Spatial Data Analysis of Complex Urban Systems

Farideddin Peiravian, Amirhassan Kermanshah, and Sybil Derrible
Complex and Sustainable Urban Networks (CSUN) Lab
Department of Civil and Materials Engineering
University of Illinois at Chicago
Chicago, Illinois, U.S.A. 60607
fpeira2@uic.edu

Abstract—Cities are complex systems that are constantly evolving to better allow people to connect with one another. Moreover, and similar to countless natural phenomena, cities exhibit inherent orders that can be captured and expressed through complex analyses of their components. Using a variety of large datasets, this work offers a ring-buffer approach to analyze the spatial characteristics of four components of Chicago urban system, namely: roads, intersections, buildings, and population+employment. The complex nature of these four components manifests itself in power-law relationships, represented by their fractal dimensions. Results show that road length and number of intersections, and to a larger degree, population+employment count and building gross floor area exhibit significantly similar properties. The proposed method could further be used to analyze large demographic, socio-economic, and other geospatial datasets with the aim to study their impacts on relevant urban systems characteristics, including mobility, connectivity, and accessibility to name a few.

Keywords—spatial data, GIS, complex analysis, urban systems, transportation networks

I. INTRODUCTION

The evolution and spread of an urban system and its components; whether it is its transportation network, or buildings, or even distribution of people themselves, happen over many years. It is the aggregated outcome of numerous individual and collective choices, each influenced by the prevailing conditions in its time. Each new change is overlaid on previous changes. In other words, any urban system and its components have a starting point when and where they were founded; tens or in some cases hundreds of years ago.

While it seems reasonable to assume that the older a city is, the less coherent its founding blocks have been, many researchers suggest [1]–[3], and sometimes demonstrate [4]–[8], that no matter how an urban system has evolved or what foundations it has been built on, from a larger perspective it has inherent order and organization. Having said that, and to better understand the complex nature of an urban system, studies have been focused on the characterization of its components rather than itself as a whole [9], because “understanding the topology of urban networks that connect people and places leads to insights into how cities are organized” [10].

The hidden, and presumably orderly, characteristics of different components of a given urban system have been a matter of interest in recent time [11]–[13]. In the case of the transportation network of a city, one can visually observe that such an order manifests itself in a self-similar pattern [14], [15]. In other words, the evolution of a transportation network is very similar to a tree trunk that grows, then splits into branches, and then each branch grows and then that also splits into more smaller branches, and so on so forth. One main difference, though, is that transportation networks create loops through branch-joining. Additionally, and noteworthy, order can manifest itself by showing similar shapes and patterns even if scales differ. This is particularly true in road networks that tend to be denser in downtowns, while keeping the same overall pattern of intersections throughout a city.

Nonetheless, the self-similar characteristic of urban systems is not restricted to their transportation networks. In fact, the spread of other components of an urban system, such as population [16]–[18], employment, and buildings [3], [19], can show self-repeating patterns as well. This phenomenon clearly fits within the realm of a branch of complex system analysis, i.e. fractals.

Indeed, as spatial and non-isomorphic systems with self-repeating patterns, cities clearly exhibit the presence of fractal entities. While this complex behavior of cities and their components has previously been studied [1], [10], new technologies and more disaggregate datasets, and in particular powerful Geographic Information Systems (GIS) tools as well as extensive geospatial data, allow for a more detailed and comprehensive inspection and analysis towards a better understanding of complex urban systems and characterization of their components.

Based on the above discussion, the objectives of this work are to: (i) analyze the characteristics of different components of a given city using a proposed ring-buffer fractal approach, (ii) determine the similarities and differences between the complex representations of those components, and (iii) explore and explain the reasons behind such similarities and differences.

Overall, this work fits within the global endeavor to analyze cities and their infrastructure as complex systems [7], [20]–[27]. Taking a fractal approach to analyze an urban system and its components offers many benefits, including the provision of a measurable metric. As we will see in the next section, each fractal possesses a particular dimension. As a result, although different components are not directly comparable (e.g., people versus buildings), comparing their fractal dimensions offers a pragmatic means to gage how they coexist and interact within the built environment.

DOI: 10.1109/BigData.2014.7004405

978-1-4799-5666-1/14/$31.00 ©2014 IEEE
II. METHODOLOGY

A. Definition

A fractal can be described as an entity that possesses self-similarity on all scales. It is important to note that a fractal needs to only exhibit similar (but not exactly the same) type of structure at all scales. Moreover, according to Mandelbrot [28]: "A fractal set is one for which the fractal dimension strictly exceeds the topological dimension." In practice, this means that while a line feature (e.g. a road) has a dimension of 1 in classical geometry, it must have a dimension larger than 1 if it is to have fractal properties.

The rough description of fractal dimension (as used in this work) of a fractal object is the exponent in the expression of the form shown in Eq. 1 [29], [30]:

\[ N(r) = a \cdot r^D \] (1)

in which \( r \) is the radius (with respect to a point of origin or center), \( N \) is the number quantifying the object \( x \) under consideration at the radius \( r \), \( a \) is a constant, and \( D \) is the fractal dimension.

The proposed ring-buffer method is based on the assumption that urban systems, and their components, evolve similar to living organisms. A living being comes to life as a single cell; let’s call it the “center”. Then it grows and spreads around that center, subject to its prevailing conditions and constraints. Similar to that, a city spreads around a point of origin, or “center” [30], [31], and then gradually expands outwards, while avoiding the physical constraints around it such as water bodies, valleys, etc. The assumption is that the spread of any component of the system, e.g. its road network, at a given point is proportional to its distance from that center.

Having determined its “center”, the urban system can then be split into rings around it, as shown in Fig. 1. The “center” could be a point, or a small area, around which the urban system has grown and evolved. By calculating the quantity of any given component of the system as a function of the distance from the center, one should be able to verify its fractal nature (if any) and extract its fractal dimension [32].

To demonstrate a fractal property, a power law relationship must be present between the quantity of the component under consideration and the radius, as discussed in Eq. 1. Nevertheless, because certain areas have to be excluded (i.e. lakes, rivers, airports, etc.), the ring areas differ from each other significantly. As a result, and instead of the quantity of the component within each ring, its cumulative density within the entire circular buffer is calculated (i.e. Eq.1 is integrated and then divided by the area). Consequently, and as shown in

\[ \log [\text{Density}(r)] = \log(b) - (2 - D) \cdot \log(r) \] (3)

i.e.: \( D = 2 + \text{Slope} \)  (Note that the slope is negative)

The regression technique used for Eq. 3 is sufficient in this case since there are only few data points; the reader is referred to [32] for a further discussion regarding statistical methods that can be used to fit power laws.

Finally, using this proposed ring-buffer approach, a question naturally arises about the selection of the center of a city, especially in cases of mono- versus poly-centric urban forms. This is an ongoing debate that does not have a definite answer to date. It is, however, irrelevant in our case since the methodology is applied to Chicago that has a well-defined center, as it will be seen in the next section.

III. APPLICATION

A. Case Study

As a case study, we attempt to investigate whether the spread and evolution of the components of one of the oldest cities in North America, i.e. Chicago, has an inherent fractal nature. Having verified the hypothesis, Eq. 3 will then be used to find the fractal dimensions of its components.

What made Chicago a unique choice was not only its long history during which it had experienced different periods of urban evolution, but also its unique morphology. Chicago is restricted on the east side by Lake Michigan (Fig. 2), which means it has only been able to expand towards the west. Moreover, two branches of Chicago River run through it from north and south, which join together to the west of the center of the city and then run eastward towards Lake Michigan. Such natural constraints on the evolution of its urban system, in addition to the man-made barriers such as its two international airports, offer intriguing challenges to the process of applying the proposed ring-buffer approach.

Fig. 1. Rings creation in the proposed method.

Fig. 2. Location Map for Chicago, Illinois (Background: Bing Maps Hybrid).
Moreover, Chicago has a well-defined center, called the “Loop”, which is Chicago’s Central Business District (CBD). It hosts the Chicago Mercantile Exchange, as well as being the city’s administrative center. It is a well-defined 1.0 km x 1.2 km rectangle surrounded by freeways and partially by the Chicago River (Fig. 3).

Fig. 3. The “Loop”, CBD of Chicago (Background: Bing Maps Hybrid).

For that reason, the “Loop” seemed to be a natural choice for the “center” (or the “heart”) of the city. Having done so, rings were then created around it with radii from 1 km to 21 km, in increments of 1 km (Fig. 4). As shown, the physical or natural barriers to the city’s expansion, such as the lake, rivers, valleys, airports, etc., have been cut out of the rings.

Fig. 4. Rings, splitting Chicago into rings from the center.

As for the urban system component to be studied, the road network is a natural choice, because it facilitates the flow of people and goods. That being said, several attributes of roads can be studied. The first attribute to consider is the road length, due to the fact that roads are not only visually similar to fractals, but also that they act as flow channels that enhance mobility and accessibility. As a second attribute, the intersections within the road network are hypothesized to have a fractal nature, as they are directly related to the roads.

A less obvious choice was the building area within the city. The rationale was that the construction of buildings and their spread resembles the expansion of living organisms. They start from a central area, and then spread and expand in different directions while avoiding natural, as well as man-made, barriers. Moreover, buildings are either the origins or destinations for a majority of the trips within a city (as places of residence or work for instance). To accurately capture the spread of buildings both horizontally as well as vertically, and since their footprints alone could not capture the vertical dimension, total building gross floor areas (equal to the footprints multiplied by the corresponding numbers of floors) were chosen as another urban system component to study.

Finally, population+employment attributes were selected as the last candidate for fractal analysis. This means that population and employment numbers were summed together to account for the fact that the downtown tends to host few households but many jobs, while the opposite is true for the suburbs. The rationale for including population+employment is that the buildings (i.e. the supply) are constructed to meet the demand, and this demand is mostly (though not completely) generated by the needs to live and work. Because of that, similar properties between roads, intersections, building gross floor areas, and population+employment are expected.

B. Data

The data for this study was obtained from different sources, as shown in Table I.

<table>
<thead>
<tr>
<th>Data</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road Network</td>
<td>U.S. Census Bureau, TIGER/Files</td>
</tr>
<tr>
<td>Census Tracts and Population Data</td>
<td>U.S. Census Bureau, American FactFinder</td>
</tr>
<tr>
<td>Building footprints, Land-Use, and Employment Data</td>
<td>Chicago Metropolitan Agency for Planning</td>
</tr>
</tbody>
</table>

Having obtained the data for each of the four urban system components, their quantities in each ring and the areas of the rings are calculated. That information is then used to calculate the densities of those components for buffers around the center at the selected radii (Fig. 5).

Fig. 5. Urban components within equi-distance rings around the Center: a) Road Length, b) Intersections, c) Population+Employment, and d) Building Gross Floor Area

IV. RESULTS AND ANALYSIS

A. Results

In order to come up with the metrics that are comparable at different radii, the density (Eq. 2, as opposed to the count $N$ from Eq. 1) of each component was calculated for buffers at the radii from 1 km to 21 km at increments of 1 km. The results are plotted in Fig. 6. The diagrams provide an opportunity to observe the variations in the overall cumulative (average) value of any given component as a function of the distance from the center.
Fig. 6. Plots of components densities vs. radius: a) Road Length, b) Intersection, c) Population+Employment, and d) Gross Floor Area

The initial observation is that all the above plots seem to exhibit power law properties, though not at the same level. As mentioned, the presence of a power law is required to prove the presence of fractal characteristics. The road length density and intersection density plots demonstrate similar patterns, both following a downward trend at a slow pace. This means that the percentage changes in the densities become less and less sensitive to the changes in the radius.

Furthermore, the population+employment density as well as the gross floor area density figures seem to show even more similar patterns. Indeed, from Fig.s 6c and 6d, density values are expectedly high at the center (i.e. small radius), but then drop sharply as one moves away from the center. Moreover, the rates of change of the slopes for both plots gradually start decreasing, demonstrating the reduced sensitivity of the population+employment and the gross floor area densities with respect to the distance from the center. This is also expected because the further away from the center, the less percentage change in the radius.

Comparing the four figures, one can conclude that the concentration of road segments and intersections within this urban system is less sensitive to the distance from the center as compared to the population+employment and the gross floor area densities with respect to the distance from the center. This is also expected because the further away from the center, the less percentage change in the radius.

To further study the potential fractal characteristics of these four components, and to determine their fractal dimensions, the same data are redrawn on log-log plots, that transform power laws into straight lines (Fig. 7).

As a mega city, Chicago has absorbed many smaller urban areas within itself during its evolution history, and therefore changes are expected in the patterns observed over the selected range of the radii, i.e. 21 km. The log-log plots indeed confirm this expectation, as there are break points in the linear trends in the plots, which one can observe in Fig. 7.

Moreover, certain features of road networks, such as highway ramps and interchanges, can sometimes artificially inflate the results, which happens here around 4-5 km. Notwithstanding, the fractal dimensions for the chosen components of the city were extracted using the piece-wise linear patterns observed in the log-log plots. Also, statistical analyses were performed on the data and the results are presented in Table II. The $R^2$ and t-stat values show that all the results are statistically significant.

### Table II. Fractal Dimensions of Components Densities

<table>
<thead>
<tr>
<th>Component</th>
<th>Radii (km)</th>
<th>Fractal Dimension</th>
<th>Std. Dev.</th>
<th>$R^2$</th>
<th>t-stat</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road Length</td>
<td>3 to 10</td>
<td>1.21</td>
<td>0.02</td>
<td>0.98</td>
<td>13</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>11 to 21</td>
<td>1.09</td>
<td>0.01</td>
<td>0.95</td>
<td>11</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>12 to 20</td>
<td>1.19</td>
<td>0.03</td>
<td>0.98</td>
<td>13</td>
<td>Yes</td>
</tr>
<tr>
<td>Intersection</td>
<td>3 to 10</td>
<td>1.19</td>
<td>0.01</td>
<td>0.95</td>
<td>13</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>11 to 21</td>
<td>1.36</td>
<td>0.02</td>
<td>0.98</td>
<td>22</td>
<td>Yes</td>
</tr>
<tr>
<td>Population+Employment</td>
<td>11 to 21</td>
<td>1.73</td>
<td>0.03</td>
<td>0.98</td>
<td>22</td>
<td>Yes</td>
</tr>
<tr>
<td>Gross Floor Area</td>
<td>11 to 21</td>
<td>1.70</td>
<td>0.02</td>
<td>0.99</td>
<td>47</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### B. Analysis

In Fig. 7, the log-log plots of both the road length and intersection densities show mild linear relationships with respect to the radius. Although there are still visible linear trends in both diagrams, they each show three parts with three different slopes. The first parts of both plots only consist of the first two points, which are within the first 2 km radius from the center. A first plausible explanation is that within 1-2 km radii, areas are small and the presence of a large park (i.e. Grant Park) affects the road density. A second plausible explanation favors the idea that, unlike the rest of the city, strong top-down planning decisions, as opposed to self-organization, were taken in that area due to its commercial importance. A third, and perhaps more likely, explanation points to the fact that the road system may have reached a point of saturation. In other words, downtown Chicago has reached its full horizontal capacity during its evolution, and its expansion has had to switch almost completely from horizontal to vertical direction (e.g. by building skyscrapers). On the other hand, road networks, unlike buildings, are restricted to two dimensions. Due to that, there could be no more new intersections or roads to build.
In contrast, both the log-log plots of population+employment and gross floor area densities show very well-defined linear relationships that start from the center and moves outwards up to the radius of 10 km, after which the slopes of both curves change.

Interestingly, the location of the change corresponds to the boundary of old Chicago City with old Cicero Township, which are now completely merged. This fact explains the sudden, yet similar, changes in the trends of all log-log plots in Fig. 7. Moreover, the diagrams show that the rates of changes of both densities have slowed down beyond the 10 km radius, which reflect the fact that after that point the sensitivities of the population+employment and gross floor area densities with respect to the distance to the center of the city have fallen, i.e. people and businesses are less reactive to a slight change in the distance to the downtown area. Nevertheless, the linear trends of both diagrams continue after the 10 km radius, though with different slopes.

Returning to Fig. 7, other than the first parts of both log-log diagrams of the road length and intersection densities, the rest of their plots show patterns similar to each other, as well as to the log-log plots of the population+employment and gross floor area densities, in that they follow two distinct linear trends with a separation point at a radius of around 10 km, thus capturing the hidden boundary between the old Chicago city and the old Cicero Township. Moreover, the decreasing slopes of the two plots represent the fact that the further one moves away from the center of the city, more infrastructure per capita (but less overall) will be required to accommodate the decreasing population.

In order to further investigate the similarity between the road and Intersection densities on the one hand and the population+employment and gross floor area densities on the other hand, the corresponding values were plotted against each other, shown in Fig. 8.

For the road and Intersection densities, the two components show a strong statistical relationship, with $R^2$ of 0.997. With respect to the population+employment and gross floor area densities, one would appreciate that their fractal dimensions are very close. This observation points to the high degree of similarity between these two different-in-nature components of the city. In fact, a plot of population+employment versus gross floor area densities show a statistically significant linear relationship between the two components, with a significant $R^2$ value of 0.997.

VI. Discussion

The trends observed in the road length and intersection densities were similar, and the values obtained for their fractal dimensions are also fairly close. On the other hand, the difference in their fractal dimensions in the 1-10 km range may suggest that the road construction capacity has been reached (as experienced every day with severe congestion), but further investigation is needed to confirm this hypothesis. That being said, the fractal dimensions calculated for the intersections are slightly and consistently higher than the corresponding fractal dimensions found for the road lengths. An explanation is that by moving away from the center of the city, the average land parcel size increases, therefore leaving less space for roads and intersections. This is particularly true for the number of intersections, since the average block size also tends to increase in suburban areas.

The two components population+employment and building gross floor area showed strong similarities with one another. In fact, their fractal dimensions are close, especially in the 11-21 km range. Although a further investigation is necessary, this result seems to attest the presence of an equilibrium between the supply (i.e. buildings) and the demand (i.e. population+employment) within that range.

The patterns of all diagrams therefore correspond to the fact that the further one moves away from the center of an urban system, the less dense it gets in terms of population and employment, building gross floor area, and transportation infrastructure (roads and intersections), something which is expected. Moreover, their rates of change, expectedly, are higher at the beginning but slow down quickly as the distance increases. This, again, is the exhibition of the power law, which is a representative of fractal behavior.

Overall, the results strongly support the hypothesis that the four components considered (population+employment, gross floor area, road length, and intersection densities) in the city of Chicago are fractals in nature, as demonstrated by the presence of power law relationships.

Nevertheless, not all fractals are exactly the same, which means that each component of an urban system possesses its own characteristics, including its own fractal dimension. Sudden changes in the behavior of these fractal entities can enable one to identify where the inherent characteristics of a system have changed. This could be a clue to the causes behind such changes, which can then be used to identify the shortcomings or deficiencies of the system.

In the case of the city of Chicago, despite its old history, all of its chosen components show fairly similar patterns. For example, they all show a change in the characteristics of the city at its old boundary. This could be used to identify hidden underlying attributes of an urban system.

VI. Conclusion

The work presented in this article offers a simple yet efficient fractal approach to the identification, analysis, and comparison of the characteristics of complex urban systems. The proposed ring-buffer method is able to expose the hidden
features of a city, even in an old and diverse city such as Chicago with its unique physical and topological barriers.

The study was able to achieve its objectives, namely: to analyze the characteristics of several components of the Chicago urban system, i.e. roads, intersections, population+ employment, and gross floor area, using the proposed fractal approach; to determine the similarities and differences between the fractal representations of those components; and to explore and explain the reasons behind such similarities and differences.

A further expansion of this work could be the application of the method used in this study to other cities, so that the results can be compared and analyzed and the findings can be used to improve the quality of the observations and conclusions from this fractal approach. As a starter, the same method could also be tested on polycentric cities.

One can also use the proposed method to study the spread of urban systems characteristics that affect the flow of people and goods, such as mobility, connectivity, and accessibility to name a few.

Further work could also focus on the inclusion of more components of urban systems, such as their utility networks (gas, electricity, water, etc.). Moreover, the proposed method can be used to explore the evolution of urban complex systems through time, i.e. temporal analysis, which could help in following their evolutionary paths and analyze the impacts of natural or man-made events on how they are shaped today.

Overall, this approach could potentially be used towards a better understanding of how a city, as a complex system, works and how its intertwined components can be analyzed, characterized, and eventually improved.

REFERENCES