Identifying Spatial Structure of Urban Functional Centers Using Travel Survey Data: A Case Study of Singapore

Chen Zhong  
Future Cities Laboratory,  
Department of Architecture, ETH Zurich,  
8092 Zurich, Switzerland  
zhong@arch.ethz.ch

Xianfeng Huang  
State Key Lab of Information Engineering in Surveying, Mapping and Remote Sensing,  
Wuhan University,  
430079 Wuhan, China  
hwangxf@gmail.com

Stefan Müller Arisona  
School of Computer Engineering,  
Nanyang Technological University,  
639798 Singapore, Singapore  
arisona@arch.ethz.ch

Gerhard Schmitt  
Future Cities Laboratory,  
Department of Architecture, ETH Zurich,  
8092 Zurich, Switzerland  
gerhard.schmitt@sl.ethz.ch

ABSTRACT
Identifying the spatial structure generated by urban movements contributes to a better understanding of urban dynamics and is crucial to urban planning applications. Despite a number of studies concerning functional urban space, related research is still in a development phase, especially using emerging urban movement data. This study proposes a centrality index and attractiveness indices for detecting the urban spatial structure of functional centers and their spatial impacts using transportation data. The basic idea of these indices is to build a relationship between the activity patterns (distribution, density, and diversity) and urban form. Accordingly, measurements, spatial analysis, and clustering methods are presented. Taking Singapore as a case study area, we applied the proposed indices and measurements to travel survey data of different years, through which centers of urban activities as well as the changing urban form are detected and compared quantitatively. Our approach yields important insights into urban phenomena generated by human movements. It represents a novel way of quantitative urban analysis and explicit urban change identification.

1. INTRODUCTION
Urban spatial structure refers to the set of relationships arising out of the urban form and its underlying interaction of people, freight, and information [1]. The impact of various spatial structures on people’s daily life, urban transportation, economic growth, social equity, sustainable urban development has long been studied [1-3]. In the past decades, urban areas have gone through strong but heterogeneous sprawl. The contemporary cities are increasingly polycentric [4]. However, the term “polycentric” is a relative and fuzzy concept that needs further clarification [5]. It is necessary to develop a quantitative measurement for a better understanding of the process of polycentric urban transformation.

The spatial structure of modern cities was shaped, in large measure, by advances in transport and communication [2]. The improvement of transportation systems, the complexity of human movements, and the distribution of urban activities represent and cause the changing of urban functions and form. Therefore, we use spatial structure as a function of travel behavior to examine the interactions between humans and built environment. Specifically, this paper aims at identifying functional urban centers that are emerging from urban activities.

A wide range of method to identify urban centers has been proposed and can be shortly reviewed. [6] proposed a method using spatial distribution of density functions and considers the peak as possible sub-centers. [7] presented a way to use a set of reference thresholds (cut-offs) which evolves additional local knowledge. With the use of statistics, parametric method has been proposed using a regression model based on density and distance [8]. and Non-parametric method are introduced after based on a smoothed density function [9]. Later, the concept of centers, especially polycentricity, has been criticized and examined by comparing morphological and functional measurements [5]. Following the discussion of functional centers, new approach is presented to measure the connectivity of individual centers to the whole urban system through human flows [10]. This paper continues this discussion, and makes use of human movement data. We use travel surveys in this paper, but other data sources, like smart card data, can be also used as alternatives.

Only recently, data sets generated by human movements have been used for urban studies. For instance, estimating dynamic...
workplace capacities [11] and identifying urban activities [12] from smart card and surveyed data; discovering regions of different functions in a city using floating car data and point of interest data [13]; inferring land uses from mobile phone activity records and zoning regulations [14]; detecting and clustering different activities and places by spatiotemporal patterns using travel surveys [15]. To identify the urban spatial structures, in [4], large scale, smart card data (Oyster card data) of individual person movements in London subway is analyzed to identify the polycentric structure and organization of the city. Similar work have been done at different scales [16-19]. The urban forms detected from dynamic data, i.e. human mobility data, reflect the use, the accessibility, and functionality of urban space in reality, thus can uncover urban problems.

This work makes two main contributions: (1) we propose a centrality index, which builds the relationship between urban activity patterns (density and diversity) and urban form, an attractiveness indices that can be used for spatial impact analysis. With the detected centers and their spatial impacts, a clear urban spatial structure appears. (2) A case study of Singapore is conducted using surveyed data of two years to test the feasibility of the proposed indices and measurements. The results are compared to earlier urban plans as an evaluation. Furthermore, comparing the results using data in 2004 and 2008, it shows that even from such a short time series, Singapore is developing rapidly towards a polycentric form.

The rest of the paper is organized as follows. In the next section, we present the proposed indices and corresponding measurements. In Section 3, background information about our case study area and data is introduced. Section 4 gives a discussion of the results. Section 5 concludes the paper.

2. CENTRALITY AND ATTRACTIVENESS INDICES TO DETECT FUNCTIONAL ACTIVITY CENTERS

Central areas are defined to have a high level of spatial accumulation and corresponding land uses, such as retail, office, compared to peripheral areas [1]. Conventional measurements are applied to morphological urban elements. For instance, density and diversity, as well as entropy, are concepts that have long been used an index of land use [20]. However, it was shown in [21] that functional changes are not tied down to morphological changes. In this paper, centers that refer to the use of urban space in reality are defined and identified: We define functional centers as spaces where people are accumulated to perform certain activities. We develop a set of indices based on the conventional ones, the two key characters - density and diversity, to quantitatively measure the spatial structure of urban activities and movements using transportation data, which represents the functional urban space. We categorize our proposed index as follows:

a) City center detection based on diversity and density index.

The centrality index measures how central an area is in terms of attracting people to do various activities, and it is a combination of diversity and density functions.

b) Spatial impact identification of centers.

The size index measures how large the centers are geographically.

The coverage index measures the spatial influence of the center, defined by the average distance people are willing to travel to the center.

The attractiveness index measures the intensity of the force that attracts people to one center, and it is based on size and coverage index.

It is important to note that centers, stronger ones and weaker ones, are relative concepts, similarly to their spatial impacts. Levels of centers are achieved by ranking centrality indices of areas. Attractiveness indices can be further used to compare the spatial impact of a center in different years. Additionally, modern cities have various functional centers, such as economic, culture, information and so on. Our focus in this paper is to define a centrality index for detecting urban centers that have multiple functions.

2.1 Diversity and Density Index

Diversity is represented by entropy here, as entropy is a more quantitative index that describes the disorder of activities. The entropy index is a concept that originates from information theory, created by C.E. Shannon [22]. It has long been adopted to measure the degree of complexity and order of a land use arrangement [20,23]. In general, the smaller the entropy, the lower the disorders of the land use. In addition, a derived formula can measure the evenness of land use arrangement.

We start from the more general definition of land use entropy, because activity entropy is derived from it. We employ a regular grid to split the whole data set into cells, according to the geographical coordinates of X and Y directions. Given a geography space $S$ and a cell size, we can split the space into $m \times n$ cells. For a cell $(x, y)$ with $J$ types of land use, its land use Entropy index is defined as

$$ E_j(x, y) = \frac{1}{K} \sum_{i=1}^{K} \left( - \sum_{j=1}^{J} P_j(x, y) \ln(P_j(x, y)) \right) $$

Where $P_j$ is the proportion of land in the use type $j$ within a cell $(x, y)$, $K$ is the number of neighborhood cells, which is used to smooth the entropy value [20]. A single land use in a cell results a 0 value of entropy.

To find the cities centers that result from agglomerating intense and diverse urban activities, we reformulate the measurement for computing information of urban activities. We extend formula (1) for measurement of activities as

**Diversity index of urban activity types** measures how mixed the activity types in one unit area, where $P_j$ is the proportion of travels to cell $(x, y)$ for the activity type $j$ during a period of time (in our experiment, we set 24 hour as a temporal unit). $J$ is the number of number of different activity types considered.

Density is a very important and commonly used index for assessing land use patterns, which represent the intensity of general distribution in one area. In the paper, it measures the intensity of visiting. We define it as:

**Density index**, measured as the proportion of people accumulated in one unit area $(x, y)$ in $m \times n$ units space $S$.
\[ D(x, y) = \frac{N(x, y)}{\sum_{i, j} N(i, j)} \]  

(2)

Where \( N(x, y) \) is the number of people heading to unit area \((x, y)\).

As indicated before, diversity and density are two indispensable indices that relate to city centers. However, none of them can represent central areas individually. To demonstrate the misinterpretation, also the relationship between density and diversity, we give some typical examples. As shown in Figure 1 (1), there can be two areas, region A and region B, having the same level of diversity, but very different densities. Figure 1 (2) shows that there can be some non-central areas with high diversity of activity types and less visiting people.

![Figure 1. Examples of misinterpretation. (1) Region A and B with same level of diversity (entropy) but different density. (2) Region C has non-central areas with higher diversity than Region D.](image)

2.2 Centrality Index: City Center Descriptor

We consider the density index and diversity index as the attributes for each area, further data clustering or analyzing can be conducted based on these two-dimensional features. But a more proper solution should be given, instead of a simple linearly combination, because the diversity and density index are two quantitative values of different dimensions and physical meanings.

To cope with this issue, we propose to apply the convolution to combine these two indices. The proposed solution – convolution produces an integrated function that is typically explained as a modified version of one of the original functions, giving the area overlap between the two functions as a function of the amount that one of the original functions is translated. In general, Convolution is a mathematical operation used as functional “add”. Given two time sequential functions \( f_1 = f_1(t) \) and \( f_2 = f_2(t) \), \( f_1, f_2 \) are the signal energy at time sequence \( t \). If these two signals “add” together, the new time-energy function \( f(t) \) will be the convolution of \( f_1 \) and \( f_2 \), as shown in equation (3):

\[ f(t) = \int_{-\infty}^{\infty} f_1(\tau)f_2(t-\tau)d\tau \]  

(3)

To write in a simple way, equation (3) is often denoted as

\[ f(t) = f_1(t) * f_2(t) \]  

(4)

However, the problem here is that the two attributes of an area, density and diversity (entropy), are still not in the same dimensions. Thus, they cannot be used directly in equation (4). Therefore, we employ probability theory to combine the density and diversity to describe the centrality index.

In a two dimensional \( m \times n \) space \( S \), we denote the density function \( D = \{D_{ij}\} \), with \( x = 1,2,..m \), \( y = 1,2,..n \). \( D_{ij} \) is the density of cell \((x, y)\) in \( S \). For each cell, there will be a function \( P_1(x, y) = f(x, y, D_{ij}) \) to denote the probability of a cell to be a city center of, based on its density. According to our hypothesis, the higher density value of a cell \((x, y)\) implies a higher probability of this cell to be an urban center. In this paper, we assume that the cell with the maximum number of visitors has the maximum possibility to be a center. So we define the probability density function as

\[ P_1(x, y) = f(x, y, D_{ij}) = \frac{D_{ij}}{\text{sum}(D)} \]  

(5)

Meanwhile, \( P_2(x, y) = g(x, y, E_{ij}) \) is a probability density function related to diversity \( E_{ij} \) at cell \((x, y)\), representing the probability of this cell to be a city center with diversity \( E_{ij} \). Similar to \( f(x, y, D_{ij}) \), we define the probability density function of diversity as

\[ P_2(x, y) = g(x, y, E_{ij}) = \frac{E_{ij}}{\text{sum}(E)} \]  

(6)

As the density of people and diversity of activities are two independent events, therefore, the joint probability density function of the two independent events is the convolution of the two probability density functions. Therefore, we can define the centrality index by a convolution of the two independent indices.

Centrality index : \( C_w \) measures the probability of an area \((x, y)\) to be the center of city.

\[ C_w = P_1(x, y) \cdot P_2(x, y) \]  

(7)

Here we use centrality to find out the central areas which has both comparatively higher diversity and density. From the formula we can see, this centrality index is the possibility of this cell to be a center, which derived from the density of people and diversity of land uses.

To be noticed, the probability functions we defined here are in simple forms. Those functions are based on our hypothesis for functional activity centers that have high density and high entropy. These functions need more deep investigation in the future, and can be changed according to a further refined hypothesis within our proposed framework.

2.3 Attractiveness Indices for Spatial Impact Analysis

Centers are attractive areas where a large population travels to carry out various activities such as shopping, eating, leisure. There are also sub centers surrounding the main central business district. Those centers exert different levels of importance and functionalities in the overall urban space. Attractiveness indices are defined to quantitatively compare the significance of centers. It is defined in this paper considering the spatiotemporal dimensions of people being drawn to one area.
The size index measures the area of centers which are at a given level of centrality.

The coverage index measures the spatiotemporal impact of the area, computed as the expected travel distance and arrival time of trips that travel to a center.

The attractiveness index measures the intensity of the force that a center can attract people.

A demonstration of the physical meanings of the indices is shown in Figure 2. Research has been done on this aspect, among them, the most significant one are central place theory and spatial gravity theory. In this paper, we derive only a simple measurement here considering the intensity of travels, and taking density of traveling people, travel distance as indispensable factors, we define the attractiveness index as follows:

\[ A(x, y) = \frac{D(x, y)}{\text{Coverage}(x, y)} \] (8)

where \( D(x, y) \) is the sum of trips that lead to a center \((x, y)\); \( \text{Coverage}(x, y) \) is the expected travel distance to center point of \((x, y)\).

3. DATA SOURCE AND STUDY AREA

We study the area of Singapore as an example. Singapore is an island city-state in Southeast Asia with an area of 710.2 km\(^2\). The current population of Singapore including non-residents is approximately 5 million. In our experiment, surveyed data – the so called Household Interview Travel Survey (HITS), is used as input. The survey is conducted by the Singapore Land Transport Authority (LTA) every four to five years to give transport planners and policy makers insights into residential travelling behaviors. About 1% of households in Singapore are surveyed, with household members answering detailed questions about their trips. The HITS data provide information including age, occupation, travel purpose, travel destination, walking time, waiting time, travelling time and so on. A report of more detailed analysis of HITS results can be found in [24]; In order to compare the changes of spatial structure, this paper uses the HITS results of 2004, which contain 51,000 validated records after our data processing and HITS results of 2008, which contain 76,923 validate records after data processing. In addition, we use the Singapore land use plan of 2008 for a comparison.


We applied our approach to surveyed data of 2004 and 2008. The results are discussed in three parts. The first part shows the relationships between diversity, density and centrality and how our defined calculation works. The second part identifies the polycentric urban form from detected centers and compares it with concept urban plan. The third part compares the results of different years. It shows that our indices provide a way of explicitly representing urban process.

4.1 Diversity, Density and Centrality

Instead of using the half-mile radius unit used in [20], we use 500 meters as the cell size to split the whole city, which is an approximate average walking distance to transportation infrastructure according to [25] and corresponds with statistical results of the surveyed data. Moreover, aggregation is also a way to safeguard individual private information. Instead of precise building-to-building mapping, we use a cell-to-cell mapping. For a fair comparison of results in 2004 and 2008, we define the activity types based on the categories given by the original surveys. We define five types of activities, namely going home, working, studying, shopping, eating, and social visiting. Social visiting includes entertainment, public service, and so on; we group those social visiting activities together, because a total of these activities take a small portion of the overall activities.

The results of diversity, density, and centrality mapping of Singapore of 2008 are shown in Figure 3. There are some incompatible diversity and density patterns of activities clearly shown in some of the areas, like the one marked by rectangles in (a) has a peak point, while in (b) contains only extremely low values. (c) is our result of centrality which is a combination of density and diversity. The density values follow a nearly exponential distribution, but not strictly. While, the distribution of centrality index follows the exponential curve very well.

![Figure 3. Diversity (or entropy) (a), density (b), and centrality (c) of urban activities in Singapore using surveyed data of 2008. X,Y axes represent geographical coordinates, and color shows normalized values.](image-url)
4.2 Polycentric Urban Form of Singapore

Figure 4 shows a centrality map of Singapore in 2008 with a three dimensional visualization. For a better visual interpretation, the centrality index is used as the height of the area. The original 2D mapping can be referred to Figure 3 (c). As indicated before, big centers and small centers are relative concept, that centers are identified by a ranking mechanism. Shown in Figure 4, centers at different levels of centrality are marked out. As an example, eight levels are classified based on natural groupings inherent of the centrality index. The break points are identified by picking the class breaks that best group similar values and maximize the differences between classes.

![Figure 4. Emerging polycentric urban form of Singapore.](image)

The level of centers (from 1 to 8) is defined by ranking the centrality index. From red to blue, eight levels are defined. Ellipses mark the centers according to their level.

After detecting those different levels of centers, the detected centers are marked with small and big ellipses where a clear polycentric urban form appears. As previously indicated, contemporary cities are complex, displaying a polycentric urban form [4]. The revised Singapore concept plan of 1991 emphasized on facilitating sustainable economic growth, and proposed the idea of decentralization. A city is planned to be surrounded by several regional centers, sub centers, and fringe centers. Polices have been carried out to intensify and expand the old city centers. Regional centers are developed to decompose human flows by shortening the distance between work and home locations. In Figure 4, the biggest center in the middle part of Singapore is downtown. It was planned even in the earliest urban plans. The four significant sub centers are: Jurong area in the east region, Tampines in the west region, Woodlands in the north region, and Seletar in the north-east region.

This result is in line with Singapore’s essential planning concept. To some aspect, this consistency of land use and activity patterns reveals the compatibility of transportation planning and land use planning in the case of Singapore. However, we also found that other emerging sub centers such as that Yishun area has a higher centrality than the planned sub centers. To some aspect, this abnormal phenomenon is an evidence of the unpredicted bottom-up changes which reshaped the urban structure in reality.

4.3 Measuring Urban Changes

A polycentric urban transformation takes years and decades to come true. In the case of Singapore, the idea of decentralization was initiated in the 1990s and the process is still going on. During a polycentric urban transformation, the functions of a highly centralized central business district (CBD) area have been gradually shared by sub centers; meanwhile, urban stock and flows are re-distributed. How is the urban process evolving in Singapore? This question can be answered by comparing data of different years.

A comparison using attractiveness indices is shown in Table 1. We count the centers with highest level of centrality. We found that the number decreased from 5 to 4 in five years. The spatial impacts of the bigger centers grew bigger while that of smaller ones were shrinking. This is shown by the increasing center size variance, which measures the difference of size of centers. We calculated the coverage of the biggest centers. It shows that travel distance is becoming longer. Furthermore, the number of travels to bigger centers as well as the density all increased. This increasing travel demand happened might because of the increasing population and activity demands. However, the global distribution of centrality index shows that the spatial impacts of biggest centers are increasing instead of decreasing and being shared by sub centers. Wherever people live, they choose to go to the biggest center instead of sub centers. This might because of the advancement of transit systems, which makes it more and more convenient to travel to the big centers, which provide more choice of services.

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of first level center</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Avg. center size variance</td>
<td>1.84</td>
<td>2.75</td>
</tr>
<tr>
<td>Coverage of the biggest center</td>
<td>10103.6</td>
<td>10149.2</td>
</tr>
<tr>
<td>Avg. number of Travels to biggest center (per 500m cell)</td>
<td>123</td>
<td>210</td>
</tr>
<tr>
<td>Avg. density of traveling to biggest center (per cell)</td>
<td>0.24%</td>
<td>0.27%</td>
</tr>
<tr>
<td>Attractiveness index</td>
<td>0.0121738</td>
<td>0.02096</td>
</tr>
</tbody>
</table>

To summarize, it shows that even from such a short time series, Singapore is developing rapidly towards a polycentricity. The results of our experiment show the effectiveness of our approach as a way to quantitatively identify the central area of urban activities and explicitly detecting urban changes.

5. CONCLUSION AND FUTURE WORKS

This paper explores the way of making insights into urban dynamics with a quantitative measurement of spatial structure using urban movement data. In particular, we propose centrality index and attractiveness indices for detecting and comparing centers urban activity and their spatial impact from travel survey data. It is a way to integrate multi-dimensional information and extract urban knowledge. Taking Singapore as an example, we used surveyed data of different years to verify the effectiveness of our indices and measurements. The quantitative approach and the results can be used as references for more explicitly representing urban dynamics to support urban plan applications.

There are still many areas to further develop in this research. First, the usage of these indices are not limited to surveyed data, but to other mobility data set such as smart card data. More experiment will be conducted in the future using multiple data sources. Second, the indices can be used not only for detecting urban activity centers but can also be further derived for detecting other
functional centers, like education centers, shopping centers. Third, we did not do a comprehensive exploration of the temporal dimension. Identify the changes along temporal dimension will be valuable for optimizing urban resources. Finally, the probability functions of diversity and density are rather simple, and whether they can strictly satisfy real situations is a problem that needs to be investigated. In the future, more parameters should be considered to define the probability functions.

In sum, this work yields important insights into urban analysis using urban sensors, like transportation data. The method as well as the results could be references for the later research on applying spatial data analysis to urban study.

ACKNOWLEDGMENTS
This work was established at the Singapore-ETH Centre for Global Environmental Sustainability (SEC), co-funded by the Singapore National Research Foundation (NRF) and ETH Zurich. The authors would like to express their sincere gratitude to the Singapore Land Transport Authority and Urban Redevelopment Authority for supporting this research.

6. REFERENCES