Fractal or scaling analysis of geographic forms and functions

Ding Ma
University of Gävle
Kungsbäcksvägen 47, 801 76, Gävle, Sweden
dingmartin1@gmail.com

Abstract

The geographic space and phenomena are inherently heterogeneous and diverse both geometrically and statistically. However, current geospatial analysis lacks ability to reveal such underlying heterogeneity, as it relies on Euclidean geometry and normal distribution statistics which focus on geometric details (such as locations, directions, and sizes) that are with a well-defined mean and small variance. This study adopts a new paradigm for geospatial analysis including topology without geometric details, fractal geometry, scaling hierarchy, and heavy-tailed distribution statistics, particularly in the big data context, for better understanding geographic forms and functions.
Keywords: fractal geometry, scaling, topology, big data

1 Introduction

The aim of this study is to better understand the underlying fractal structure of the geographic forms and functions. Geographic forms refer to how geographic space looks; i.e., the geometry of geographic space at various levels of resolution, including cities, countries, and the entire world. Geographic functions refer to how geographic space works, indicating the impact of geographic forms on human activities within the space. Geographic space or the Earth's surface is neither smooth nor regular. It is generally recognized that the Earth's surface is diverse and heterogeneous (Goodchild 2004), which consists of far more small geographic features than large ones. However, geographic information science and systems confront challenges in understanding the instinct heterogeneity of the geographic space because the traditional geospatial analysis is founded on and dominated by Euclidean geometry and Gaussian statistics (Jiang 2015). As is well known, Euclidean geometry and Gaussian statistics are only good at handling respectively regular shapes and similar things, but the shapes of all meaningful geographic features (e.g. mountains, rivers, trees) on the Earth's surface are irregular and their sizes are quite dissimilar to each other. In these respects, the study adopts a new paradigm based on the fractal geometry and Paretian statistics for geospatial analysis.

2 Fractal, scaling, and topology

The fractal geometry represents rough and infinitely heterogeneous shapes. Developed by Benoit Mandelbrot, it denotes the power law relationship between details and scales (Mandelbrot 1982). Not limited to this framework, the study relies on the new definition of fractal (Jiang and Yin 2014): a set or pattern is fractal if the scaling of far more small things than large ones recurs multiple times. Based on the new definition, scaling and fractal are interchangeable. Paretian statistics is within the new definition, referring to far more small things than large ones, and should mathematically be defined as heavy-tailed distributions. The typical heavy-tailed distributions are power law distribution, exponential distribution, and lognormal distribution. Hence, the fractal or scaling structure of the geographic space implies the heavy-tailed distributions of the geographic features.

To enable us to see the fractal or scaling structure, the topological perspective or topology of the geographic features works as the key prerequisite (Jiang 2013a). Topology refers to the properties that remain unchanged under the distortion of geometric space. In other words, topology focuses on the relationships among objects and neglects their geometric details, i.e. location, length, sinuosity etc. Different from the topology in the conventional geographic information systems which concerns how graphic primitives (points, polylines, polygons, and pixels) interconnect, the topology in this study is the relationship among meaningful geographic features (e.g. a continuous named street rather than its contained meaningless segments or vertices). This refined topology leads us to perceive the fractal or scaling structure of the geographic space. As the case with the road network at a city or bigger scale, the street-street topology reveals the underlying scaling pattern of far more less-connected streets than well-connected ones (Jiang et al. 2008). In this connection, topology and fractal or scaling linked closely with each other and form the theoretical foundation of this study.

3 Data and methods

The study is situated in the big-data context. The diverse nature of emerging geospatial big data plays an important role in shifting the paradigm of the geospatial analysis (Jiang and Thill 2015), and thus is powerful for obtaining new insights of geographic forms and processes. The study benefits considerably from the volunteered geographic information (VGI) (Goodchild 2007) coming from social media platforms that are OpenStreetMap (OSM), Twitter, Brightkite, and Gowalla. To be more specific, the study employs the global OSM historical database, the countrywide geo-referenced tweets and check-in locations. Each record of the data sets stores accurately the spatio-temporal information at the individual level. With the help of massive VGI data, the study generates two types of geographic representations, i.e. natural streets and natural cities, to examine human activities over the geographic space across city, country, and world scales. Natural streets are a collection of individual street segments with good continuity; Natural cities refer to spatially clustered human activities from individual locations or street blocks.

To effectively discover the scaling properties of the geospatial big data, the study proposes a complexity science methodology that primarily relies on head/tail breaks (Jiang 2013b) and its induced ht-index (Jiang and Yin 2014), together with power-law detection and complex network analysis. Head/tail breaks method derives effectively and visualize the scaling hierarchy of data with a heavy-tailed distribution, and its induced ht-index indicates the number of hierarchical levels. Through the proposed methodology, the study demonstrates clearly the underlying scaling pattern of the geographic space and how human activities are shaped by the scaling structure. The analyses and findings include: (1) Predict urban human activities on the natural streets (2) Analyse the city tweets and their densities based on street blocks; (3) Find the correlation between social and geographic aspects of human activities in US; (4) Explore the global mapping activity in OSM.

4 Contributions of this study

The contribution of the study is multi-fold. First of all, the new definition of fractal is extended from conventional top-down to now bottom-up manners, and further facilitates a more in-depth understanding of the fractal. The deeper understanding of fractal enables us to detect the scaling property through head/tail breaks and power law statistics of the massive VGI data at city, country and global scales. The derived scaling hierarchies, power law metrics, and network measures provide new insights into the scaling structure of the geographic space based on natural streets and natural cities, and further help us to understand how human behaviour or activities are shaped by such scaling structure.

References

Goodchild M. F. (2004), The validity and usefulness of laws in geographic information science and geography, *Annals of the Association of American Geographers*, 94.2, 300–303.

Goodchild M. F. (2007), Citizens as sensors: the world of volunteered geography, GeoJournal, 69(4), 211-221.

Jiang B. (2013a), Volunteered geographic information and computational geography: New perspectives, *In Crowdsourcing geographic knowledge: Volunteered geographic information (VGI) in theory and practice*, editor. Sui D. Z., Elwood S. and Goodchild M. F., Dordrecht: Springer.

Jiang B. (2013b), Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution, *The Professional Geographer*, 65(3), 482 – 494.

Jiang B. (2015), Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity, *GeoJournal*, 80(1), 1-13.

Jiang B. and Thill J. C. (2015), Volunteered Geographic Information: Towards the establishment of a new paradigm, *Computers, Environment and Urban Systems*, 53, 1-3.

Jiang B. and Yin J. (2014), Ht-index for quantifying the fractal or scaling structure of geographic features, *Annals of the Association of American Geographers*, 104(3), 530–541.

Jiang B., Zhao S., and Yin J. (2008), Self-organized natural roads for predicting traffic flow: a sensitivity study, *Journal of Statistical Mechanics: Theory and Experiment*, July, P07008.

Mandelbrot B. B. (1982), *The Fractal Geometry of Nature*, W. H. Freeman and Co.: New York.