

Research on Rural Economic Development Based on Data Mining

Xiaoya Ma* , Zhipeng Li

School of Economics and Management
Beijing Jiaotong University
Beijing, China
17120569@bjtu.edu.cn

Xiang Xie

Industry-oriented research university development strategy
research center
Beijing Jiaotong University
Beijing, China
406182963@qq.com

Abstract—The development of rural e-commerce and rural tourism has provided new ideas for the reform of rural economic construction. The rapid expansion of rural e-commerce represented by Taobao Village shows people the enormous potential of the rural e-commerce market. While e-commerce has changed the structure of rural economic development, the rapid development of rural tourism has also aroused people's concern. This article analyzes the development of Taobao Village and rural tourism and finds that the development of rural tourism is relatively balanced. The development of Taobao Village is highly spatially autocorrelated, which means that the development of Taobao Village will be affected more by the surrounding Taobao Villages. Finally, the article put forward relevant policy recommendations based on the analysis results.

Keywords—E-commerce; Data mining; Rural tourism; Rural economy

I. INTRODUCTION

With the rapid development of the economy and the vigorous promotion of the information age, products based on high-tech are constantly blended with traditional industries. While providing convenient services to people, it also greatly advances the process of Informationize. The rapid economic development is also actively promoting the prosperity of the tourism industry. The heavy life pressure and the stereotyped daily routine of the city have prompted more and more people to join the rural tourism during weekend and holidays. Rural cultural atmosphere and relaxed and comfortable environment also greatly enhance their quality of life.

In the development of Taobao Village, Que Beiyi and others analyzed the data from 2013-2014 on Taobao Village in China and found that the development of Taobao Village has an unbalanced spatial pattern, mainly in eastern coastal cities, and Taobao Village derives rural endogenous Cities and cities rely on two development models [1]. Zhao Junyang and others used data of Taobao Village from 2009 to 2016 to explore the spatial and temporal distribution and pattern change trends of Taobao Village under the three scales of the country, province, and city [2]. Based on data from Taobao Village and tourism villages, Liu Chuanxi analyzed the flow pattern and its influencing factors of rural villages in Zhejiang Province and found that the overall nodes are clustered, and the impact on their nest distribution is mainly due to factors such as traffic, resources, and government [3]. Xu Zhibang studied the

statistical data of Taobao Village from 2013 to 2015 and found that Taobao Village spreads to neighboring regions with Jiangsu and Zhejiang as the core, and the overall space is a ladder-like distribution [4]. Xin Xiangyang proposed the formation of a diffusion mechanism in Taobao Village: the accumulation of early-stage foundations (resources, technologies, and policies), and the creation and later evolution of later-stage evolutionary entrepreneurs with commercial thinking [5]. Cao Lingling and others used the ISM to analyze the factors affecting the clustering effect of Taobao Village and put forward targeted suggestions for improving the clustering effect of Taobao Village [6-10]. Liu Yajun believes that the poverty alleviation policy should be changed from "blood donation-type poverty alleviation" to "hematopoietic poverty alleviation", and the internalization ability of farmers through self-employment will not change with the environment. It is an effective and lasting poverty alleviation method [11].

In the theoretical research of tourism informatization, domestic scholars have done a lot of research. Li Junxi proposed the concept of tourism information science [12-16], and analyzed and constructed the system's framework structure of the discipline, which mainly includes the research object, Areas, themes, and goals. Abdulrasool D and Manoufali M have developed a smart guidance system based on sound reminding blind people, which can sense the user's position and make reminders by installing some hardware sensor devices such as wireless radio or bluetooth devices indoors [17, 18]. At present, smart travel projects have been run abroad: Belgium's "Identity City" project, South Korea's "I Tour Seoul" project, Singapore's "Smart Travel Program" project. The construction of these tourism informatization has highlighted the intelligence and convenience in the tourism process.

II. DATA SOURCES AND RESEARCH METHODS

The research data is mainly from the research report data of China's tourism villages, the statistics of the National Bureau of Statistics of China, and the government's public data of some provinces. According to the standards set by the Ali Institute, Taobao Village refers to a rural area with an administrative village as its unit, with more than 100 online shops or more than 10% of the active online shops occupying more than 10% of its administrative village households and the E-commerce transactions amounted to more than 10 million yuan. Demonstration tourism villages are based on the government's

This research was financially supported by China Railway Corporation Research Plan Project (2017D003-D, 2016D001-A).

list. The time dimension of this study is from 2014 to 2017. The list of Taobao Village and the past four years of tourism villages, addresses, coordinates and other data have been collected.

This research mainly uses social network analysis method, nuclear density analysis, barycenter analysis, SDE analysis and other methods to analyze the tourism rural data. The following elaborates on these research methods and measurement results.

A. Rural Tourism Social Network Analysis

This article first sorts out the detailed information of these tourist villages based on the list of 2017 national tourist villages, and then imports them into Gephi. The social network drawn is shown in Fig. 1. In order to count the geographic location information of the tourist town, and the four-year address of Taobao Town was converted into coordinates and imported into ArcGIS. The distribution map of Taobao Town 2017 was shown in Fig. 2.



Fig. 1. Social network diagram.

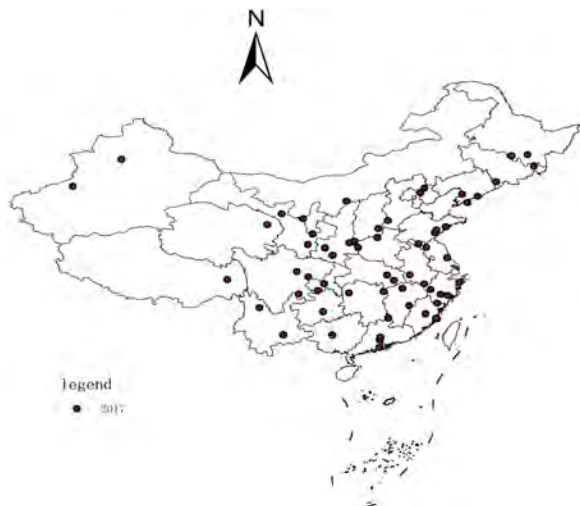


Fig. 2. Tourism town 2017.

B. Nuclear Density Analysis

In statistics, nuclear density estimation is a non-parametric method to estimate the probability density function of random variables. Based on limited data samples, the reasoning of populations is completely dependent on sample data to avoid the intervention of prior knowledge of human experience. The sample is estimated to a maximum degree. Assuming that x_1-x_n is an independent, identically distributed sample selected from the population of the density distribution function f , and $f(x)$ is the value of f at point x , the formula for estimating the nuclear density is obtained:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x-x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \tag{1}$$

In (1), K denotes the function kernel, and h denotes a selected bandwidth (window, neighborhood) parameter. The kernel with subscript h is called the scaling kernel and $x-x_i$ represents the distance between the estimated point x and the event point x_i . This study is to further study the agglomeration characteristics of rural tourism in 2017 by importing the address coordinates of tourism rural areas in 2017 into ArcGIS and applying ArcGIS's Kernel Density analysis to obtain the 2017 tourism rural nuclear density map (shown in Fig. 3).

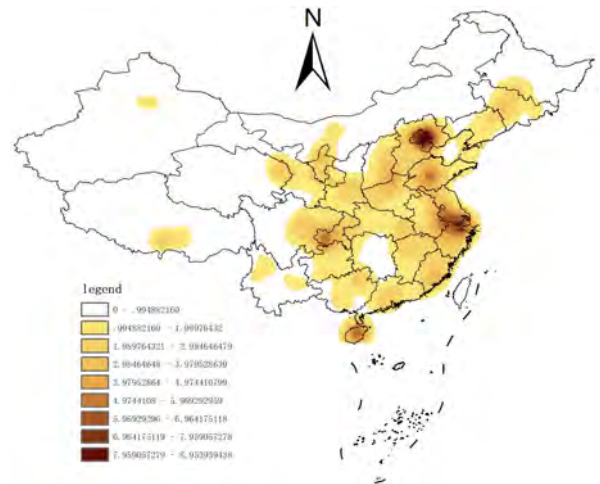


Fig. 3. Kernel density.

As can be seen from the map, the better-developed tourism rural areas in 2017 are mainly concentrated in the central and eastern parts of China and the southeastern coastal areas. The economic, technological, and other factors may be related to the facilities and policies of these places. Rural tourism has established a strong link between cities and rural areas, alleviating urban pressures, stimulating rural economic development and reducing the gap between urban and rural areas. At present, rural tourism websites are not centralized, information is scattered, and it is difficult to give full play to the advantages of the Internet. There are also some rural villages that use traditional media such as newspapers, radio and television in their propaganda. They have poor interactivity and are unable to attract large numbers of tourists. The adoption of a network integration platform can solve this problem well. Through the combination of web templates and

customization, it establishes its own dedicated page, integrates information platform for promotion, communication, booking, payment, and review, and improves the visibility and service quality.

C. Analysis of the Trajectories of Tourism Village

The center of gravity analysis is mainly to find a key element. The sum of the elements of the elements in the rest of the area is the smallest, which is similar to the process of finding the center of gravity of an object. The center of gravity of this study is to find a tourism village that has been at the centre of gravity for the 2014-2017 period. The tourist village that is in the center of gravity has the shortest distance between the year and the rest of the tourist villages. The center of gravity analysis can be expressed as:

$$X = \frac{\sum_{i=1}^n R_i X_i}{\sum_{i=1}^n R_i}, \quad Y = \frac{\sum_{i=1}^n R_i Y_i}{\sum_{i=1}^n R_i} \quad (2)$$

X and Y in (2) are the coordinates of the center of gravity of the tourist village in the region. X_i , Y_i represent the coordinates of the i-th tourist village in the region, R_i represents the distance from the i-th tourist village, and n is the total number of tourism villages in the current year.

This study separately calculates the location of the center of gravity of tourist villages in 2014-2017. The results are shown in Fig. 4 below. The study found that: The center of gravity of tourist villages in 2014-2016 has always been moving in Henan Province, and 2017 shifted southeast to Hubei Province. The general trend of tourism and rural development is to the southeast, indicating that China's rural tourism develops rapidly in the southeast, and the development quality is higher, because high tourism quality can become a model tourism village.

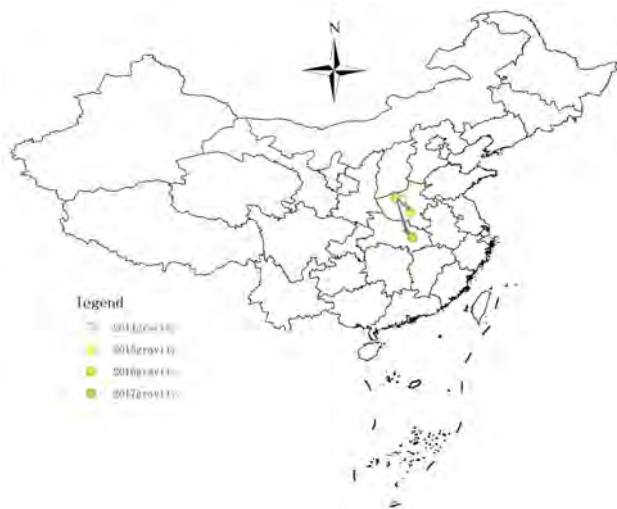


Fig. 4. Change of center of gravity.

After the analysis of the center of gravity, this paper proceeds to perform standard deviation ellipse analysis on the tourism rural data of 2014-2017. As shown in Fig. 5 below, it is found that the directional performance of the data for 2014-

2017 is not obvious, it shows that the geographical conditions have less impact on excellent tourism villages.

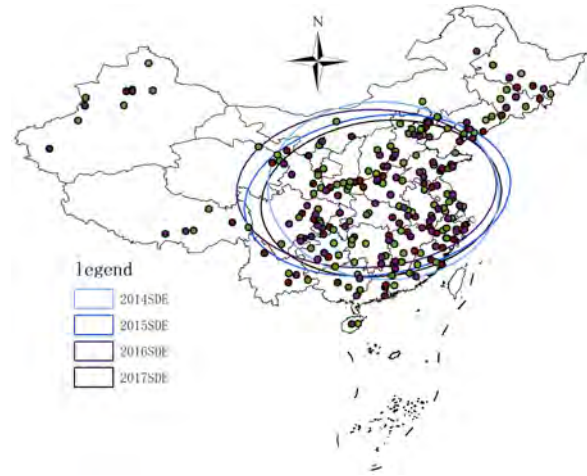


Fig. 5. SDE analysis.

III. TOURISM INFORMATION CONSTRUCTION PROGRAM

With the continuous improvement of China's economic strength and the doubling of national income, a larger and larger proportion of the living expenses are devoted to tourism consumption. since the reform and opening up, people's tourism awareness has also deepened. Therefore, the increasingly strong economic support and far-sighted consumption concepts have contributed to the vigorous development of the tourism industry. According to statistics released by the National Tourism Administration, in 2015, China's domestic tourism exceeded 4 billion people, and tourism revenue exceeded 4 trillion yuan. Among them, the number of inbound tourists reached 133 million in 2015, exceeding the number of tourists. The overall trend of tourism research is increasing year by year. The number of relevant research papers has increased from 200 in 2000 to about 600 in 2010.

At present, artificial guides are increasingly unable to meet the needs of tourists. The quality of tour guides is uneven, and the issue of high fees and charges also affects the experience of tourists. The popularization of smart phones and the advancement of positioning technology are also promoting the emergence of electronic guides. At present, there are already some electronic guide software on the market, but it is difficult to achieve real-time indoor and outdoor positioning, automatic trigger explanations, and fewer coverage spots. In recent years, artificial intelligence has gradually integrated into people's lives. Integrating mobile phone APPs, artificial intelligence and guides to form a new tour guide model will play an increasingly important role. It can liberate tourists' hands and eyes. It is not necessary for tourists to use their mobile phones to scan and click on and play continuously while scanning. It is safe and convenient, and there are no expensive fees for manual guides, and service quality can also be continuously optimized. In addition, the intelligent electronic tour guide can also change the ticket checking mode of the scenic spot, realize electronic ticket inspection, standardize the tourism market,

accurately obtain the location of tourists, and improve the convenience of the service and the security of the scenic area. It is believed that in the future electronic guides will have a broad space for development.

One important aspect of big data is that the Wikipedia can organize tourism resources. Each site's electronic resources include explanations (voice and textual picture information), and strategies. Travel enthusiasts can share their own strategy of the attraction for other tourists to see. The explanation information of each attraction can be modified. If the tourism enthusiast feels that the interpretation of an attraction is not very accurate or the explanation is not good enough, he can create his own scenic spot to explain the information and go through the review of the background review team to determine whether to modify the current attraction. Over time, it will form some complete electronic resources of scenic spots and constitute a Wikipedia of tourism resources. There may be many other problems in the process of tourists' travel. For example, the inscription in the monuments is vague, or it is difficult to understand the meanings of the monuments. We can use the photo identification technology to compare the tourists' historic sites with the photos of the ancient monuments stored in the system. Since the photographs of the monuments are individually packaged according to the spots, after obtaining the geographical location of the tourists, and only need to make a comparison of the corresponding site database. The corresponding speed and accuracy can be greatly improved, so as to provide better explanation service for tourists.

IV. CONCLUSION

Rural tourism and e-commerce poverty alleviation measures have a better effect than other forms of rural poverty alleviation. After a series of entrepreneurial training or tourism development practices, local farmers have acquired some basic enrichment qualities, and these experiences will be internalized in the awareness of farmers. Among them, the probability that they will be poor again will be reduced. Secondly, in the rural development, we should pay attention to the development of characteristic industries in rural areas and preserve the original ecological service environment. Make full use of local resources, geography, humanities, and develop tourism, and gradually form industrial agglomeration. The use of e-commerce platform for those tourists to provide featured products such as mailing, after-sales and other services, but also can attract more customers to buy or even purchase specialty products in the field, increase the flow of rural tourism, and rural tourism to promote the development of e-commerce, so e-commerce development Reinvest in rural tourism. Based on the data research of Taobao Village and tourism villages, this paper proposes the following relevant recommendations:

- 1) Strengthen infrastructure construction
- 2) Focus on cultivating rural E-commerce talents
- 3) Pay attention to the development of e-commerce associations
- 4) Promote industrial integration

ACKNOWLEDGMENT

First of all, I would like to take this opportunity to express my sincere gratitude to my Professor, who has given me important guidance on the paper. Without his help and encouragement, my paper would have been impossible. Besides his help with my paper, he has also given me much advice on the methods of doing research. Last but not least, I would like to express my gratitude to all the friends who have offered me help. I will continue to do my best in research.

REFERENCES

- [1] B.D. Diao, K.L. Chen, and L. Ding, "Spatial distribution pattern and its influencing factors in Taobao Village of China," *Tropical Geoscience*, vol. 37(01), pp. 56-65, 2017.
- [2] J.Y. Zhao, J.H. Ding, and X.Y. Wang, "Temporal and Spatial Distribution and Evolution of China's Taobao Village under Different Scales," *World Geography Studies*, vol. 26(6), pp. 73-82, 2017.
- [3] C.X. Liu and D.J. Tang, "The Spatial Pattern of Rural Mobility in Zhejiang and Its Influencing Factors: Based on the Analysis of Taobao Village and Tourism Village," *Journal of Zhejiang Agricultural Sciences*, vol. 28(8), pp. 1438-1446, 2016.
- [4] Z.B. Xu and Z.H. Wang, "Spatial distribution characteristics and driving factors of "Taobao Village" in China," *Economic Geography*, vol. 37(01), pp. 107-114, 2017.
- [5] X.Y. Xin and J.J. Xin, "Space-time evolution and formation mechanism of Taobao Village agglomeration," *Areal Research and Development*, vol. 37(01), pp. 11-15, 2018.
- [6] L.L. Cao, X.L. Qin, and X.X. Wu, "Analysis of Factors Affecting the Cluster Effect of "Taobao Village"," *Jiangsu Agricultural Sciences*, vol. 45(12), pp. 311-315, 2017.
- [7] X.S. Shi and T. Wei, "Research on the development model of rural e-commerce industry clusters and the study of spatial differences: A survey of Taobao Village in Jiangsu Province," *Modern Economic Research*, vol. 2017(11), pp. 118-125, 2017.
- [8] M. Falk and E. Hagsten, "E-commerce trends and impacts across Europe," *International Journal of Production Economics*, vol. 170(1), pp. 357-369, 2015.
- [9] C. Cheng and D. Ding, "Internet + Agricultural E-Commerce: The Development Path of Modern Agricultural Informatization," *Information Science*, vol. 34(11), pp. 49-52, 2016.
- [10] Y.Y. Wang, Y. Xie, and M. Wang, "Research on Rural E-commerce Relationship Network and Local Construction in the Background of Precise Poverty Alleviation: A Case Study of Wuhua County, Guangdong Province," *World Geography*, vol. 26(06), pp. 119-130, 2017.
- [11] Y.J. Liu and X.M. Chu, "Research on the Evolution of China's "Taobao Village" Industry," *China Soft Science*, vol. 2017(02), pp. 29-36, 2017.
- [12] J.Y. Li, L. Liu, and J.L. Sun, "Tourism Information Science: A Research Framework," *Tourism Journal*, vol. 2011(6), pp. 72-79, 2011.
- [13] L.Y. Zhang and M. Liu, "Basic Concept and Theory System of Smart Travel," *Tourism Journal*, vol. 27(5), pp. 66-73, 2012.
- [14] L.G. Nie, Y.M. Li, and Y.H. Yuan, "Intelligent travel planning based on clustering analysis algorithm," *Computer Development & Applications*, vol. 25(2), pp. 28-30, 2012.
- [15] L. Wang and L.S. Zeng, "Exploration of Intelligent Cloud Guide System Based on Android Platform," *Electronic Design Engineering*, vol. 20(12), pp. 21-23, 2012.
- [16] M. Kenteris, D. Gavalas, and G. Pantziou, "Near-optimal personalized daily itineraries for a mobile tourist guide," *Computers and Communications, IEEE*, 2010, pp. 862-864.
- [17] D. Abdulrasool and S. Sabra, "Mobile-embedded smart guide for the blind," *Dijon, France: Sprinr Verb*, 2011, pp. 571-578.
- [18] M. Manoufali, A. Aladwani, and S. Alseraidy, "Smart guide for blind people," *International Conference and Workshop on Current Trends in Information Technology, IEEE*, 2011, pp. 61-63.