are professional workers, economic status, education level, household characteristic, population movement and employment. Figure 6 displays the spatial distributions of the five clusters. Based the U-matrix of the five PCA components, the five clusters are labeled and interpreted below.

Cluster 1. *High Density, Intellectual Area*. Regions recognized as this type of social area are characterized by good employment opportunity and high education level. This type of social areas locate mostly in the central city and the inner suburb sub-districts.
Cluster 2. *High Density, Migrant Concentrated Area.* This category of social areas is characterized as having high proportions of migrants, high frequencies of movement, and rapid growth of employment opportunities. This social area includes 16 sub-districts within inner suburbs and two outer suburb sub-districts far away from the urban core.

Cluster 3. *Lower Density, Urbanized Suburb Area.* This type of social areas dispersed primarily in the outer suburbs. It is recognized as having lower population density, higher income and car ownership.

Cluster 4. *High Density, Commercial Area.* This type of social areas is found in several regions in the inner suburb edge. This type is characterized by high employment ratio and educational level, suggesting a concentration of industry-oriented land use type.

Cluster 5. *Outer Suburb Area.* Low population, low employment ratio and low education level were dominant characteristics in this type of social areas, which are labelled as areas of agricultural population. These sub-districts concentrated in the outer suburbs of the city.

**Discussions**

Comparing the social-spatial structure identified in years 1990, 2000 and 2009, it is apparent that dramatic changes had taken place in the city during the two decades of time. From year 1990 to 2000, transformation of social space has mainly occurred in two ways. First, the urban area identified in 1990 had separated into different social area types, indicating a functional separation within the area. New social areas, including a new high density commercial area and intellectual area of higher densities have emerged in the outskirts of the urban core and inner suburb areas, resulted primarily from the expansion of markets and the formation of CBD areas. Several outer suburb sub-districts, recognized as migrant concentration areas in previous
years, had been incorporated into these two categories as well, while new migrants concentrated more in the inner suburb belt in 2000. Another noticeable change is that several sub-districts in the urban core area present similar characteristics with those of the inner suburb area, which can be attributed to a suburbanization process that causes relocation of urban residents to suburb areas. Second, the predominence of a urban residents population was broken once the homogeneous outer suburb area, known for its concentration of agricultural population, had developed further. This also reflects the impact of suburbanization and the development of new satellite towns.

From the year 2000 to year 2009, suburbanization had further progressed. Transformation of social space can be seen to have unfolded in many ways including the following. First, the intellectual area had expanded, especially into suburb areas adjacent to the intellectual area prior to this time period. Second, the commercial area was enlarged, represented in both the expansion of previously dispersed sub-districts in inner suburbs and the newly developed areas of outer suburb to the north of the urban core. This change reflects continuing suburbanization and is also a result of improvement in transportation networks and services. Thirdly, in all directions, more outer suburb areas had been urbanized during this time period, but particularly concentrated in a clustered pattern in the southeastern part of the city. A noticeable change is the decrease of population density in the urban core area. Population growth, planning mechanisms, and dramatically increasing housing prices in the urban core forced people to move to the more affordable periphery of the city. At the same time, the expansion of transportation networks has allowed people to commute between the periphery of the city to the urban core, which makes living and working in different areas more feasible. On the other hand, some more wealthy people moved away from the city core. They choose alternative housing in the outskirts of the
city where they could have access to better living environments and avoid the crowdedness of the city. Increasing private car ownership also lessened the barrier of a long commute. A series of Olympic-related relocation projects also played an important role in this process. A lot of new cultural facilities, such as Olympic Game Park, had been constructed recently. Lots of old industry sites have been revitalized and some noxious industrial enterprises were moved to suburban area for the overall city imagery. Meanwhile, multiple instances of relocation have occurred in Beijing due to various reasons such as demolition of residential areas for pre-Games construction, forcing local residents to move out of the old urban districts and to planned construction projects.
4 Conclusions and Future Research

This study examines and tracks the changing social areas at fine spatial levels in the transforming Beijing. From 1990 to 2009, this study covers two decades of the post-reform years, in which the most dramatic urban transformation really took place. Particularly, by combining currently available city statistics with remote sensing data using geographic interpolation techniques, this study provides insights into the most recent changes of social spatial patterns in Beijing, reflecting the pattern after Beijing’s hosting the 2008 Olympic Games which stimulated far-reaching impacts on the municipality.

Our study indicates that Beijing’s urban growth has restructured Beijing’s social space with observed spatial patterns over time. The findings may have important implications to planning and public policy making. The change in social-spatial patterns over time highlights the emergence of new social area types and the formation of more disaggregated social-spatial structures. New types of social areas, which we labeled as the commercial area and the intellectual area resulting from increased economic activity, broke a once homogeneous social structure. Those two social areas represented high status zones where a considerable proportion of residents were well-educated professional workers. Our study shows an expansion trend of high status zones during the time, indicating a possible strengthened separation between the affluent and the poor. With transitions from more planned economy to more market-oriented economy, people with high social status will take advantage of redevelopment to upgrade, while lower social status population always lag behind. However, there are many important questions remaining to be addressed in future studies. For instance, how the restructuring has changed its resident’s lives? Whether and how will the restructuring impact its population differently for
different social groups? What are the implications to policy makers to maintain a socially sustainable Mega-City?

From the perspective of analysis techniques, a challenge worthy of further consideration is the inconsistency of census variables. Though general social spatial patterns could be observed, interpreted and explained, one limitation of this study is that the change of patterns through different time periods may be influenced by variation of the variables. This may cause faulty findings, particularly at a more precise spatial level.

References


