

Smart City and Public Cultural Resources Demand Based on Big Data: A Case Study of Tianjin Library

Lei Li, Shaojun Ma, and Yiping Wang

Abstract—With the rapid development of data mining and cloud computing technologies, smart cities and public cultural resources are closely integrated. Public resource allocation based on big data is used as a technical means of urban governance and monitoring to continuously meet the diversity and complexity of residents' demand for public cultural resources. For this paper, Geographic Information System (Arc GIS10) and Data Discovery System are used to analyze the spatial distribution characteristics of Tianjin's public cultural resources and explore residents' demand for public cultural resources. Meanwhile, some suggestions for the allocation of public cultural resources will be proposed based on it.

Index Terms—Smart city, public cultural resources, spatial distribution, visitor experience, data mining, library.

I. INTRODUCTION

In the context of data-driven public service innovation, many social organizations, business enterprises and individuals have actively participated in the new public service governance reform. The allocation of urban public service resources has attracted much attention in the new era [1]. In 2017, the rate of China's urbanization reached 58.52%, and some studies have shown that the level of basic public services is closely related to the process of urbanization. However, the 13th Five-Year Plan for Promoting Equalization of Basic Public Services issued by the State Council of China points out some problems including China's poor quality of basic public services and unbalanced development, which make it difficult to satisfy residents' demand for urban public cultural products and services.

"Advocating universal reading and building a learning society" is an important part of the report on the work of the Chinese government. Library is considered as the most common cultural place in urban public cultural resources, and is endowed with a variety of functions (learning, reading and information consultation) as well as social roles (cultural dissemination, communication and sharing).

Manuscript received August 3, 2019; revised February 12, 2020. This work was supported in part by the National Natural Science Foundation of China (Grant No. 71203153; NO. 71874120); Social Sciences Major Project Foundation of Tianjin Education Committee (Grant No: 2017JWZD02) Project of Science and Technology Development Strategy Research and Planning of Tianjin (Grant No. 2017ZLZLZF02000). Fund for Beiyang Young Scholars Program (Grant No. 2018XRX-0023) and the Public Resource Center of Tianjin University.

Lei Li and Shaojun Ma are with the College of Management and Economics, Tianjin University, China (e-mail: lilei@tju.du.cn, mashaojun0212@tju.edu.cn).

Yiping Wang is the Software & Communication School, Tianjin Sino-German University of Applied Sciences, China (e-mail: 985981088@qq.com).

Nowadays, library planning and design innovation are constantly enriching the urban public cultural space form, promoting the upgrading of urban public resources and cultural services, and attracting visitors such as urban residents, tourists and so on. How to build a smart library in the context of "Internet +" to meet the needs of residents for urban public cultural services is a question worth thinking. Generally speaking, it includes two aspects: (1) Spatial distribution: the allocation of urban public cultural resources, which refers to the basis of regional distribution and allocation; (2) Visiting experience: satisfaction of urban public cultural resources, which refers to people's perception of cultural sites, affecting people's judgment on the quality of urban cultural life.

II. LITERATURE REVIEW

A. Public Cultural Resources and Service Supply

Wang S S *et al.* (2019) believed that urban cultural space is closely related to the quality of urban public services and social economic value, however, too much attention has been paid to the current urban economic growth while the balanced development of urban cultural space has been neglected [2]. Zhang T *et al.* (2017) used TOPSIS time variant decision fusion model to make empirical analysis of public service supply in 34 cities in China, finding that the level of government satisfaction with the supply of public cultural services is low [3]. Loughran (2014) found that the public cultural resources of American cities are not as balanced as economic stratification, and public cultural services under neo-liberalism will aggravate the long-standing social public capital [4]. Aptekar S (2015) believed that community gardens are one of the forms of public cultural resources, but the interaction of residents with different levels, races and immigrant status is gradually emerging in diversified and high-end urban communities [5]. Breaking the conflict of public cultural resources will help to change residents' views on urban cultural space (community gardens).

B. Smart City and Public Cultural Resources

Smart cities are closely integrated with the allocation of public cultural resources. The allocation of public resources based on big data has become a technical means of urban governance [6]. ICT and IoT are considered as important measures to enhance the competitiveness of cities [7]. Li S *et al.* (2013) believed that the sharing of public cultural resources is an innovative measure in the construction of urban public cultural service system, and the cultural sharing project based on cloud computing broadens the transmission path of urban public cultural resources,

improves the efficiency of public cultural services, and reduces effectively the management and operation costs [8]. Information and intelligence technology can solve the phenomenon of "mismatch" of public cultural services accurately in time, which can meet the individual and diversified public cultural needs of residents [9]; Koukopoulos Z (2017) proposed a public cultural service platform suitable for intelligent urban environment, providing digital services and intelligent management of cultural heritage online, which can effectively preserve and disseminate tangible and intangible cultural heritage information to meet the needs of residents' public cultural resources [10].

Library is one of the carriers of urban public service supply. OCLC is a global library cooperative organization that manages, integrates and shares library resources of different natures (public libraries, university libraries, professional libraries)^①. At the same time, the status of Library in the public cultural system is becoming increasingly important. Li L (2018) took Tianjin libraries of China as an example to analyze the current situation of public service supply using the Thiessen Polygon method [1]. Ke P (2016) investigated public cultural service institutions/organizations in Yunnan Province by questionnaire, founding that libraries and related cultural service institutions (stakeholders) are partnerships rather than competitive relationships [11]. Therefore, the development strategy of libraries should focus on internal conditions, external environment and stakeholders. Hammond (2002) found that the inefficiency of the operation of the British public library system is caused not only by over-subscription to publications, but also by the service level which needs to be reorganized [12].

In a word, there have been sufficient studies on public cultural services under emerging technologies, especially on library as a public cultural resource. However, there are few literatures to analyze the distribution of public cultural resources from the perspective of residents, and few literatures related to public cultural resources experience perception. Therefore, this study supplements this issue.

III. RESEARCH IDEA AND METHODOLOGY

A. Research Idea

Library is a relatively open public cultural institution that provides residents with a cultural space for reading and enquiry services. It is also an important public cultural resource of the city to enrich and improve the daily life and cultural quality of the residents.

This study uses the location data of urban public cultural resources (taking Tianjin Library as an example), using TF-IDF (term frequency-inverse document frequency) algorithm to calculate the key areas of urban public cultural resources, and summarizes the problems existing in the allocation of urban public cultural resources. Therefore, the study is mainly divided into two sub-studies as follows:

Study 1 (The spatial distribution characteristics of urban public cultural resources)

The first step is to use Baidu map to obtain the latitude and longitude data of urban public cultural resources

(Tianjin Library) and the distribution of major traffic routes in Tianjin. Because the main public transportation mode directly to the library is provincial road and subway route, this study will select the above two indicators for analysis;

Arc GIS is the main tool to visualize the spatial distribution characteristics of current scientific research. The second step is to use the geographic information system software (Arc GIS 10.2) to map the relationship between urban public cultural resources and traffic routes to determine their configuration status.

Study 2 (Study on the visiting experience of urban public cultural resources)

The first step is to use python to crawl the comments of Dianping on urban public cultural resources, then to make data cleaning, correlating the sample of comments with the space area to obtain the most compelling cultural resources;

According to Pareto's Law, the second step is to select 31 libraries with the most comments as the new sample set. Using TF-IDF algorithm to produce the specific perception experience of residents in obtaining urban public cultural resources;

Finally, on the basis of data analysis of the above public cultural resources, combined with the overall layout of urban public cultural resources, traffic layout and experience perception results, the residents' actual demand for urban public cultural space environment and suggestions for resource allocation are proposed.

B. Research Method

Based on the background of big data, IoT and cloud computing, this paper acquires urban public cultural resource data through network reptiles, pattern recognition, data mining and performs spatial coupling analysis on the first-hand data obtained.

Based on TF-IDF algorithm, keywords for evaluation and discussion of public cultural resources are extracted [13]. TF-IDF algorithm is a commonly used weighting technique for information retrieval and data mining. It has achieved successful experience in dealing with different fields in the era of big data [14]. This method is much mature in extracting keywords with different importance levels [15], using to indicate how important a word is to the content of a document text or corpus. TF (Term Frequency) can be used to reflect the frequency of word appearance, and IDF (Inverse Document Frequency) can be used to measure the general importance of words. This algorithm can effectively filter words with no practical meaning to the subject of the document, and highlight the key words at the same time, which can be more accurately and accurately in counting the key words than the traditional simple word frequency statistics.

This paper uses the TF-IDF algorithm to carry out keyword cloud computing on the comments of urban public cultural facilities. Its main formula is:

$$TF\text{-}IDF = TF_{ti} \cdot IDF_t = \frac{n_{ti}}{N} \cdot \log \frac{|D|}{n_t}$$

where N_{ii} is the number of times the word t appears in the document i , and N is the sum of the number of occurrences of all words in the document D , D is the total number of

between library resources accessibility and comments posted by readers online.

C. Data Analysis

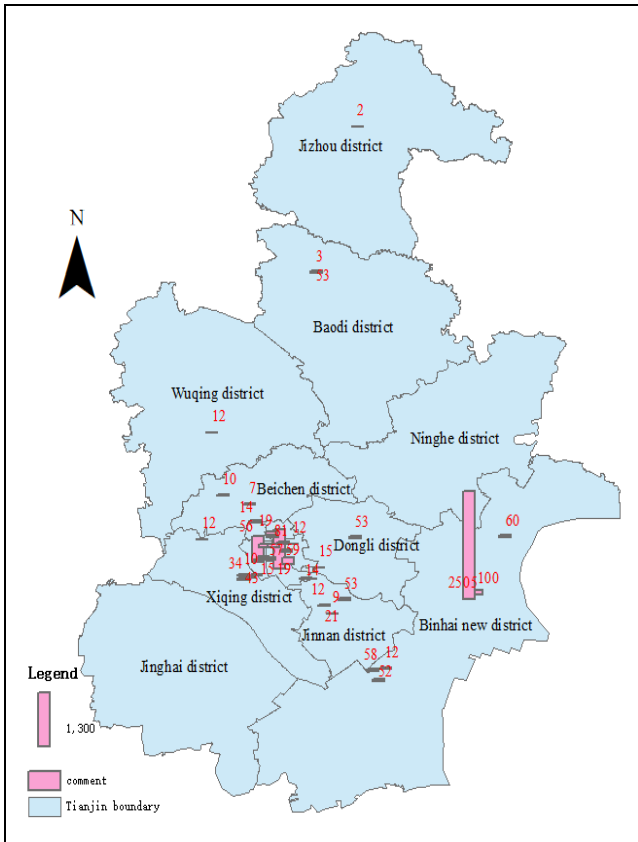


Fig. 3. Review distribution map of Tianjin Library.

First, "Books" and "Study room" are two important parts for residents when visiting libraries. It has a high frequency of words, which reflects the most important needs of residents when visiting libraries, for example:

- "The collection is rich, the environment is very comfortable, the location is near the cultural center, and the surrounding human environment is very deep, so children like to come here.
- "There is a study room on the first floor, and a lot of people study in. I find the feeling of college in the library. But there are more children on weekends."

Secondly, the transportation facility around the library are an important factor in the residents' demand for public cultural resources, appearing keywords such as "Parking Area" and "Light Railway", residents have the following comment:

- "The library is far from the city center, and the traffic is not very convenient, so I finally chose self-driving, parking area is very large, which can park a lot of cars."

In addition, the aggregation of public cultural resources can enhance the satisfaction of residents' public cultural demand. For example, the comment of the Tianjin Library (Pingjiang Road) is based on the perspective of the cultural center:

- "After the opening of the subway, it is more convenient to go to the library, which is adjacent to the Tianjin

Grand Theatre, Tianjin Art Museum and Tianjin Museum. The surrounding transportation is convenient, and there is an underground parking lot."

It's worth noting that we found that some words appeared, such as "New Area", "Architectural Style", "Internet Celebrity", "Good-Looking" and "Art Gallery". These words all come from Binhai Library. Combined with the introduction of the library, we find that it was officially launched in October 2017. With large collection of books and ingenious design, it has become popular on the Internet. It is known as "the eye of Binhai", and since then, it has attracted thousands of residents to visit and experience.

Combined with the word frequency map of the visiting experience of Tianjin Library, this study objectively and deeply analyses the characteristics of each category contained in the two core subjects (residents and libraries) from the perspective of residents' experience (Fig. 5). This map reflects that the supporting facilities of the library will directly affect the residents' visiting experience, including books, self-study places, atmosphere and seats of the library. In addition, it also includes such factors as traffic convenience, appearance design and so on.

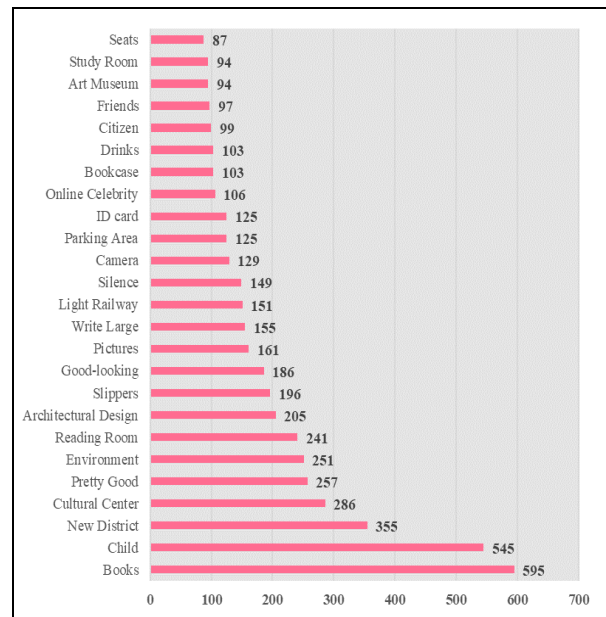


Fig. 4. Frequency chart of visiting experience based on TF-IDF algorithm.

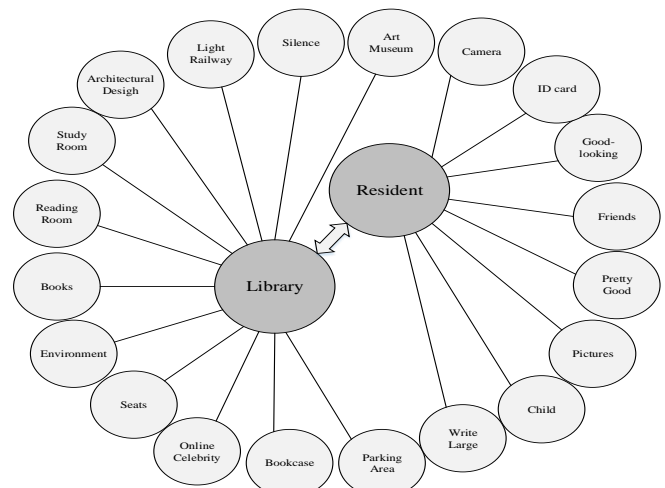


Fig. 5. The characteristics of visiting experience of library in Tianjin.

V. DISCUSSION

Combined with the data mining and data analysis above, we can make the following conclusions according to the spatial distribution of public cultural resources in Tianjin and the residents' visiting experience.

A. Conclusions for Results

1) The allocation of public cultural resources is still dominated by the central city

From the perspective of the spatial distribution of public cultural resources, this paper takes Tianjin library as an example to find the geographical coordinates of Tianjin administrative area and library resources. The allocation of public cultural resources in Tianjin is mainly distributed in the central city. In addition, the public cultural resources of Binhai new district with national strategic significance are gradually improved. Therefore, in the construction of smart cities, it is still necessary to first implement the goal of equalizing public services, and continuously satisfy the needs of residents to obtain public cultural resources.

2) Public cultural resources rely on the accessibility of transportation

This paper firstly focuses on the transportation vehicles frequently used by residents, using the mapping between library distribution characteristics and traffic routes to explore the spatial distribution of public cultural resources. The research shows that the spatial distribution of Tianjin library is arranged around public transportation; the accessibility of the main modes of transportation of residents is an important basis for the distribution of libraries. Therefore, to obtain the public cultural resources of residents, we need to improve the construction of public infrastructure around public cultural resources.

3) Residents have a high level of satisfaction with public cultural resources

This paper starts with the residents' ratings and comments on public cultural resources, and uses the big data web crawler and TF-IDF algorithm to obtain the satisfaction of residents' public cultural resources. The study finds that residents generally have higher satisfaction with Tianjin library experience. In addition to the richness of the library's collection resources and the comfort of the learning space, the library's design and intelligent system have become another important reason for attracting residents to visit.

B. Limitations

First of all, this paper mainly conducts data mining and cloud computing from the text of online review. In view of the incomplete information of the residents participating in the review, it is impossible to comprehensively and deeply analyze the correlation between demographic characteristics and other variables. Future research can be conducted from the online questionnaire, and the combination of research, quantitative and qualitative analysis can also be used for demonstration. Secondly, this study only uses Dianping as the data source. In the future, the corresponding research will widely use a variety of social media data to analyze, combined with multi-source data and check-in with geographic location information. The data can be studied to attract more visitors from different regions, so as to more

scientifically and comprehensively summarize the behavioral characteristics of residents visiting the library. Finally, this study uses TF-IDF algorithm to analyze the word frequency of library public resources comments, but it can only analyze the visiting experience of Smart City Library from macroscopic analysis, and fails to analyze word frequency connection and specific reasons from microscopic analysis.

NOTES

①<https://www.oclc.org/>

②<http://www.dianping.com/>

③The method of merging similar data in this study is as follows: first, determine the geographical location coordinates of each library and sort them according to this, and secondly, compare and analyze the libraries of the same geographical coordinates, and find the information of the library's website and the names used in the face of uncertain public resources.

④ The downtown area of Tianjin includes six administrative regions: Heping District, Nankai District, Hedong District, Hexi District, Hebei District and Hongqiao District.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Lei Li conducted the research; Shaojun Ma and Yiping Wang performed the experiments and analyzed the data; Shaojun Ma designed the structure and wrote the paper.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support provided by the National Natural Science Foundation of China (Grant No. 71203153; NO. 71874120); Social Sciences Major Project Foundation of Tianjin Education Committee (Grant No: 2017JWZD02) Project of Science and Technology Development Strategy Research and Planning of Tianjin (Grant No. 2017ZLZLZF02000). Fund for Beiyang Young Scholars Program (Grant No. 2018XRX-0023), the Public Resource Center of Tianjin University, and the Postgraduate Research and Innovation Program in Tianjin (No.2019YJSB185, 2019YJSB186).

REFERENCES

- [1] L. Li, J. Kong, and C. Gu, "Research on the configuration of urban public cultural facilities: A case study of Tianjin libraries," in *Proc. the 4th International Conference*, ACM, 2018.
- [2] S. S. Wang, Y. Zhang, and Y. Q. Xu, "A quantitative analysis of inequality of urban cultural space distribution in Xi'an," *Science China (Technological Sciences)*, 2019.
- [3] T. Zhang and F. Wang, "TOPSIS time variant decision fusion model evaluation for internet of public service things," *Cognitive Systems Research*, 2018.
- [4] L. Kevin, "Parks for profit: The high line, growth machines, and the uneven development of urban public spaces," *City & Community*, vol. 13, no. 1, pp. 49-68, 2014.
- [5] S. Aptekar, "Visions of public space: Reproducing and resisting social hierarchies in a community garden," *Sociological Forum*, vol. 30, no. 1, pp. 209-227, 2015.

- [6] L. Li, T. Hao, and T. Chi, "Evaluation on China's forestry resources efficiency based on big data," *Journal of Cleaner Production*, 2016, pp. 142.
- [7] C. Nijkamp, "Smart cities in Europe," *Urban Insight*, vol. 18, no. 2, pp. 65-82, 2011.
- [8] S. Li and D. Wang, "The study of culture sharing project based on cloud computing," in *Proc. International Conf. on Business Intelligence & Financial Engineering*. IEEE Computer Society, 2013.
- [9] Y. Ying, T. Xialing, and T. Wei, "Study on governmental cultural resources purchase management based on public information behavior big data," in *Proc. International Conf. on E-education*, ACM, 2017.
- [10] Z. Koukopoulos, D. Koukopoulos, and J. J. Jung, "A trustworthy multimedia participatory platform for cultural heritage management in smart city environments," *Multimedia Tools and Applications*, 2017.
- [11] P. Ke, X. Chen, and Y. Liu, "Library strategic environment in the public cultural service system in China," *Library Management*, vol. 37, no. 4/5, pp. 195-209, 2016.
- [12] C. J. Hammond, "Efficiency in the provision of public services: A data envelopment analysis of UK public library systems," *Applied Economics*, vol. 34, no. 5, pp. 649-657, 2002.
- [13] W. Zhang, T. Yoshida, and X. Tang, "A comparative study of TF* IDF, LSI and multi- words for text classification," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2758-2765, 2011.
- [14] H. C. Wu, R. W. P. Luk, and K. F. Wong, "Interpreting TF-IDF term weights as making relevance decisions," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 131-137, 2008.
- [15] W. X. Rao, L. Chen, P. Hui, "Move: A large scale keyword-based content filtering and dissemination system," in *Proc. the 32nd International Conference on Distributed Computing Systems*, 2012, pp. 445-454, Macau, China: IEEE.

unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



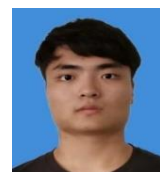
Lei Li was born in Qingdao, Shandong province in 1980, and he is a professor in the College of Management and Economics at Tianjin University now.

His research interests include urban sustainable development, intelligent urban agglomeration construction, public resources, and performance management. He has published in technological forecasting and social change, sustainability, journal of cleaner production, ecological indicators, environmental engineering & management journal, mathematical and computer modelling, and transport policy.



Shaojun Ma was born in Henan province in 1993, and he is a Ph.D student in the College of Management and Economics at Tianjin University.

He has participated in national and international conferences and published papers in different journals. His research interests include public policy analysis and management, public service resources allocation, and data mining in urban agglomeration.



Yiping Wang was born in Henan province in 1998, and he is a university student in Software & Communication School at Tianjin Sino-German University of Applied Sciences. His research interests include resources allocation and data mining.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits