Advances in Economics, Business and Management Research, volume 100
International Conference of Organizational Innovation (ICOI 2019)

USING FRACTAL THEORY TO STUDY APPLICATION IN LIBRARY AND INFORMATION SCIENCE

Yu Kaijun (Library, Shanghai University of Medicine &Health Sciences, Shanghai, China)
Ruiyi Gong (Library, Shanghai University of Medicine &Health Sciences, Shanghai, China)
Shanshan Hu (Library, Shanghai University of Medicine &Health Sciences, Shanghai, China)
Song Luo (Department of Equipment, The Affiliated Hospital of Guizhou Medical University, Guizhou, China)

Email：luosong@gmcah.com

Abstract—As a nonlinear science, fractal theory has already been becoming an important method in the study of complex system. In this paper we make a preliminary study on the application of fractal in library and information science in the following three respects ： 1) Fractal and classical bibliometrics law; 2) Uniform bibliometric model; 3) Fractal and library. Fractal theory has a new application prospect in digital libraries and knowledge management.

Keywords—Fractal theory, Library science, Information science, Bibliometrics, Knowledge management

I. INTRODUCTION

Since the 1970s, the advent of nonlinear science with chaos theory, fractal geometry and orphan particle theory has marked the beginning of a new era in scientific development. The concept of nonlinearity represents a brand-new concept of nature. It unites simplicity and complexity, order and disorder, certainty and randomness, inevitability and contingency in the natural[1].

Fractal theory has been widely used in many disciplines such as mechanical science[2], astronomical meteorology[3], life sciences[4], and information science[5] since its establishment in the mid-1970s by Mandelbrot[6]. Fractal theory provides us with a new worldview and methodology that recognizes that changes in spatial dimensions can be either discrete or continuous.In the book *The Fractal Geometry of Nature* published in 1982[7], Mandelbrot discovers the Zipf's law of the frequency or occurrence of words through the so-called dictionary tree, and generalizes it into various social and natural phenomena[8].This shows that these natural and social phenomena have good fractal characteristics.

Library science and information science play an important role in the rapidly developing modern society. At present, the understanding and exploration of fractals in Chinese library and intelligence circles mainly focus on the application of the three laws of bibliometrics. Scholars at home and abroad focus on studying equations, parameter estimation, and fitting optimization to find a mechanism to explain the relationship between these three laws. Although many achievements have been made, the theoretical model for the unification of the three empirical laws is still under study. Fang shu and other scholars have examined some examples of Zipf's law, and Yin Zhenzhong studied the fractal properties of Lotka's law[9-10].

In recent years, with the in-depth application of network information technology in library science and information science, library and iservices have undergone profound changes.Therefore, it is necessary to deeply study the application of fractal theory in library and information science, and draw on the theoretical methods of nonlinear science for research on fractal and classical bibliometrics, unified bibliometric models, and library fractal management.

II. FRACTAL AND FRACTAL DIMENSION

A. *Fractal*

Fractal theory is a science that specializes in the study of fractal geometric features, quantitative representations, and their applications. The idea of fractalism first appeared in the paper "*The length of the British coastline*" published by Mandelbrot in Science in 1967, and proposed the concept of fractal geometry during the lectures of the French Academy in 1973[11]. In 1975, his monograph "*Fractal: Shape, Opportunity and Dimension*" was published as a symbol of the birth of fractal theory. In the 1980s, his book "Fractal Geometry in Nature" gave a fractal definition as a set of Hausdorff-Besicovitch dimensions strictly larger than its topological dimension. In 1986, he further defined fractals as their constituents and overall objects that were similar in some way. Such objects are extremely fragmented and complex and cannot be described by Euclidean geometry, but these objects are self-similar or self-affinity systems.

B. Fractal dimension

Fractals are irregular, non-uniform complex geometric systems, all of which have non-integer spatial dimensions. Non-integer dimensions are called fractal dimensions, denoted as D. Fractal dimension is an important parameter to measure the irregularity and complexity of a geometric set or natural object, which illustrates the irregularities and complexity of the research objects in time and space[12]. In 1919, Hausdorff studied the properties of singular sets, first proposed the idea of fractal dimension, and defined the Hausdorff measure and dimension theory. The dimension of the space is not saltation, but can be continuously changed. It can be an integer (one-dimensional, two-dimensional, three-dimensional, etc.) or a fraction. On this basis, many scholars have proposed more than a dozen different dimensional calculations and expressions according to different research directions, such as self-similarity dimension, box dimension, information dimension, correlation dimension, generalized dimension and Lyapunov dimension[13].

III. FRACTAL AND CLASSIC BIBLIOMETRICS LAW

Scientometrics is based on two types of empirical laws, namely the law of the distribution of bibliometrics and the law of dynamic growth aging. The common feature of the three famous laws is that they are derived from the statistics and research of data sources in a large number of literatures.

A. Bradford's Law and Fractal

Bradford's Law: If the various scientific journals in a subject area are arranged in descending order of the number of papers, these journals can be divided into the core area with the largest number of papers and several areas with the same content as the core area. At this time, the number of journals in the core area and each successor area is 1: $N:N^2$. N is called the Bradford Discrete Coefficient, and its mathematical description (Limkuler function) is:

$$R(r) = a\ln(1+br) \qquad (1)$$

In which, $R(r)$ is the accumulation of papers in the 1 to *r* grade journals, *r* is the journal grade number, *a* and *b* are constants. Bradford's law has fractal features, which contains self-similarity. Take the derivative of both sides of equation (1).

$$R'(r) = ab/(1+ab) \qquad (2)$$

The Mandelburt fractal law is:

$$g(r) = G/(1+Hr)^n \qquad (3)$$

In equation (3), $g(r)$ is a density function, G and H are constants, $n=1/D$ is a Mandebolt dimension, and D is a system fractal dimension. Comparing equation (2) and (3), when $n=1$, $D=1$,that is: Bradford's law has the feature of 1 dimension.

B. Ziff's Law and Fractal

For a language text, all words in the language are sorted in descending order of use, and the order or serial number of each word represents its position in the sequence table. Therefore, the relationship between the frequency of occurrence of a word and its ordering or serial number can be simply expressed as:

$$f_r = C \cdot r^{-1} \qquad (4)$$

This means that the product of the frequency at which a word appears and it's sort or serial number is the same as its order or serial number is constant. This law was first proposed by physicist E. U. Condon[14]. But because of the *Human Behaviour and the Principle of Least Eff*ort published by the philosopher Zif, it is now called Zipf's law. Therefore, equation (4) should be more accurately called Zipf-Condon's law[15].

The further extension of Zipf's law is:

$$f_r = C \cdot r^{-1} \Rightarrow C \cdot r^{-dz} \to f_r = C \cdot (r-a)^{-dz} \qquad (5)$$

Where C is a constant, f_r is the frequency of each word in a longer article, and r is the number assigned to the word corresponding to f_r . Initially, Ziff proposed $d_z=1$ through a large number of statistics, but after a lot of verification, d_z was between 0.7 and 1.04. Comparing equations (4) and (5), the relationship between d_z and fractal dimension D is:

$$d_z = 1/D \qquad (6)$$

C. Lotka's Law and Fractal

The famous American scientologist Rotka published the famous paper "*The Frequency Distribution of Scientific Productivity*" on June 19, 1926[16]. His research results for the first time quantitatively revealed the quantitative relationship between the author and the literature.

Lotka's law is a law of experience summarized from experimental data. The content described has statistical self-similarity, and self-similarity is one of the important properties of fractal theory. Lotka's law can be expressed as:

$$f(x) = c/x^{\alpha} \qquad (7)$$

Where c is the characteristic constant. Equation (7) is called generalized lotka's law. If $\alpha = 2$, lotka's law is expressed as:

$$f(x) = c / x^2 \quad (8)$$

Equation (8) is the original mathematical expression given by lotka, so it is called the classical lotka law. A large number of experimental data show that the negative power distribution law given in equation (7) is universal in a wide range of subjects or subjects. But then people tested the professional data of various disciplines and found that A was not all equal to 2, but a variable between 1.2-3.8[17]. And strictly speaking, the α values are almost all non-integer, in fact, has a fractal meaning.

Among the three empirical laws, Bradford's law is the law of document dispersion, that is, using the journal grade number to measure the cumulative distribution of papers related to a subject. Zipf's law is to measure the frequency distribution of words in an article by assigning the serial number of the word. Lotka's law is to measure the author's distribution by the number of papers written by the author. They all use a certain scale to measure a subject, moreover, the observations change with the change of the "scale", the "scale" shrinks, the observation increases, and vice versa. Therefore, we designed $f(x)$ as the observed distribution density measured by scale x for a subject. According to the scale of fractal theory, we derive the relation between $f(x)$ and x as:

$$f(x) \propto x^{-D} \quad (9)$$

Perform logarithmic operations on both sides of the above equation can obtain a equation for calculating the eigenvalues:

$$D = \log f(x) / \log(1/x) \quad (10)$$

Equation (9) is called the negative power law statistical fractal in information science, and D is the fractal dimension in information science. For Bradford's law, $f(x)$ is the cumulative number of related papers, and x is the journal rank number. For Zipf's law, $f(x)$ is the word frequency, and x is the ordinal of the word. For lotka's law, $f(x)$ is the number of paper authors, and x is the number of papers. D is an extremely important parameter with profound fractal significance, which characterizes the unevenness of distribution.

IV. Unified scientific measurement model and fractal

The scientific measurement model has seven principles, namely: actor network principle, transformation principle, space principle, quantitative principle, synthesis principle, central-peripheral or nucleation formation principle and unified principle of cumulative advantage. Based on this, a unified scientific measurement model for introducing fractal index and cross fractal index can be established. The model can unify all the laws of bibliometrics and scientometrics, including Zipf's Law, Bradford's Law, Lotka's Law and Pareto's Law, Science's Law of Exponential Growth, and Brooks-Alamos Library literature aging law, and can be better applied to the expression of empirical values[18].

According to the quantitative principle, the model conversion function equation that satisfies the monotonic transformation can be expressed as:

$$T(x) = \frac{dF(x)}{dx} \quad (11)$$

Therefore, for irreversible transformation, the base equation of the model can be expressed as:

$$\frac{dF(x)}{dx} = -k \frac{F(x)}{(x+m)^{\varphi}} \quad (12)$$

This is an ordinary differential equation of a separable variable.

$$\frac{dF(x)}{f(x)} = -k \frac{dx}{(x+m)^{\varphi}} \quad (13)$$

The integration interval ranges from 1 to x.

$$\int_{F(1)}^{F(x)} \frac{dF(x)}{F(x)} = -k \int_{1}^{x} \frac{dx}{(x+m)^{\varphi}} \quad (14)$$

V.LIBRARY FRACTAL MANAGEMENT

In 1992, professor Wanneck, President of the German society of engineers, proposed the concept of "fractal management"[19]. Fractal management theory uses the concept of self-similarity in fractal geometry to treat subsystems, treat departments and even employees in enterprise systems as a fractal element.

A. Feasibility of Library Fractal Management

The library is an open, unbalanced complexity system. Its functions and structures are self-similar and have fractal features within a certain range. For public libraries, libraries include natural science management books, social science books, tool books, digital resource platforms, etc. The library's book-centric matrix organizational structure lays a good foundation of self-similarity for its implementation of fractal management. The distribution of the management of each library has similar characteristics, so the fractal management of library books is feasible.

B. The Fractal Characteristics of Library Management

The library management structure consists of many different fractal units that share the basic platform of the organization. The higher the similarity of each fractal unit, the more resources it can share, the lower the cost of acquiring shared resources, and the closer the internal connection is. Library management is composed of many different types and different levels of personnel interaction, which can be divided into different hierarchical structures from multiple angles. According to the information flow and capital flow, the library management establishes the fractal unit of personnel structure, utilizes the library resource and information sharing platform and mechanism, establishes the team cooperation and coordination mechanism, makes the structure flat, reduces the management level, and realizes the rationality of personnel arrangement. The autonomous units of departments and posts are divided according to library management classification. In summary, the library has obvious fractal features such as hierarchical, finite iterative nesting, self-similarity, self-organization, self-coordination, and dynamic evolution within the scale.

VI. CONCLUSION AND PROSPECTS

In the era of big data, the characteristics of library digitization and virtualization become more and more obvious, and the nonlinear characteristics and characteristics of non-equilibrium systems are presented.

As a nonlinear and complex library knowledge management system, the relationship between the variables of knowledge sharing cost is very complicated, with uncertainty and randomness, and it may be Chaos. Fractals provide important methods and mathematical models for describing the irregularities, complexity or chaos of knowledge management systems. Therefore, the study of library knowledge management and the fractal and chaotic characteristics of digital libraries may be worthy of further consideration and research in the field of library science and information science.

ACKNOWLEDGMENTS

This work was financially supported by Library Information Research Fund for University Library Alliance in the Yangtze River Delta Region(Grant number 2016B018).

REFERENCES

[1] Zhao Jian, Lei Lei, Pu Xiaoqin. Fractal Theory and Its Application in Signal Processing. Beijing: Tsinghua University Press, 2008.
[2] Rinaldo A, Rodrigueziturbe I I, Rigon R, et al. Self-organized fractal river networks. Physical Review Letters, 1993, 70(6):822-825.
[3] Fossum P K, Frøyd L, Dahlhaug O G. Design and Fatigue Performance of Large Utility-Scale Wind Turbine Blades. Journal of Solar Energy Engineering, 2013, 135(3):14-16.
[4] Puetz S J, Borchardt G. Quasi-periodic fractal patterns in geomagnetic reversals, geological activity, and astronomical events. Chaos Solitons & Fractals, 2015, 81(12): 246-270.
[5] Parsa M, Maghsoudi A, Ghezelbash R. Decomposition of anomaly patterns of multi-element geochemical signatures in Ahar area, NW Iran: a comparison of U-spatial statistics and fractal models. Arabian Journal of Geosciences, 2016, 9(4):1-16.
[6] ChamorroPosada P. A simple method for estimating the fractal dimension from digital images: The compression dimension. Chaos, Solitons & Fractals, 2016, 91(8): 562-572.
[7] Lin Hongyi, Li Yingxue. Fractal Theory and Exploration of Singularity. Beijing: Beijing Institute of Technology Press, 1992:22.
[8] Mandelbrot B.B. The fractal geometry of nature. Macmillan,1983.
[9] Fang Shu, Li Houqiang. Lotka's Law and Fractal Theory. Library and Information Service, 1989 (6): 9- 14.
[10] YI Zhenzhong. Research on the Fractal Properties of Lotka's Law. Journal of Shandong University (Engineering Science), 2007,37(4):118-121.
[11] Mandelbrot B.B. How long is the coast of Britain. Science, 1967, 156(3775): 636-638.
[12] Mandelbrot B.B. Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands. Proceedings of the National Academy of Sciences, 1975, 72(10): 3825-3828.
[13] Mandelbrot. Fractal object. Beijing: World Book Publishing Company, 1999: 17.
[14] Zhao Laiyuan. Interpretation of Gruss Offset. Library and Information Service, 1995(3):18-19.
[15] CONDON E. U. Statistics of vocabulary. Science, 68(1928) 1733.
[16] ZIPF G.K. Human Behaviour and the Principle of Least Effort. Adisson-Wesley Press, Inc Cambridge, 1949.
[17] ROUSSEAU R. The number of authors per article in library and information science can often be described by a simple probability distribution. Journal of Documentation, 1994(50) :134-141.
[18] EGGHE L. Consequences of Lotka's law for the law of Bradford. Journal of Documentation, 1985(41):173-189.
[19] Zhang Xiaoping, Gan Renchu. Fractal Management of Enterprises. Foreign Economics and Management, 1999, 21(8):6-10.