ResearchLEAP

Journal of International Business Research and Marketing
Volume , Issue ,
journal homepage: www.researchleap.com

An Entropy Approach to China Big Data Ecology

Tao Li
Associate Researcher, Innovation & Development Research Center, Nanjing University of Science & Technology, Nanjing, China
MBA, Associate Researcher, Shanghai University of Finance & Economics, Shanghai, China

DOI: *10.18775/jibrm.1849-8558.2015.34.3003*
URL: *http://dx.doi.org/10.18775/jibrm.1849-8558.2015.34.3003*

Keywords:
Big data ecology, Entropy, Knowledge innovation, Open innovation

ABSTRACT

This paper analyzes the application layer of Internet Big Data in China from the perspective of Information Entropy, which studies the evolution and characteristics of the big data application layer in China. The application layer faces public users directly, from which the feedback effects reach commercial-end and government-end. By utilizing a combination of entropy approach and method of concatenated substitution, the structural evolution characteristics of public preferences are well-reflected. Discovering that the entropy increment mainly comes from newly-born groups, though such groups cause the disorder of entire application layer, the left-side of inverse U shape corresponds the national strategy of innovation ecology in Internet Big Data. With a dialectical and rational view of the system's entropy increase, we must also be aware of the economic transformation incentive policy in accelerating the establishment of an open innovation system for the "Mass Entrepreneurship and Innovation". The local government should be encouraged to promote the Internet Big Data industry innovation in policy coordination, enhance the construction of the fundamental level of the ecological system, and pay close attention to the level change of consumers' demand, the innovation performance of commercial organizations, and the abuse of public information.

1. Introduction

In the late twentieth Century, mobile phone, Internet, Big Data, Blockchain, cloud computing and innovation marked the fourth industrial revolution for human beings. In 2012, Obama administration released "Big Data Research and Development Initiative" (Whitehouse, 2012) whose purpose was to enhance the ability of the United States refining knowledge and acquiring large amounts of complex digital information, and solving some of the country's pressing problems and challenges.

In 2014, China wrote Big Data into the government work report for the first time. The report pointed out that the establishment of new industries, entrepreneurship innovation platform, in Big Data and other aspects of catch-up, leading the future industry development (Li Keqiang, 2014).

Thus, there is consistency between China and the United States on the overall strategy of the government promoting Big Data, and there are differences in target segmentation. In the report, the Chinese government advocates leading the industry with pioneering innovation. It is in line with the strategy of promoting economic growth with the "double engine" of "Mass Entrepreneurship and Innovation."

This paper attempts to study the following questions: what characteristics of the evolution of China's Internet big data ecosystem are, and what policy level and industry level can be extracted from the empirical results.

2. Literature Review

2.1 Theory of Information Entropy

Since Shannon (Claude Elwood Shannon, 1961-2001) introduced entropy into information theory to measure the uncertainty degree of a random event in 1948 (Shannon, 1948), entropy theory has been widely used. Economics from the perspective of information entropy has vital implications for the development of information society (Zhang Chuwen, 2001). Zhang and Yu analyzed the positive evolution factors of the industrial ecosystem by using entropy and dissipative structure theory (Zhang Wenlong & Yu Jinglong, 2009).

The entropy theory in natural science has a significant difference from the theory of information entropy in Social Science. The greater the information entropy is, the greater the degree of disorder is, and the more information it carries. The information entropy in social science is not monotonically increasing. The Big Data industry and Big Data ecology and organizations in the social sciences have the ability of self-evolution. Entropy increase and entropy decrease exist in the process of society and industry ecological evolution at the same time. Therefore, during the research of industrial ecosystem in the social problems, we should not only analyze the industry ecosystem entropy trend but analyze the influence factors of the increase or decrease of entropy as well.

2.2 Big Data Industrial Ecology

“Big Data” was mentioned in “Nature” in 2008 for the first time, which marks the evolution of human's knowledge from large amounts of data and large data sets to "Big Data", and from computer science, data science to commercial applications. GOOGLE proposed the term "Big Data" in the background of European Large Hadron Collider, genome sequencing Sanger center, and massive data storage and computation (Nature, 2008). Beyar summarized the characteristics of Big Data as 3Vs - volume, velocity, variety (Beyer & Laney, 2012). Scholars studied the characteristics, modeling and processing methods of big data, which provides the research foundations for the theory and methods of big data. In this paper, the business decision process of Big Data in the current stage is summarized as follows: Target requirements, Algorithm design, Data storage, Data preprocessing, Data interpretation, and Quantitative decision making. Based on relative theories such as 4Vs (Volume, Variety, Velocity, Value (Cukier & Schönberger, 2013)), 5Vs (Volume, Velocity, Variety, Value, Veracity (Bernard, 2014)), 5Rs (Relevant, Real-time, Realistic, Return on investment (Stidston, 2014)), HACE (Heterogeneity, Autonomy, Complexity, Evolution (Wu, Zhu, Wu, & Ding, 2014)), and BigKE (Big Data Knowledge Engineering (Wu et al., 2015)), this paper defines the nature and scope of Big Data as follows: Big Data, whose nature is data and information, is a series of procedure for processing, computing and storing an endless stream of data through a medium of transmission by means of a predetermined set of information gathering methodology.

Fransman pointed out that personal computers and communications technology spawned the Internet, further transforming it into a new information communication ecosystem – New ICT Ecosystem, then a four-layers model proposed by Didier Lombard, suggesting that such new ICT ecosystem plays a vital role in economic growth and social progress (Fransman, 2010). Li & Wei proposed not only an Internet-based self-organization model and non-self-organization model based on Chinese industrial cluster informatization but a modular process of Virtual Industry Cluster as well (Li Bin & Wei Chuanyong, 2012). Luo analyzes and compares the development of Big Data industry in China and in America, raises the core view of the transformation, upgrading of information technology enterprises, and the protection of intellectual property rights of software technology (Luo Tao, 2013). Di proposed a dichotomy (market share), and a three-point method (marketing model), and a five-point model (value model) of China's Big Data industry classification (Di Liya, 2014). Chi summarizes a cross-regional, cross-format and cross-industry development path of the Big Data industry cluster (Chi Lian, 2015). Liu holds an opinion that the essence of the Internet industry cluster upgrading of is integration, transformation, and innovation (Liu Zhou, 2015). Qu proposed that the government should increase Internet infrastructure in the upgrading of the Internet industry cluster, and create a pattern of cross-regional and inter-industry cooperation (Qu Beihang, 2016). From a marketing perspective, Han discusses the use of Big Data for accurate marketing and user information security in E-commerce (Han Hui, 2017). Deng discusses the difficulty of sharing Big Data and the difficulty of dealing with low-quality data in E-commerce (Deng Zhilong, 2017). Che takes Amazon Co as the research object and analyzes Porter's five forces model. It points out that Amazon is facing the situation of increasing threat from competitors in the online and offline markets in the era of the Big Data (Che Siqi, 2017). Yang puts forward the system of government process management, public policy management, intelligent medical care and intelligent policing for the development trend of Chinese E-government (Yang Minggnag, 2017).

Scholars and researchers also study the open innovation and commercial ecology. Freeman suggests that enterprises interact with customers, suppliers, governments, universities and other institutions to promote innovation (Freeman & Boeker, 1984). Moore, based on cases such as APPLE and IBM, proposes a coevolutionary view of business ecology that enterprises are in an innovative environment that is both competitive and cooperative (Moore, 1993). Iansiti suggested that enterprises should set their company strategy according to their niche, and the company's strategy should be adjusted with the change of ecological circle (Iansiti & Levien, 2004). Yun and Jeong analyzed the patents of four major cities in Korea and found that the depth and breadth of open innovation also increased as the process of knowledge development in the cities continued to expand (Yun, Jeong & Yang, 2015). Yun studies the dynamics of (Open Innovation Economic System) OIES's open innovation economy and answers the question of economic growth after the capital economy, suggests it necessary to introduce a combination of technology, market and social structure to SMEs, start-ups and innovative economic development in open innovation economy (Yun, 2015). Yun, Won, and Park establish conceptual models for the analysis and prediction of the dynamic effects of open innovation (Yun, Won & Park, 2016). Hence, Open Innovation promotes the healthy development of business ecology.

Based on the above relevant research results, the development of Internet industry, combining with the characteristics of Chinese enterprise "Internet plus" business model and the Big Data industry, this paper summarizes three layers' model for the Big

Data Ecology: Infrastructure and resources, commercial deployment, the application layer. (Exhibit 3.1, further discussion in 3.2)

3. Quantitative Description of Application Layer Structure in Big Data Ecosystem

3.1 Information Entropy Perspective

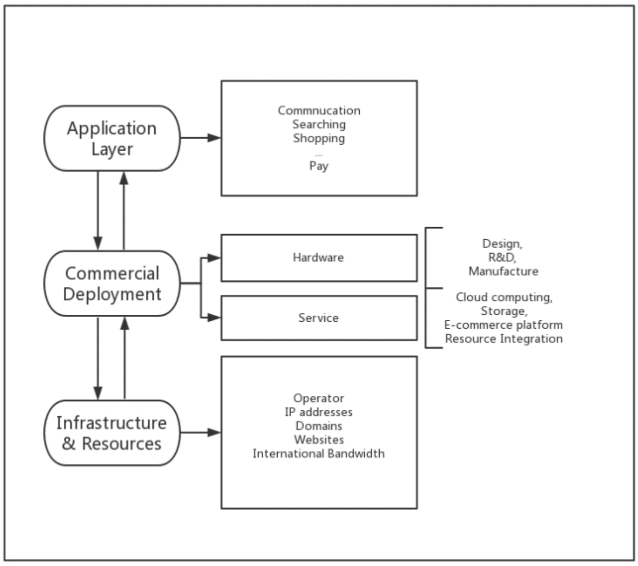

Figure 1: Three Layers' Model for China Big Data Ecology

Big data ecosystem and the external environment always commence information exchange. Meanwhile, the size and number of communities change dynamically within an ecosystem. According to the definition and formula of information entropy.

$$S = -\sum p_i \ln p_i$$

Shannon's information entropy has the characteristics of non-equiprobable (non-equal probability) constraint so that it can describe uncertainty and disorder for any system. Hence, the ability to measurement uncertainty for any event enables a broader applying scope of information entropy than those of thermodynamic entropy and statistical entropy. In the application layer of Big Data, consumers select services based on their needs. The classification of services is equivalent to grouping consumers into clusters, within which the change in the number of consumers leads to the change in the structure; the number of consumers in a group increases, and the proportion of the group as a whole increases; since that a part of the consumer selecting a new type of service produces a new grouping, it is a new species of the entire group; given a fixed number of consumers, the emergence of new groupings will reduce the proportion of others to some extent. Hence, as time goes on, the structure of the application layer continuously "in the effect of external disturbance and the internal fluctuations under the structure on Succession and changes, showing spontaneous and irreversible evolution characteristics, fully meet the predetermined assumption of dissipative structure system"(Geng Haiqing, Gu Shuzhong & Guo Dongmei, 2004).

3.2 Scope Definition of Big Data Ecosystem Application Layer

The Big Data ecosystembased on the Internet in application layer shows that the public users exchange information through hardware terminals and software platforms within, which should aim at matching of demand and supply and the distribution of resources via the Internet. Its ecology boundary should be defined as any platform and application that from online resource match processing to background data storing.

As mentioned in the literature review, this paper summarizes three layers' model for Big Data Ecology (Figure 1): Infrastructure and resources – first layer, commercial deployment – second layer, application layer – third layer.

The foundation of the Internet's Big Data environment comes from the Internet infrastructure and resources provided by the government and operators, as the foundation of earth ecology from the sun, water and air. Commercial enterprises and social organizations of profit or non-profit commence deployment involving hardware design, development, and manufacturing, at the same time providing open source and closed source software development, cloud computing, cloud storage, e-commerce platform, offline and online resource integration services. The business deployment is not only for enterprise users but also for individual professional users. The hardware design and research should focus on the efficient processing and utilization of the basic resources. In the application layer, large, small, own micro enterprises and organizations commence Internet transformation for part of the supply chain nodes, to establish their information systems, such as ERP, CRM and so on. Public users interact with enterprises and organizations for e-commerce activities in the application layer.

The Internet is one of the important carriers for the flow of Big Data. Generally, public users access to the Internet via PC or mobile terminals, commercial enterprises access the Internet with their internal LAN terminals, Big Data platform dealing with data sending from IoT terminals scattered around.

As an important carrier for Big Data information flow, it is necessary to enhance the construction of the broadband network, speed up network bandwidth phase-by-phase, learn policy incentives and business models from leading countries, which enables faster and broader e-commerce interaction.

3.3 Applying Information Entropy Approach

Big data ecosystem is built on the media of Internet, the motivations for human needs and exchanges are mapped from offline to online. Because of its intrinsic benefit drive and value strategy, commercial companies change the original closed type information management system into a modern ERP system with more open information input. Quick response from the front end to customer needs, which is a good customer experience that gathers customer base. Thus, a stable user community being divided by specific requirements is formed within the application layer. The gathering of multiple user communities becomes the application layer of the Big Data ecosystem. Therefore, the change in user communities reflects the change in the application layer of Big Data ecology.

3.4 Equilibrium & Dominance

Assuming an extreme situation that there is only one cluster or group in the application layer, for example, instant communication, the information entropy has a minimum value Smin=0; on the contrary, if there are a variety of groups with the same size, H1=H2=……=Hm=H/m, then the information entropy has a maximum value Smax=lnm in application layer. Generally, the two extreme situations will not occur, because the application layer always presents the situation of diversification and innovation with its information entropy value ranging in the interval [0, lnm] implying the complexity of application layer's cluster structure. Moreover, equilibrium degree and dominance degree reflect the degree of balance and dominance of the structure among clusters.

According to the formula of information entropy,

$$S = -\sum p_i \ln p_i$$

the equilibrium degree of the application layer in the Big Data ecosystem can be expressed as:

$$E = -\sum_{i=1}^{m} p_i \ln p_i / \ln m$$

S stands for the value of information entropy. When Smin = 0, implies E = 0, in which situation the complexity of the application layer is the lowest; when Smax=lnm, implies E = 1, in which situation the complexity is the highest. The dominance degree is expressed as D = 1-E; when E = 0, implies D = 1, which reflects a very prominent characteristics of application layer structure in the ecosystem; when E = 1, implies D = 0, which on the contrary, explains an unobvious characteristics of application layer structure – that is, there is not any cluster that makes its proportion higher than the other clusters.

4 Entropy Analysis of Application Layer Structure in China Internet Big Data Ecosystem

4.1 Research Design and Data Sources

This research takes China's big data industry as the research object. Since China's Big Data strategy released in 2015, provinces and cities promulgated planning document pertaining to the national strategy by January 2017. In "Big data Industry Development Plan (2016 - 2020)", national goal exceeds 1 trillion yuan. According to "China's Local Government Big Data Development Planning Analysis Report", as of 2020, the total output value of Big Data industry from local governments' plans are more than 2,840 billion yuan, which already far exceeds the national goal (Big Data Committee, 2017). Based on the definition of Big Data ecology structure mentioned in Ch3, internet and IoT both belong to infrastructure and resources (first layer), commercial deployment (second layer), in which the environment building and hardware establishing contribute to the platform that helps application layer acquire Big Data information flow.

Statistic results from the "Report" indicate that, at present, China's Big Data industry mainly concentrated in Beijing, Shanghai, Guangdong, Jiangsu and other eastern developed regions; Inner Mongolia, Ningxia, Guizhou, Chongqing and other central and western regions, where the economic development is in great demand for infrastructure construction, is growing rapidly in Big Data industry scale; these areas' investment layout of Internet and IoT has more room to play. Therefore, the eastern region of China Big Data industry ecology mainly based on the application layer, whose purpose is for commercial services and activities. The eastern region has more advanced business activities, talents, techniques, and innovations.

The central and western regions are mainly based on land and energy resources with large-scale infrastructure construction. If the East, Central, West, and Northeast region find their own Big Data industry development advantages that achieve regional characteristics of unrepetitive and non-homogeneous industrial development model, the industrial regional imbalance and redundancy are able to be avoided, and homogeneous competition and disorder of industrial clusters can be prevented.

The source data comes from "China Internet Network Information Center" report. In 1997, CNNIC lead a survey of Chinese internet development with related organizations. So far, CNNIC has successfully released 38 national Internet development statistics reports. CNNIC's previous reports with rigorous and objective data provided an important basis for decision-making of government, enterprises and other circles to master the development of China's Internet.

4.2 Entropy Evolution Characteristics of Application Layer Structure in China Internet Big Data Ecology

Based on the analysis of internet utilization status structure from 2013 to 2016 (Figure 2; Appendix: Table 1), the change of information entropy in the structure shows a significant trend of increase which implies a continuous rising in the complexity of application layer in four consecutive years. The complexity of the structure represents the uncertainty and disorder of the clustering groups in the application layer over time. It has a coupling relationship with the Information Processing Innovation Project by the Ministry of Industry and Information Technology of PRC in 2011, enterprise data operation proposed by Alibaba in 2013, Big Data emerging industry innovation platform from government work report in 2014, and opening a new ecological pattern for "Mass Entrepreneurship and Innovation" in "Action Outline for Promoting The Development of Big Data" by State Counsel in 2015 (State Council of PRC, 2015).

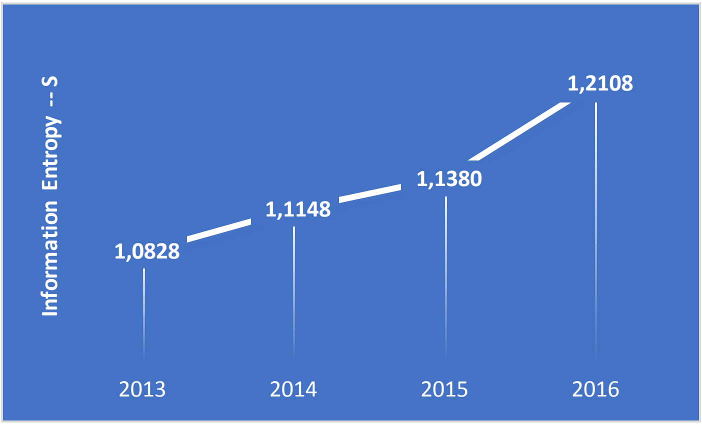

Figure 2: Information Entropy -- S

4.3 Factor Analysis for Entropy Value

The equilibrium and dominance results (Appendix: Table 1; Figure 3) from 2013 to 2016 tell a decrease in equilibrium degree from 2013 to 2015 and then a slight increase in 2016, while the dominance degree goes an opposite path.

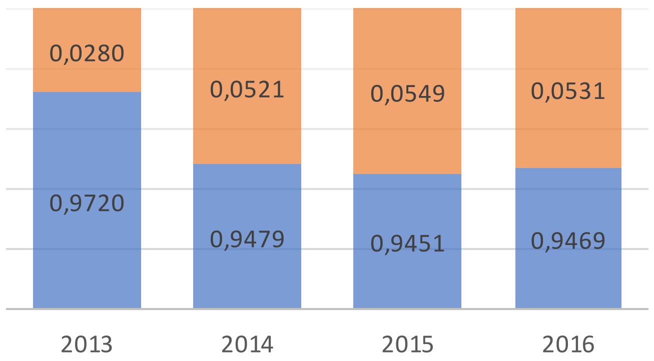

Figure 3: Equilibrium (blue) & Dominance (red)

To explaining it, a method of Concatenated Substitution is applied for a quantitative decomposition for analyzing driving factors.

The factor analysis via concatenated substitution on information entropy value of application layer structure changes shows a significant contribution to the entropy increase (increase in disorder degree) from "Online Investment" and "Online Stock/Funds" which were newly born clusters from 2013 to 2014 (Figure 4); "Online Education" – newly born cluster in 2015 – and "Online Pay" contributed the most to the entropy value from 2014 to 2015 (Figure 5); from 2015 to 2016, “Online Takeaway”, “Webcast” and “E-Government” – all newly born clusters – made the highest contribution, meanwhile the other clusters had reductions to a certain extent, which implies that these clusters formerly existed evolved in an orderly direction (Figure 6).

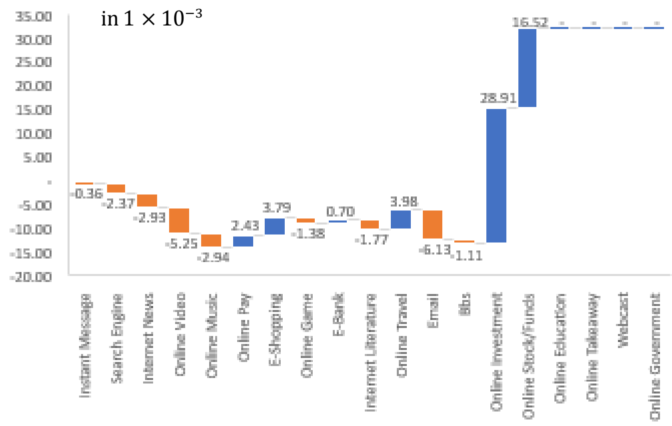

Figure 4: Waterfall Diagram of Influence Factor, 2013~2014

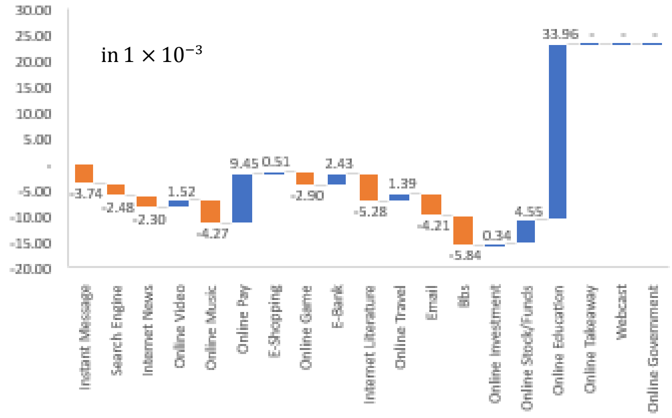

Figure 5: Waterfall Diagram of Influence Factor, 2014~2015

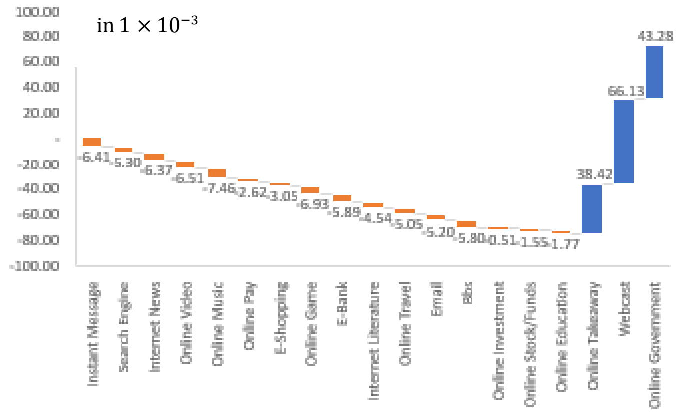

Figure 6: Waterfall Diagram of Influence Factor, 2015~2016

5. Findings & Conclusions

a. It is suggested that regulators and telecom industry should coordinate to accelerate the balanced regional development of the Internet in China.

Carrying forward the construction of Internet infrastructure to economic undeveloped areas and wireless network coverage without blind angles are the next five years' China Internet, IoT, Big Data development policy footstones and underlying stones, developing an Internet Economy and E-commerce in order to provide high-efficiency drainage for offline products and services through the Internet Big Data platform.

b. At the current stage, there is a strong relationship between the potential of open innovation in the development of Big Data background, and credit system of individuals that involve in the Internet Big Data ecology.

The development of an open innovation system should begin with the establishment of a perfect social credit system, inter-enterprise trust, and the Internet spirit of enterprise and individual. The government sets up and improves the regional open innovation community pilot project, which makes better regional Government - Enterprise - Public innovation climate and closer trust relationship. With linking up the pilot areas, China's mode of opening and innovation keeps constant clarifying.

c. For SMEs and startups, the establishment of an open innovation system and radiation effects of regional pilot projects plays a catalytic role in the implementation of "Mass Entrepreneurship and Innovation" national strategy.

Svirina et al. have pointed out that OI (Open Innovation) has a significant impact on improving public efficiencies and commercial economies, moreover, with the aid of the government or society, commercial enterprises that establish OI will achieve higher performance and shorter payback period (Svirina, Zabbarova & Oganisjana, 2016). In the current Chinese public business environment, SMEs and citizens with potential entrepreneurial awareness and desire are still on the level of relatively blank stage in acquisition and maintenance of manpower, weak in funding and knowledge system pertaining to IP rights and patent technology. Yun studies the dynamics of OIES – Open Innovation Economic System – and answers the question that knowledge economy is the macroeconomic growth point, further suggests that it is necessary to introduce a combination of technology, market and social structure to SMEs, start-ups and innovative economic development in open innovation economy (Yun, 2015). In the mainstream Chinese "Internet plus" model of light assets entrepreneurship, economic transformation strategy of "Mass Entrepreneurship & Innovation" is essentially the government providing technology, patent, an intellectual support platform for public entrepreneurship to break through the bottleneck.

d. Policy synergy in government promoting innovation in Internet Big Data Industry.

The government and regulators should pay attention to the construction of the basic level of the ecological system, pay attention to the level change of consumer demand of the public, adopt the Balanced Scorecard for reasonable evaluation of commercial innovation performance, draw lessons from other countries' government information disclosure positive effects on industrial ecosystem, and guard against and prevent leakage of commercial organizations and abuse of individuals' public information.

e. Regulators should stand in the height of Big Data industry layout, with the national economic development plan, to promote Big Data industry in various regions of the unified planning and to a position with the vision of continuous development, to avoid homogenization of competition and industrial cluster disorderly situation.

Learn from other industry overcapacity and the unreasonable industrial structure, the construction of large data industry and ecosystem should be alert to the enterprises' over-optimistic estimates of rate of return and flocking investment for their own short-term profit-making, or subsidy pursuing goals, that is, to get involved in a part of the industry chain and the acquisition of policy financing or financing concept with the national strategy and policy incentives. Policy subsidies should be based on strict project application and progress acceptance.

f. A Big Data applications layer ecosystem of Internet banking and financial innovation, strict review and due diligence mechanism should be established, the use of technical means to monitor and supervise the Internet financial instruments.

Loans and financing of financial institutions should be established on the basis of strict approval by government departments, Financial institutions should do well in auditing and due diligence for the use of loans and financial innovations. The securities regulatory authorities should timely and appropriately supervise and manage the excessive speculation of listed companies.

g. The carrier's transfer of Internet users' public sentiment poses challenges to supervision.

With the transfer from "BBS forum of PC terminal" to "Instant Messaging application with self-media function adapted to mobile terminal", the healthy development of webcast, as an emerging industry of Internet, depends on the development and application of the social sentiment censorship mechanism of Big Data and AI (Figure 7; Figure 8).

Data shows that there is a certain extent negative correlation between the number of instant messenger users and the number of BBS users (Pearson ecoefficiency: -0.5398). The change of the network public sentiment carrier of this side poses a great challenge to the information management of the Big Data application layer. Because of the self - media function of instant messaging tools, the public is faster in the dissemination of hot information and personal views. Compared to BBS, self-media is lack of the audit mechanism and the real-time supervision, also the privacy of the public users involves in the problem.

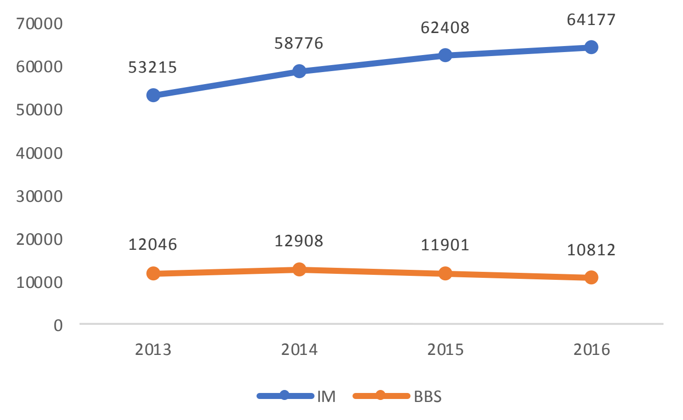

Figure 7: The Transfer of Public Sentiment Carrier under the Background of Internet Big Data Development

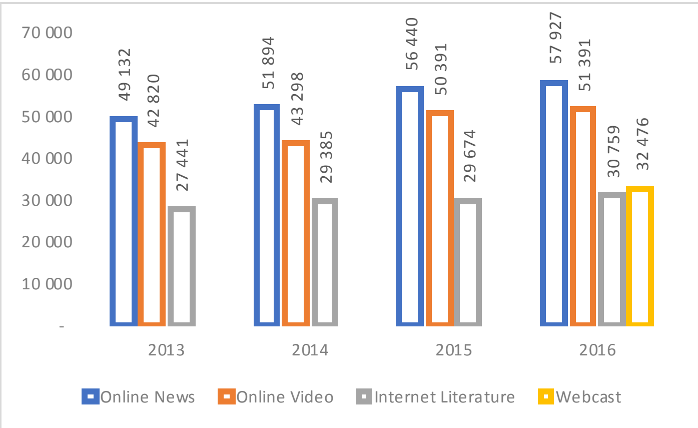

Figure 8: Trends in Other Public Sentiment

Therefore, under the Big Data environment, the monitoring system of network public sentiment should have the functions of hot spot identification, subject tracking, and automatic summarization of information, trend analysis, sudden event analysis and statistical report. As the main executor of the network public sentiment monitoring, government and the regulators need to cooperate with the enterprises providing instant messaging and self-media functions, simultaneously, a unified specification on the enterprises that master the public users of privacy information.
h. The structural changes of the Big Data application layer are essentially the derivation and promotion of the Maslow demand hierarchy in public demand from offline to online.

The Big Data application layer is a dissipative system of material and exchange in which commercial Big Data application platforms and public users participate is evolved over time. The matter is the carrier of information, and the transmission of information is realized by energy, thus in the process of continuous exchange of Internet data and information, the order of existing groups and structures is constantly increasing, and the emergence of new groups and structures will increase the disorder. Its essence reflects the change of public demand and the promotion of Maslow's demand level. The mapping of the public needs in the Internet world, implemented by the business organizations, is a significant external manifestation of the evolution of the industrial structure in Big Data application layer.

i. A complete inverted "U" trend requires a "Top-Down" policy structure that continuously stimulates the national strategy of “Mass Entrepreneurship & Innovation”.

Generally, with the increase of time and per capita income, the evolution characteristics of structural information entropy show obvious "inverted U" (Geng, Shu-Zhong & Guo, 2004). According to entropy change evolution from 2013 to 2016 of application layer structure (Exhibit 4.2-a), the application layer has undergone an entropy rising phase that could mean early stages of the Internet's Big Data strategy and policy. It cannot be predicted with certainty that when the top of the "inverted U" type comes. Whether the "inverted U" trend will be completed may depend on several factors:

- whether public demand level be mapped well on the Internet;
- whether business organizations acutely capture public demand and provide more efficient solutions for resource allocation;
- whether the target strategy of industrial ecology innovation and
- "Mass Entrepreneurship & Innovation" can continuously and effectively stimulate the innovation desire of commercial organizations and the public at the policy implementation level.

j. An entropy strategy maintained on the left side of "inverted U" type –maintenance of the newly born groups in structure leading to continuous increase in ecosystem’s entropy and rising disorder degree in application layer – is, in essence, the maintenance of the diversity, symbiosis, open coordination and self-organization evolution of innovation ecology and innovation environment.

The application layer of the Internet Big Data ecosystem with characteristics of innovation is not an isolated, closed "ecosphere", it is about providing an open environment into which foreign/external innovation continues to enter, and with the environment, structure increasingly complicated in that the information entropy value rising via entries and acquisitions of new level and new information. What’s more, the increase of information entropy in ecosystem requires not only entries of new information, new energy and new “species” that offsets the information entropy drop caused by declining in the activities of existing “species”, and requests establishing corresponding mechanism and methodology of digesting and transforming, which is closely related to the formulation of policies and regulations.

6. Contributions and Limitations

This paper tries to make an analysis of application layer of China Internet Big Data ecological system, explaining the reasons for its entropy increase in structural evolution. Because of the rapid development of the Big Data industry and the fact that objects in the physical world carrying information which can be collected and calculated, the boundary of the concept of Big Data is unclear. Similar to the mathematical and philosophical viewpoint that "All things are numbers" (Bowra, 1959) by Pythagorean School of ancient Greeks, "all things are information" (Bao Jigang, Xia Shutao & Liu Xinji, 2013) of information theory came into existence naturally, so that even unrealistic ideas such as "all things are data" is being proposed from time to time. However, subject to technology and the "cost-benefit" viewpoint of Utilitarianism (Maneschi, 1996), not all data and information can be collected and processed. This article uses a concatenated substitution approach to analyze the quantitative effects of various factors, without considering the order of the factors that may have small different effects. An improved approach that applying calculus can fix the shortcomings of the traditional approach used in this paper and make further analysis. According to marketing theory, consumers are in the upper level of the ecology – application layer – and the analysis of this layer can be combined with the psychology of Alderfer's theory and Maslow's hierarchy of needs theory.

Acknowledgments

This paper was received 'Best Paper Award' at SOItmC 2017 conference.

The author would like to thank Prof. Lei Ma, and the support from NSFC Funding Program No.71272164.

References:

- Bao Jigang, Xia Shutao, & Liu Xinji. (2013). Information, Entropy, Economics, Economic Science Press.
- Bernard, M. (2014). Big Data: The 5 Vs Everyone Must Know.
- Beyer, M. A., & Laney, D. (2012). The Importance of 'Big Data': a Definition.
- Big Data Committee. (2017). Analysis Report on Development Plan of Local Government's Big Data in China: Big Data Committee.
- Bowra, C. M. (1959). The Greek experience.
- Che Siqi. (2017). Difficulties That Amazon faces in Big data Era. Economic Research Guide (6), 147-149.
- Chi Lian. (2015). Formation Path of Big Data Industry and Dynamic Mechanism of Industrial Cluster Development. Commercial Economy Research (17), 66-68.
- Cukier, K., & Schönberger, V. M. (2013). Big Data: A Revolution that Will Transform How We Live, Work, and Think: Hodder & Stoughton.
- Deng Zhilong. (2017). Innovation of E-Commerce Marketing Model Based On Big Data. Economic Research Guide (7).
- Di Liya. (2014). Research on the Development of Large Data Industry in China. Scientific and Technological Progress and Countermeasures (4), 56-60.
- Fransman, M. (2010). The New ICT Ecosystem. General Information. Crossref
- Freeman, J., & Boeker, W. (1984). The Ecological Analysis of Business Strategy. California Management Review, 26(3), 73-86. Crossref
- Geng Haiqing,Gu Shuzhong & Guo Dongmei. (2004). Analysis of Urban Household Energy Consumption Structure Evolution Based On Information Entropy -- Take Wuxi as an Example. Journal of Natural Resources, 19(2), 257-262.
- Geng, H. Q., Shu-Zhong, G. U., & Guo, D. M. (2004). Analyses on evolution of household energy consumption structure based on information. Journal of Natural Resources (02), 257-262.
- Han Hui. (2017). Discussion on the Development of E-Commerce in Big Data Environment. Technology Outlook, 27(3).
- Iansiti, M., & Levien, R. (2004). Strategy as ecology. Harvard Business Review, 82(3), 68-78, 126.
- Li Bin & Wei Chuanyong. (2012). the Construction and Development of Virtual Industry Cluster: Model and Path Choice. Economic Research Guide (4), 210-211.
- Li Keqiang. (2014). Report on Government Work - At the Second Session of the Twelfth National People's Congress in March 5, 2014.
- Liu Zhou. (2015). Research On The Upgrading Of Industrial Clusters And "Internet Plus". Science and Management of Science and Technology, 36(8), 73-82.
- Luo Tao. (2013). American Experience in Big Data Industry and China's Countermeasures. High Technology and Industrialization, 9(5), 54-57.
- Maneschi, A. (1996). Jules Dupuit: A sesquicentennial tribute to the founder of benefit-cost analysis. The European Journal of the History of Economic Thought, 3(3), 411-432. Crossref
- Moore, J. F. (1993). Predators and Prey: A New Ecology of Competition. Harvard Business Review, 71(3), 75-86.
- Nature. (2008). Nature. Big Data.
- Qu Beihang. (2016). Research On The Upgrading Of Industrial Clusters And "Internet Plus". Economic Research Guide (19).
- Shannon, C. E. (1948). The Mathematical Theory of Communications, I and II. Bell System Technical Journal, 27. Crossref
- State Council of PRC. (2015). Action Outline for Big Data Development. Group technology and production modernization, 32(3), 51-58.
- Stidston, M. (2014). Business Leaders Need R's Not V's: The 5 R's Of Big Data.
- Svirina, A., Zabbarova, A., & Oganisjana, K. (2016). Implementing open innovation concept in social business. Journal of Open Innovation Technology Market & Complexity, 2(1), 20. Crossref
- Whitehouse. (2012). Obama Administration Unveils "Big Data" Initiative.
- Wu, X., Chen, H., Wu, G., Liu, J., Zheng, Q., & He, X., et al. (2015). Knowledge Engineering with Big Data. IEEE Intelligent Systems, 30(5), 1. Crossref

- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data Mining with Big Data. IEEE Transactions on Knowledge & Data Engineering, 26(1), 97-107. Crossref
- Yang Minggang. (2017). Trends in Government Big Data. Software and Integrated Circuit (1), 30-31.
- Yun, J. H. J. (2015). How do we conquer the growth limits of capitalism? Schumpeterian Dynamics of Open Innovation. Journal of Open Innovation: Technology, Market, and Complexity, 1(1), 17. Crossref
- Yun, J. H. J., Jeong, E. S., & Yang, J. H. (2015). Open innovation of knowledge cities. Journal of Open Innovation: Technology, Market, and Complexity, 1(1), 16. Crossref
- Yun, J. H. J., Won, D. K., & Park, K. (2016). Dynamics from open innovation to evolutionary change. Journal of Open Innovation: Technology, Market, and Complexity, 2(1), 7. Crossref
- Zhang Chuwen. (2001). Economic Implications of Information Entropy and Some Practical Issues. Journal of Hunan Business College (01), 97-99.
- Zhang Wenlong, & Yu Jinglong. (2009). Application of Entropy and Dissipative Structure Theory in Industrial Ecology Research. Social Scientists (2), 118-121.

Appendix

Table 1: Entropy Change Of Application Layer Structure For China Internet Big Data Ecosystem

	2013 Number of users In 10,000	2014 Number of users In 10,000	2015 Number of users In 10,000	2016 Number of users In 10,000
Instant Message	53,215	58,776	62,408	64,177
Search Engine	48,966	52,223	56,623	59,258
Internet News	49,132	51,894	56,440	57,927
Online Video	42,820	43,298	50,391	51,391
Online Music	45,312	47,807	50,137	50,214
Online Pay	26,020	30,431	41,618	45,476
E-Shopping	30,189	36,142	41,325	44,772
Online Game	33,803	36,585	39,148	39,108
E-Bank	25,006	28,214	33,639	34,057
Internet Literature	27,441	29,385	29,674	30,759
Online Travel	18,077	22,173	25,955	26,361
Email	25,921	25,178	25,847	26,143
Bbs	12,046	12,908	11,901	10,812
Online Investment		7,849	9,026	10,140
Online Stock/Funds		3,819	5,892	6,143
Online Education			11,014	11,789
Online Takeaway				14,966
Webcast				32,476
Online Government				17,626
Grand Total	437,948	486,682	551,038	633,595
S	1.0828	1.1148	1.1380	1.2108
E	0.9720	0.9479	0.9451	0.9469
D	0.0280	0.0521	0.0549	0.0531