IMPACT: International Journal of Research in Business Management (IMPACT: IJRBM);
ISSN(Print): 2347-4572; ISSN(Online): 2321-886X
Vol. 10, Issue 8, Aug 2022, 1–16

THE ANALYSIS ON TOURIST SOURCE MARKET OF CHIKAN ANCIENT TOWN BY COUPLING GIS, BIG DATA AND NETWORK TEXT CONTENT ANALYSIS

Jun-Jie Tanm & Ruei-Yuan Wang

Research Scholar, Guangdong University of Petrochem Technology, Sch Sci, Maoming - 525000, People's Republic of China

Received: 29 Jul 2022 **Accepted:** 02 Aug 2022 **Published:** 04 Aug 2022

ABSTRACT

The China cultural exhibition tourism of Chikan Ancient Town (CAT) overseas is the key construction project of Guangdong Province, also is the no.1 cultural tourism industry project of Jiangmen city. To ensure the sustainable development of tourist attractions, people should deeply realize the critical factors such as tourists' cognition, emotional expression and satisfaction of the tourist destination, so as to grasp the psychological characteristics of tourists and their tourism behavior, and to continue expanding the TSM. Based on the support of Big Data, Network Text Content Analysis (NTCA) and the application of Geographical Information System (GIS), this paper comprehensively analyzes the TSM of CAT from the perspective of macro market and tourists' characteristics by accessibility analysis, NTCA and other methods. The study shows that the TSM of CAT is mainly the central and western cities in the Pearl River Delta (PRD). The place is the main TSM with the highest degree of tourist inflow, while other cities are relatively low but have great market potential. The semantic network of tourist comments presents a three-layer structure of "core layer-secondary core layer-edge layer". Finally, the paper puts forward development and management suggestions to improve the awareness, satisfaction and brand image recognition of CAT.

KEYWORDS: *Geographical Information System (GIS); Big Data; Network Text Content Analysis (NTCA);Chikan Ancient Town (CAT);Tourist Source Market(TSM)*

INTRODUCTION

The analysis of tourist source market (TSM) is of great significance for a new tourist attraction, which is related to the design of commercial development scale of tourist area, the economic benefits of pan-tourist area and whether it can meet the expectation of project construction. Domestic and foreign research on TSMs focus on influencing factors, space-time structure, market segmentation and market prediction (Quan et al., 2012). Huang et al. (2021) used Network Text Content Analysis (NTCA) and ROST Content Mining (ROSTCM6) software to study tourists' image perception of Xiaoyaojin Park in Hefei. Zhou et al. (2019) divided the tourist market potential of China's World cultural heritage into five levels by using GIS technology. Hu et al. (2014) analyzed the tourist potential of small and medium-sized cultural tourism cities based on SPSS correlation analysis and ArcGIS spatial analysis technology.

After consulting a large amount of literature, it is found that most of the previous analysis methods for TSM are analyzed by SPSS analysis method (Hu et al.,2014), build matrix model (Xiao,2010), Shift—Share Method(SSM) analysis(Xiao,2010), questionnaire survey method(Bai,2017;Li,2017;Su,2016;Tai and Qiu,2011)and SWOT analysis(Su

and Ma,2005). In recent years, more researchers use network texts evaluation to analyze tourist sentiment of a scenic spot (Fan et al., 2021; Wang and Lu,2016;Tang and Wang,2016). For the analysis of the potential of TSM, GIS technology is mostly used (Zhou et al., 2019; Hu et al.,2014). However, for the tourist attractions in CAT, because we don't know the playability of the new scenic spot and the lack of tourist evaluation, it is not reliable to take the evaluation of different types of scenic spots in the same region as the analysis conditions of the tourist market of the new scenic spots.

In view of this situation, unilateral analysis of TSM cannot correctly understand the real TSM of a tourist destination, but using Big Data, analysis (Wang et al., 2021), combined with GIS technology analysis, macroscopically, we can analyze the tourist market of new scenic spots in a certain area from the most basic tourist source place, tourist source emotion and tourist source accessibility. CAT is a tourist attraction of "old scenery reconstruction", there are both existing know ability–previous tourists' comments on old scenic spots and unknown – style, type and visibility of the reconstructed landscape. It is more appropriate to analyze the potential tourist market of CAT with the application of GIS technology supported by Big Data.

STUDY SIT AND DATA

The Processing Framework

This paper uses GIS technology, data crawling and Big Data, analysis technology to achieve accurate and targeted data acquisition, and then processes the obtained basic data in accordance with specific methods to form usable data for research. Taking all kinds of factors into consideration, this paper analyzes the tourist market of CAT by using spatiotemporal analysis and tourist group analysis. The research process based on GIS analysis and Text Semantic Analysis (TSA) includes data collection, classification, preprocessing, analysis, conclusion drawing and other steps by web crawler tools. The overall processing framework is shown as Figure 1.

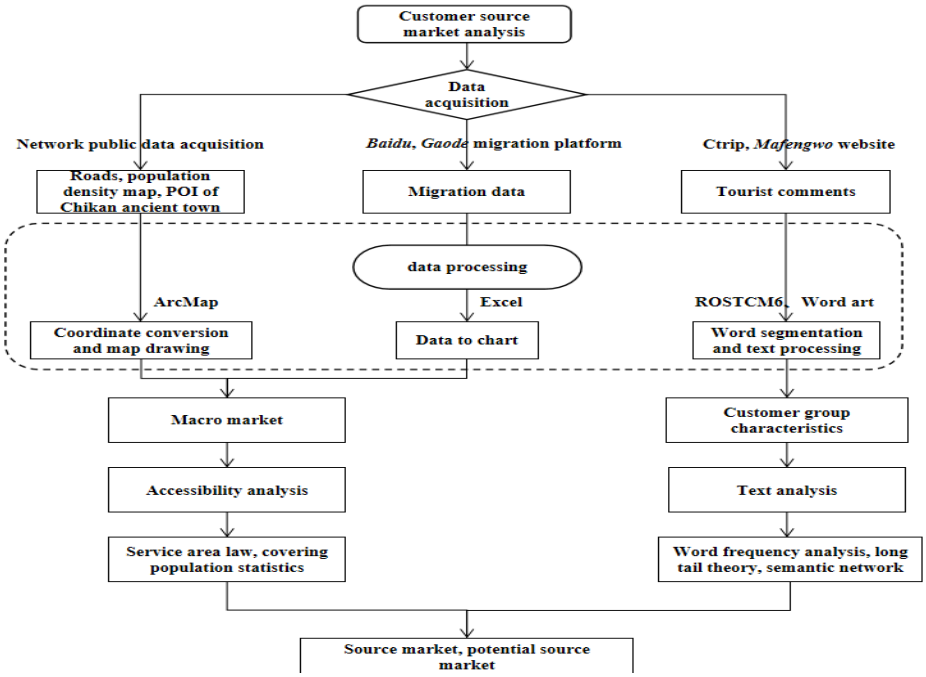

Figure 1: The Overall Processing Framework.

Study Site

CAT is located in Kaiping City, Jiangmen City, Guangdong Province (shown as Figure 2). With a history of more than 360 years, CAT is a town full of China characteristics, a national key town and a famous town of China history and culture (Chen and Lv, 2021). CAT has gathered buildings with China and Western characteristics during the period of the Republic of China as well as strong cultural customs of Wuyi area. It has been the location of many films, but with the passage of time, the protection of historical relics and buildings is facing great challenges. In order to promote the development of tourism economy in Kaiping and respond to the task of revitalizing rural tourism on the 18th National Congress of the Communist Party of China, the Project of Overseas China cultural tourism exhibition in CATwas officially launched in 2014, which is of great significance to tourism development, cultural relic protection and regional economic development. According to the latest progress of the project, CAT will be completed by September 2022, and partial trial operation will be carried out at the end of 2022.

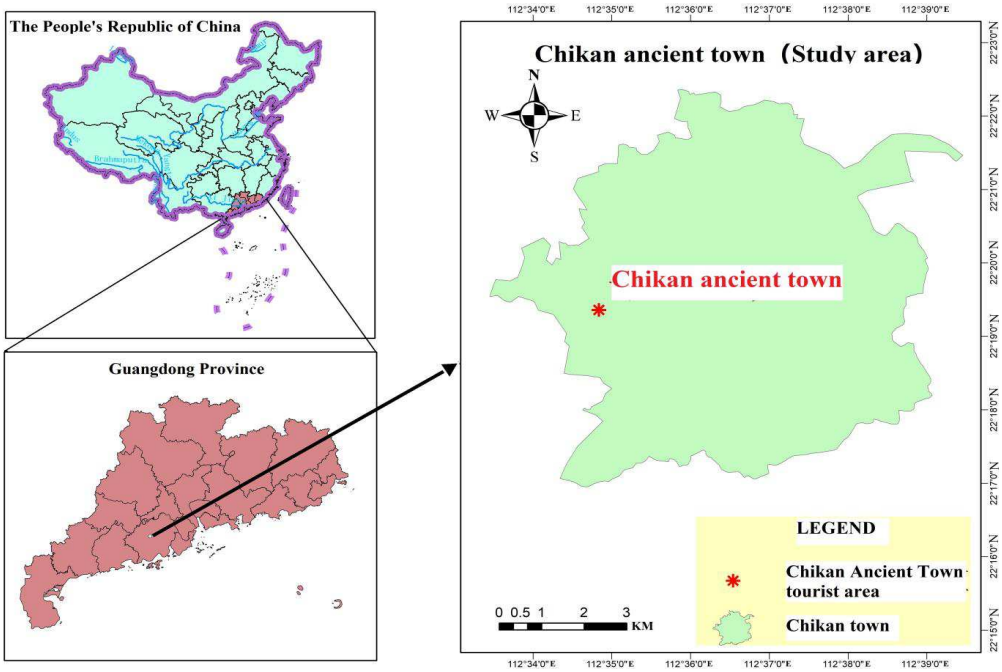

Figure 2: The CAT of Study Area.

Source of Data

The study data mainly falls into two categories: Firstly, the Big Data, of various websites crawled based on crawler technology, including tourist comments of "old scenery" of CAT, source data of people moving into Jiangmen city during the National Day holiday in 2021, POI (Point Of Interest) of CAT. Secondly, National road, provincial road (2021), vector data of Guangdong county boundary (July 2021), Population density map of Guangdong Province (2020). In addition, tourist comments comes from Ctrip (www.ctrip.com) and Mafengwo (www.mafengwo.com), Acquisition time: April 14, 2022,Passenger flow in Jiangmen migration migrating data from Baidu platform (www.qianxi.baidu.com) and Gaode migration Big Data, platform (https://trp.autonavi.com/migrate/page.do), Acquisition time: May 1st, 2022, The POI of CAT comes from planning cloud encyclopedia(http://guihuayun.com/poi/). Vector data such as national highways and population density maps of Guangdong Province are both from online public data.

METHODS AND ANALYSIS

Processing of Tourist Comment Data

Due to the limitation of the collector used to crawl data, some tourist comments could not be displayed or follow-up analysis could not be carried out. In order to prevent such data from causing large errors to the study analysis, 523 valid comments were obtained by manually removing the invalid data from a total of 533 comments.

After the pre-processing of the tourist comments mentioned above, the text mining tool (ROSTCM6) developed by Professor Shenyang was used to conduct vocabulary segmentation and high-frequency vocabulary statistics for the tourist comments extracted. The results of initial vocabulary segmentation are not ideal, and some of the low-frequency effective vocabulary cannot be separated out normally. After adding valid vocabulary to the "Custom vocabulary list", the effect of vocabulary segmentation is significantly improved, and the vocabulary frequency analysis of the text after vocabulary segmentation is carried out. In this paper, only the entry frequency greater than 10 are listed (Table 1as below).

Table 1: Chikan Ancient Town's Overall Cognitive High-Frequency Vocabulary

Ranking	Entry	Vocabulary Frequency	Ranking	Entry	Vocabulary Frequency
1	Ancient Town	388	41	Having an Antique Flavour	21
2	Architecture	197	42	Zili Village	21
3	Kaiping	173	43	Amorous Feelings	20
4	Characteristic	143	44	Hour	20
5	Arcade	142	45	Visit	19
6	Watchtower	115	46	Unfortunately	18
7	History	76	47	Ancient Architectural Buildings	18
8	Style	69	48	Jiangmen City	17
9	Local	65	49	Ancient	17
10	Film and Television City	61	50	That Year	17
11	Film	54	51	Whole	17
12	Scenic Spot	53	52	Guangzhou	17
13	Guangdong	52	53	Many Years	16
14	Culture	51	54	Convenient	16
15	Shot	50	55	Full	15
16	Rice Hot Pot	48	56	Tofu	15
17	Taste	46	57	Specialty	15
18	Local	45	58	Building	14
19	Travel	44	59	Road	14
20	China	40	60	The Republic of China Era	14
21	hometown of Overseas Chinese	33	61	Proposal	13
22	Tofu Horn	33	62	Scenery	13
23	Retain	33	63	Modelling	13
24	European Style	32	64	Nothing	13
25	Admission Ticket	30	65	Close	12
26	Let the Bullets Fly	29	66	Opposite Bank	12
27	China and Western Combined	29	67	Kilometer	12
28	Opening to the Outside World	28	68	Travel	12
29	Century	27	69	Style Street	11
30	South of the Five Ridges	27	70	Having Dinner	11
31	Movies	27	71	A Large Number	11

Table 1: Contd.,

32	Preservation	25	72	Eel with Rice	11
33	Reform	25	73	Museum	10
34	Business	25	74	Really	10
35	Feel	23	75	Not Big	10
36	The Grandmaster	23	76	Ecology	10
37	Photograph	23	77	Whole	10
38	Development	22	78	Wind and Rain	10
39	Books	21	79	Scene	10
40	Government	21	80	Playful	10

Processing of Migration Data

Since the impact of the Corona Virus Disease 2019 (COVID-19) outbreak, the tourism industry in China's everywhere has suffered a heavy blow, and the number of tourists in many tourist destinations has dropped sharply, thus the relevant data is no longer representative. In order to better conform as well as let the analysis results in this study are more consistent with the development status of tourism industry in the context of pre-epidemic research. Based on the migration data during the National Day holiday of 2021 when the epidemic was at its low point, this study analyzes the potential market of CAT.

Because the data obtained by Gaode migration Big Data, platform is in days, the data of a certain day cannot fully reflect the migration situation in this period of time. Therefore, this study takes the data of the 7 days on the 2021 National Day holiday (i.e. from October 1 to 7) and uses the mean value of data as the reference index of migration intention in this period. And presented in the visual form of charts, as shown in Figure 3.

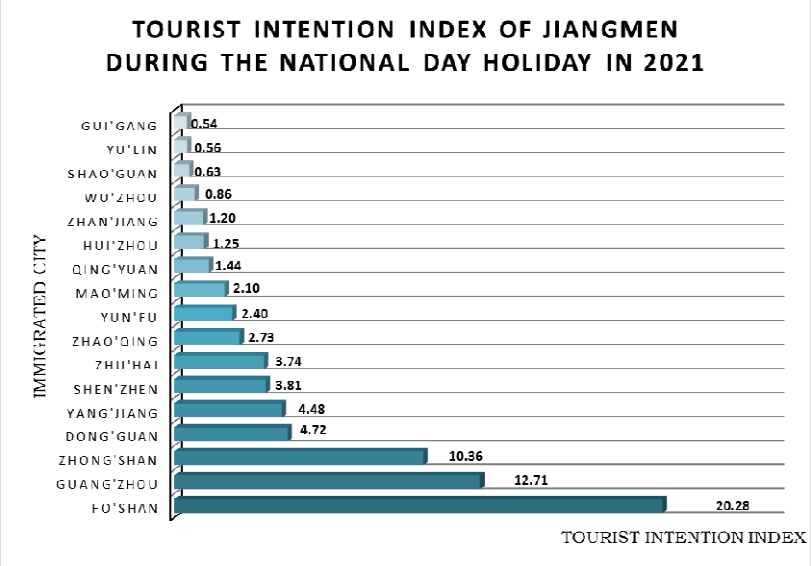

Figure 3: Tourist Intention Index of Jiangmen during the National Day Holiday in 2021.

Processing of Vector Data

For POI (Point Of Interest) data of CAT extracted from planning cloud encyclopedia, the attached coordinates are Baidu coordinates. QGIS software is used to convert Baidu coordinates into WGS_1984 by using relevant plug-ins, and text coordinates are converted into vector layers in the QGIS. For vector layers of provincial road, national road and Guangdong county boundary obtained from the network, their coordinates are uniformly transformed into geographical coordinates WGS_1984. After processing the above vector data, the following figure is obtained (shown as Figure 4).

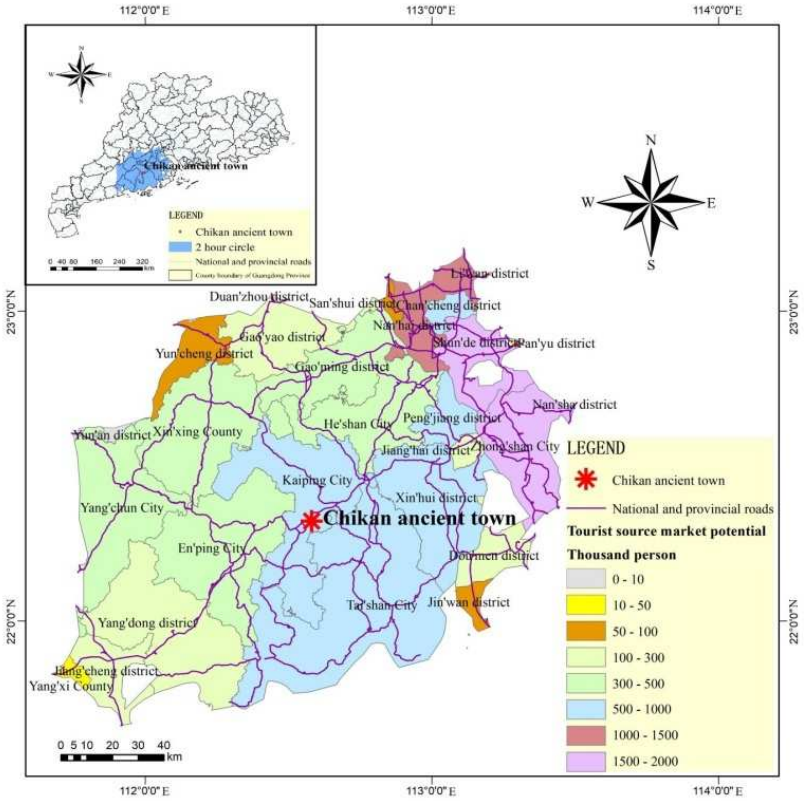

Figure 4: The Map of 2-Hour Circle TSM Analysis.

Accessibility Analysis Method

Accessibility is used to measure the spending time from a point in the region to the target tourist attractions, characterize the convenience of tourists to the tourist attractions in the region, and help to clarify the relationship between tourist attractions and transportation networks. This study calculates the accessibility of the target tourist attractions based on the traffic network, and combines the population density data to present an advanced analysis of the accessibility area obtained by a certain assignment to cover the population statistics. The primary study ideas are the combination of the following two methods.

Service Area Method Based on GIS Network Analyst

Using the Wizard tool, create a network dataset based on the network vector data of ShapeFile synthesized by national and provincial roads, define the network source data and its role in the network, and specify the connectivity and network attributes in the network. Then, create the target point of the calibration of the network location tool, independently assign the driving speed of the simulated car as V, calculate the time required to pass through each line segment based on this speed in the road network attribute table, and take this time as the impedance attribute to create the network data set. The main calculation formula is as follows

$$t = L/v \text{.....(1)}$$

In 1, ‘L’ is the network distance from a node in the built road network to the target point, and ‘t’ is the time from the node to the target point at the speed of ‘v’. An isochronal circle centered on the target point is established.

Zonal Statistics as Table Method

Using the "display partition statistics in table" tool in ArcMap to obtain the data set in the grid layer - pixel (COUNT), area (AREA), minimum value (MIN), maximum value (MAX), D-value (RANGE), mean (MEAN), standard deviation (STD), sum (SUM). The principle (as Figure 5) is to compute multiple statistics using all predefined statistics and output them in a table.

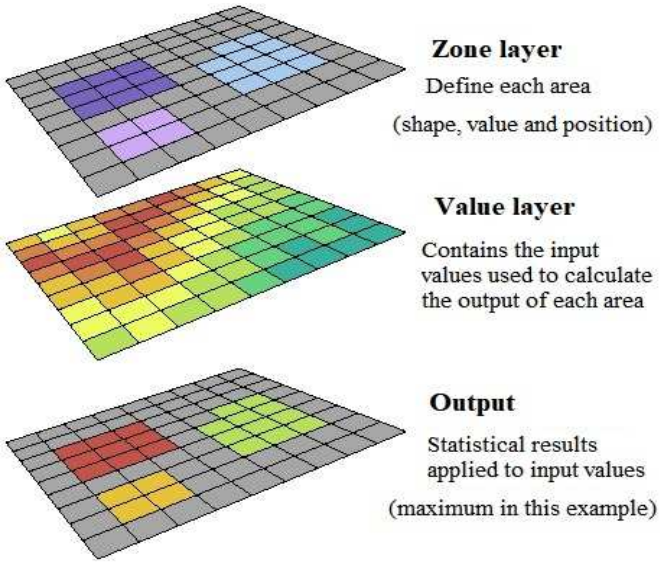

Figure 5: Sample Input and Output of Partition Statistics (Source: https://desktop.arcgis.com, Modified)

Text Analysis Method

Text analysis is a study method for objective and quantitative descriptions of explicit content (Zhou and Qiao, 2022). The biggest advantage of text analysis is that it can obtain tourists' complete psychological perception. In this paper, ROSTCM6 text mining software is used for vocabulary segmentation, high-frequency vocabulary statistics and semantic network construction of text files. This method quantifies the language of tourists and presents it in the form of figures and charts.

Vocabulary Frequency Analysis (VFA)

VFA is a kind of statistical analysis for the number of important vocabulary in the text to find the core information hidden in the text content. Thus, VFA is a relatively elementary but very effective text mining method, which conducts quantitative analysis of text content through statistics of high-frequency vocabulary (Sun and Ni, 2018).

Semantic Network Analysis (SNA)

SNA is proposed in 1968 by J.R.Quillian as an explicit psychological model of human associative memory. SNA is mainly based on VFA, focusing not on vocabulary themselves, but on the relationship between various vocabularies. This method can deconstruct the semantic path between syntax and concept of web text content, so as to identify the association and meaning of text vocabulary (Zhou and Tian, 2017).

RESULTS ANALYSIS

Macro Market Analysis

Relying on the powerful spatial analysis ability of GIS, the real road network is simulated to construct the traffic isochron in a scientific and reasonable way, and the graphs of the radiating population in the spatial range are obtained by combining geospatial Big Data, and population density data (Figure 4 and Table 2).According to the China Travel Vehicle and Boat Association, the number of self-driving tourists in China accounted for 77.8% of the total number of domestic tourists in 2020. With the improvement of people's quality of life, the per capita car ownership in China is increasing year by year, and the number of self-driving tourists will continue to grow in the next few years. Thus, the accessibility analysis of scenic spots is particularly important for the analysis of the tourist market. This paper focuses on the potential analysis of the tourist market

According to the time spent by vehicles on the way to travel, the tourism market is divided into three levels. The first level is the tourism market with a two-hour drive, the second level is the four hour drive tourism market, and the third level is the tourism market with a driving distance of more than four hours(Li and Huang,2021). In terms of self-drive travel, more tourists tend to take short trips, and tourists from the first level of the tourism market occupy the majority. Many places even formulate corresponding "two-hour travel circle" to attract tourists. For a tourist destination, those who have not travelled are its potential tourists, and those who have travelled will come again. Therefore, the potential source market can be regarded as the total population within a certain range of tourist destinations (Zhou et al.,2019). As can be seen from Figure 4 and Table 2, within the two-hour circle established with CAT tourist area as the center, the main potential city (county) level markets are Zhongshan city, Shunde district and Nanhai district, each with more than one million potential tourists. The second is the surrounding areas of CAT, such as tourists of Kaiping city, Taishan City, Xinhui District, Pengjiang District, etc.

CAT is located within the jurisdiction of Jiangmen city, and the intention of Jiangmen tourists to move in can indirectly reflect the source of tourists in CAT. Figure 6 shows that during the National Day holiday in 2021, the willingness indexes of tourists from Yangjiang city and Yunfucity to move to Jiangmen are 4.48 and 2.40 respectively, ranking only the 5th and 9th among all cities. It can be seen from Figure 6 that during the National Day holiday in 2021, the proportion of tourists actually moving into Jiangmen from Yangjiang city and Yunfucity is 3.24% and 2.00% respectively, ranking 4th and 8th among all cities. Based on the above analysis of immigration intention and actual immigration, it is found that the number of tourists from Yangjiang city and Yunfucity who move in Jiangmen city is very low, which is different from the number of potential tourists shown as Table 2 (Yangjiang city: 847.87 thousand persons, accounting for 9.04%; Yunfu city: 410.73 thousand persons, accounting for 4.38%) has a great difference. Only from the perspective of accessibility, both Yangjiang city and Yunfucity are potential major TSMs of CAT, and the market is highly developable, which is particularly important for CAT to formulate good planning and publicity to attract tourists from these two major tourist markets in the future.

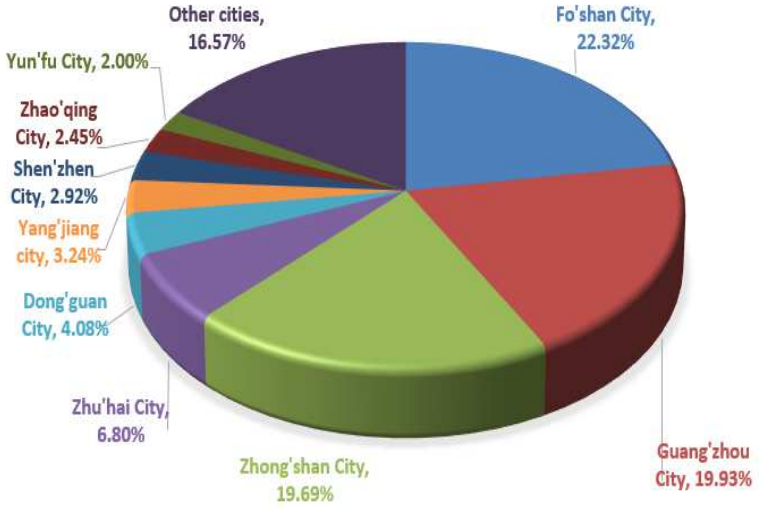

Figure 6: The Proportion of Actual Passenger Flow Sources in Jiangmen during the 2021 National Day Holiday.

Table 2: Hour Circle TSM Analysis Table

Prefecture Level City	County Level City / Municipal District	Tourists Market Potential/(Thousand Persons)	Total/(Thousand Persons)	Ratio
Zhong'shan city	-	1801.54	1802.54	19.22%
Yang'jiang city	Yang'chun city	303.99	847.87	9.04%
	Yangxi County	13.05		
	Yang'dong District	295.42		
	Jiang'cheng District	235.41		
Zhao'qing city	Gao'yao District	278.45	287.19	3.06%
	Duan'zhou District	8.74		
Zhu'hai city	Jin'wan District	72.49	320.22	3.42%
	Dou'men District	247.73		
Jiang'men city	En'ping city	387.34	3405.00	36.32%
	He'shan city	397.41		
	Kai'ping city	532.94		
	Tai'shan city	707.66		
	Xin'hui District	642.39		
	Jiang'hai District	128.01		
	Peng'jiang District	609.25		
Guang'zhou city	Nan'sha District	0.30	26.38	0.28%
	Pan'yu District	14.42		
	Li'wan District	11.66		
Fo'shan city	Gao'ming District	386.80	4078.78	43.50%
	San'shui District	52.72		
	Shun'de District	1676.56		
	Nan'hai District	1157.31		
	Chan'cheng District	805.39		
Yun'fu city	Luo'ding city	0.97	410.73	4.38%
	Xin'xing County	328.58		
	Yun'an District	5.80		
	Yun'cheng District	75.38		

Tourist Group Evaluation Analysis

Vocabulary Frequency Analysis (VFA)

Python crawler technology is used to preprocess all the online comment data of CAT tourist attractions which obtained from the website (Jia, and Zhai, 2021), and then obtain the vocabulary frequency in Table 1. Using the VFA data in Table 1 and the wordle analysis method, the data visualization diagram is made with the boundary of Chikan town as the scope (shown as Figure 7). It is not difficult to see from Figure 7 that shows the significance of tourist comment data keywords in different forms, such as frequency, font, color, etc., the larger the font, the more tourists evaluate it, and the higher the evaluation frequency, indicating that tourists pay more attention to the content (Cheng et al.,2020),thus reflecting the importance of these keywords to tourist attractions. From Figure 7, it can be clearly seen that “Ancient towns”, “Architecture”, “Characteristic”, “History”, “Watchtower”, “Kaiping” and other comments are keywords with a high probability of occurrence, which are very prominent in the visual performance of the wordle visualization, and are the key contents of tourists' concern captured in the huge tourist comment data.

By sorting out the high-frequency vocabulary in Table 1 and Figure 8, it can be found that the high-frequency vocabulary for tourists' cognition of CAT is primarily divided into three categories, including destination vocabulary, landscape vocabulary and gourmet vocabulary.

Figure 7: Visualization Diagram of Wordle Data of Tourist Comment Content Information.

Destination vocabulary: “Ancient town”, “Kaiping”, “South of the Five Ridges”, “Zili village”, “Guangzhou” and other vocabulary appear more frequently, indicating that tourists have a strong identity with tourist destinations and have a clear understanding of the region where CAT is located. The vocabulary "Guangzhou" appears in related comments like "There was Guangzhou Bay in ancient times, and there is Chikan Old Street now." "It's similar to the arcade in Guangzhou, but there are differences." "It's similar to Beihai Old Street in Guangzhou, on a smaller scale." and so on. These comments all point out that the landscape of CAT is similar to many places in Guangzhou, indicating that the unique style characteristics of CAT of overseas China have not been displayed to the greatest extent. For some tourists, it lacks freshness and uniqueness, lacks attraction, and cannot create a positive primacy effect.

Landscape vocabulary: “Arcade”, “Watchtower”, “Film and Television City”, “Ancient Architectural Buildings”, “Style street” etc. Comments include: "European Style Street is still very beautiful, with buildings with China and Western combined and China stone bridges." "The town is all arcade buildings, ancient and quaint, quite local characteristics and style." and so on. It shows that the main reason for attracting tourists to visit is the characteristic landscape of south of the Five Ridges Wuyi area in CAT, tourists can get fun from the antique southern town, enjoy the scenery and human feelings. And “Watchtower” is the gold card of Kaiping's tourism industry, and the tourist attractions related to "watchtower" are very close to CAT, the vocabulary has appeared many times in the tourist comments of CAT, which shows that for most tourists, CAT is not the only tourist attraction they visit here, but a way to visit multiple scenic spots including CAT and Kaiping watchtower. For the development of tourism in Kaiping City, the establishment of tourist areas with CAT and Diaolou scenic spots as the core is a major potential development direction of tourism development.

Gourmet vocabulary: “Rice Hot Pot”, “Tofu horn”, “Eel with rice”, “specialty” and other high-frequency gourmet vocabulary, which highlight the characteristics of local cuisine. The appearance of such high-frequency vocabulary indicates that the focus of tourists here is mainly reflected in the local cuisine culture except for the features of the scenic spot. And the frequency of such words is so high, which shows that for tourists, the local characteristic cuisine has high satisfaction, strong acceptance and a good reputation. It is a major link in the tourist consumption structure, and the characteristic cuisine has become an important potential market to attract tourists. Therefore, in the follow-up development and operation process, CAT should increase the integration degree of characteristic cuisine and scenic spot, so as to make it a famous brand in the development of scenic spots.

The Long Tail Theory (LTT) is a new theory proposed in 2004 by The American Chris Anderson. For a long time, people only pay attention to the top 20% of the products sold and define them as the mainstream, while the LTT believes that the remaining 80% of the products have a huge potential market, which is called "niche market"(Cheng et al.,2020). In VFA, the LTT is still applicable, and low-frequency long-tail vocabulary other than high-frequency vocabulary is still the worth focus of attention. Through long-tail vocabulary analysis, tourists' product cognition can be further identified. Among the four guarantees of people's tourism, "clothing, food, lodging and transportation", among the 523 comments collected, there is only one entry about lodging- "Hotel", and the word frequency is only 4 times, which reflects from the side that in the past, tourists' demand for " lodging" during tourism activities in CAT was not very high, due to tourists being mainly one-day tours, and there were few overnight people. Secondly, CAT is only less than half an hour's drive from the Kaipingdowntown, so tourists have more accommodation options. However, in the reconstruction process of CAT, "lodging" is not a factor that doesn’t need to be considered at all.

With the improvement of China economy, people have a higher pursuit of tourism accommodation, and then appeared the concept of "Boutique B&B(Bed and Breakfast)". Boutique B&B is an important direction for the diversified, personalized and high-quality development of the B&B industry (Hou and Hu,2022). For CAT, which strives to build a new landmark of cultural tourism in ancient towns in China, the emergence of boutique B&B can meet the needs of tourists seeking high-quality travel. At the same time, tourists can experience more local cultural characteristics from boutique B&B. Whether for CAT or Kaiping overseas China hometown culture, this is an important path for cultural output. Therefore, in the accommodation development of CAT project, we should focus on refinement and delicacy, have a unique style, and pursue quality rather than quantity, so as to develop the middle class crowd in the TSM.

Semantic Network Analysis (SNA)

Although the main characteristics of CAT can be reflected through word frequency analysis and LTT analysis, it cannot reflect the relationship between the central vocabulary and marginal vocabulary and the structural relationship between the overall vocabularies. By drawing a network semantic map, the relationship between various elements can be intuitively displayed, which is conducive to mining the characteristics of the tourists group at a deeper level (Cheng et al.,2020). In the SNA, vocabulary segmentation is firstly carried out on the tourist comment text of CAT, then high-frequency vocabulary are extracted and some meaningless vocabulary are filtered, and the following semantic network graph is finally generated through the function of "Social network and Semantic analysis" in ROSTCM6, as shown in Figure 8.

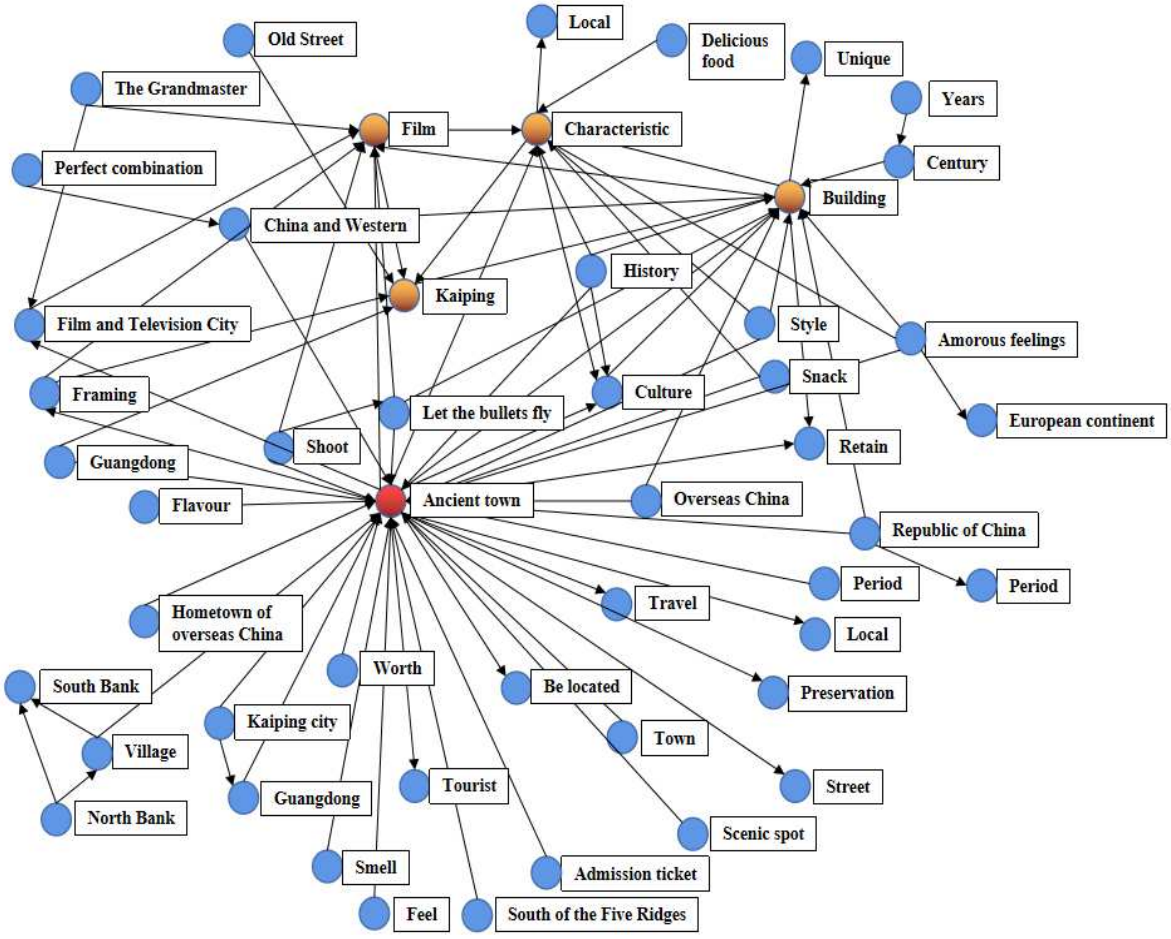

Figure 8: Semantic Network Analysis (SNA) Diagram of the Overall Image of CAT.

From the perspective of hierarchical structure, the semantic network diagram mainly presents the hierarchical structure characteristics of "core - margin" and "core - sub core - margin", and multi-layer margin subgroups or sub core subgroups are formed around important nodes in the diagram.

The first layer is the core layer, and the vocabulary only contains "ancient town". As the analysis object of this paper is CAT, the vocabulary "ancient town" has the highest frequency and is the core. The levels of "ancient town - be located" and "ancient town - hometown of overseas China" with the structure of "core - margin" reflect that tourists have a high awareness of the location and style of the ancient town.

The second layer is the secondary core layer, containing "Kaiping", "characteristics", "building", "film" and other vocabulary. Chikan films and television cities used to be the shooting place of many film and television dramas. This feature can be seen in hierarchical structures such as "film - The Grandmaster" and "film - Let the Bullets Fly", indicating that characteristic buildings are an important reason for the location of film and ancient town in modern times, and is deeply recognized by tourists. Therefore, the tourist market aimed at film and television culture has great development potential television works. Film and television culture is a unique regional characteristic culture developed in Chikan.

The outermost layer is the margin layer, which mainly includes vocabulary such as “Delicious food”, “Years” and “Style”, reflecting tourists' all-round cognition of CAT in terms of positioning, facilities and culture. This is consistent with the above research conclusions.

CONCLUSION

In this study, GIS technology is used to analyze the first-level TSM of CAT from a macro perspective such as accessibility, network distance, and zonal statistics. Meanwhile, based on the text analysis model, tourists' overall cognitive image of CAT before reconstruction is studied, and the following conclusions are drawn:

- Based on the macro analysis of time cost and actual situation, the main TSMs of CAT are Foshan city, Zhongshan city, Zhuhai city, Yunfu city, Yangjiang city and Guangzhou city, etc. According to the analysis, Yunfu city and Yangjiang city are the most exploitable tourist markets.
- CAT is a town with China characteristics and a famous historical and cultural town, but its scenic spot fails to highlight its characteristics as the hometown of overseas China to the maximum, and the TSM has different perceptions of the overall cultural image of CAT.
- Tourists who visit the scenic spot of Kaiping Diaolou and Villages can become the potential source market of CAT in the near future in the case of satisfaction.
- According to the analysis of the basic guarantee of tourists' travel, tourists have a greater sense of identity for the cuisine of CAT, and do not show a strong tendency for lodging.
- Tourists prefer the film and television culture as well as characteristic architecture of CAT. Tourism image, traditional architecture and cultural landscape are the most popular tourist scenes in CAT.
- Combined with the main research results, the following development suggestions are proposed.

Explore New Tourists’ Source Market and Improve Tourists’ Source Structure

The economic development has the characteristics of vulnerability and are easily affected by other factors. A stable TSM can ensure the steady economic development of CAT, but for CAT, sustainable development is the basis for maintaining economic vitality. Therefore, for the potential TSM, local and government should introduce corresponding policies to attract such tourists. For example, Shaoguan issued relevant policies and 17 A-level scenic spots were open to Dongguan citizens for free .Guizhou's preferential policies for tourists from Guangdong, Shandong, Chongqing and other places not only promote the economic development of Guizhou, but also expand the travel options of Guangdong tourists, achieving a win-win situation.

Strengthen the Integrated Marketing Way, Expand the Spatial Structure

Although CAT got a large number of repeat visitors in the past business process, the new tourism resources still have a continuous attraction. The government can promote the reconstruction of tourism space of CAT, optimize the integration of resources, form a joint force with surrounding tourism resources, establish a tourism circle, shorten the travel time of tourists and improve the comfort level of travel.

Improve Tourism Service Facilities and Optimize Consumption Structure

The local cultural characteristics of the diet should be deeply explored, developing characteristic cuisine, standardized management of local specialties in the hometown of overseas China, reject the "three no" products (respectively are products without manufacturer name, production site and production hygiene license code) and crack down on fake and shoddy products. Reasonably develop lodging places represented by high-quality B & B, and effectively ensure their quality. Develop overall tourism from point to area and create an all-round tourism market, so as to expand the TSM.

Although this study can simply reveal the tourist market of CAT tourist attractions at the level of accessibility and customer group evaluation by relying on Big Data, and GIS technology, due to the lack of relevant data at the level of social economy and population structure, the accuracy of the research cannot be further improved, and it is impossible to propose the optimal tourist market development strategy. However, due to the in-depth analysis of the first level (main) tourist market and the real perception of tourists in CAT, this survey still has a certain reference basis for the further development and management of the overseas China cultural tourism exhibition project in CAT.

ACKNOWLEDGEMENTS

The author is grateful for the research grants given to Ruei-Yuan Wang from GDUPT Talents Recruitment (No.2019rc098), Peoples R China under Grant No.702-519208, and Academic Affairs in GDUPT for Goal Problem-Oriented Teaching Innovation and Practice Project Grant No.701-234660.

REFERENCE

1. *Bai, Y. (2017). Investigation and analysis of domestic tourist market in Chaoyang City. Journal of Bohai University (PHILOSOPHY AND SOCIAL SCIENCES), (04), 73-76. doi:10.13831/j.cnki.issn.1672-8254.2017.04.017*
2. *Chen, Q., &Lv, S. (2021). How to solve the financing difficulties of cultural tourism projects. Jiangmen Daily, A10*
3. *Cheng, Y., Zhu, J. &Mou, X. (2020). Product cognition, emotional expression and satisfaction of ski tourists -- Text Analysis Based on online evaluation. Journal of physical Education, (04), 15-21. doi:10.16419/j.cnki.42-1684/g8.2020.04.004*
4. *Fan, M., Chen, Y., & Zhang, H. (2021). Hunan tourism source market structure and destination terrain image perception – An Empirical Analysis Based on geographical marker photos. Economic Geography, (12), 223-232. doi:10.15957/j.cnki.jjdl.2021.12.024*

5. *Hu, H., Li, R., Hu, R. & Zhou, L. (2014). Analysis of tourist potential in small and medium-sized cultural tourism cities – from the perspective of tourism support. Journal of Arid Land Resources and Environment, (06), 183-188. doi:10.13448/j.cnki.jalre.2014.06.031*
6. *Huang, Y., Pan, H., & Li, Q. (2021). Research on image perception of Hefei Xiaoyaojin Park Based on network text analysis. Sichuan Architecture, (01), 4-7.*
7. *Hou, Y., & Hu, H. (2022). Analysis of spatial distribution characteristics and driving forces of boutique B & B in Yangshuo County. Journal of Guilin University of Technology*
8. *Jia, Y., &Zhai, J. (2021). Visual analysis of tourist comment data based on Python crawler technology. Journal of Anyang Normal University, (05), 51-54. doi:10.16140/j.cnki.1671-5330.2021.05.013*
9. *Li, T. (2017). Survey and analysis of Pingliang health tourism market. Tourism Overview (second half of the month), (08), 67-69.*
10. *Li, L., & Huang, D. (2021). Industrial market analysis of Characteristic Towns under Rural Revitalization -- Taking Shaoguan hot spring characteristic town as an example. Modern Rural Science And Technology, (08), 1-2.*
11. *Quan, H., Zhao, L., Chen, T.,& Yang, Z. (2012).Empirical analysis of the market structure of inboundtourists– a case study of Jiangsu Province. Economic Geography, (01), 146-152. doi:10.15957/j.cnki.jjdl.2012.01.009*
12. *Su, H., and Ma, Y. (2005). Analysis of tourist source market based on affinity and SWOT method - Taking the British tourist source market in Xi'an as an example. Journal of Arid Land Resources and Environment, (05), 146-149*
13. *Su, A. (2016). Market survey and current situation evaluation of tourist sources in Guilin in the era of high-speed rail. Tourism Overview (the second half of the month), (24), 153+155*
14. *Sun, X., & Ni, R. (2018) Product cognition, emotional expression and brand image perception of Chinese Cruise tourists -- content analysis based on online comments. Geographical Research, (06), 1159-1180*
15. *Tang, Y., & Wang, Q. (2016). Analysis on the structure and behavior mode of tourist market in Langya Mountain scenic spot of Chuzhou City. Journal of Chifeng University (Natural Science Edition), (07), 113-115. doi:10.13398/j.cnki.issn1673-260x.2016.07.048*
16. *Tai, X.,& Qiu, F. (2011). Research on the tourist market of Xi'an Datang Furong garden. Economic Research Guide, (13), 177-178+201*
17. *Wang, Z., & Lu, M. (2016). Analysis of spatial structure, tourist behavior and perception of tourism market in Xiangxi Prefecture. Commercial Science Research, (04), 67-74*
18. *Wang, R. Y.; Lin, P. A.; Chu, J. Y.; Tao, Y. H.; Ling, H. C. A decision support system for Taiwan's forest resource management using Remote Sensing Big Data. Enterprise Information Systems, 2021, 1-22. https://doi.org/10.1080/17517575.2021.1883123*

19. *Xiao, L. (2010) Research on the development strategy of inbound tourism in Hangzhou Based on the analysis of tourist source market. (Master's thesis, Zhejiang Normal University),https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2011&filename=2010241307.nh*
20. *Zhou, X.,&Qiao, S. (2022) Research on destination ecological environment perception based on Web big data text analysis -- a case study of Yalong Bay tourist destination in Sanya. Modern Business, (05), 27-29. doi:10.14097/j.cnki.5392/2022.05.056*
21. *Zhou, Y., Deng, Z., Shi, J., Ma, D. & Zheng, X. (2019). Research on the market potential of China's world heritage tourism destinations. Economic Geography, (04), 216-222. doi:10.15957/j.cnki.jjdl.2019.04.026*
22. *Zou, Y., & Tian, F. (2017). Image of Nanchang tourism space in online virtual community. Resources Science, (02), 314-324*