

Discriminating Homecoming Visitors Using Mobile Signaling Data

Yonggang Tang^{1, 2, a}, Huijie Lin^{3, b*}, and Mingjian Mo^{4, c}

¹School of Tourism Data, Guilin Tourism University, Guilin, Guangxi, China;

² Guangxi Key Laboratory of Culture and Tourism Intelligent Technology, Guilin, Guangxi, China.

³ School of Outdoors Sports, Guilin Tourism University, Guilin, Guangxi, China.

⁴ Guangxi Tourism Data Center, Guilin Tourism University, Guilin, Guangxi, China.

^a tangyg@gltu.edu.cn, ^b 75814813@qq.com, ^c 573748615@qq.com

Abstract. This study utilizes mobile signal data as an emerging data source, aiming to construct a framework that can efficiently distinguish homecoming visitors from other types of tourists. The rich behavioral features contained in mobile signal data are deeply analyzed. A two-stage clustering method based on machine learning, which integrates multiple aspects such as feature engineering, model construction, and performance evaluation is designed to ensure the accuracy and generalization ability of classification results. The main data source is the signaling data of the three major telecommunications operators in China. Using the signaling data from China's three major telecommunications carriers as the primary data source, this study verified the number of tourists in the Guangxi, as well as in Nanning, Guilin, and Yulin cities during the 2019 National Day Golden Week. The results indicate that the method presented in this paper offers a viable new approach for the statistical analysis and data mining of large-scale tourism figures.

Keywords: Homecoming visitors; Mobile signal data; Machine learning; Tourist classification; Analysis of Tourism Behavior

1. Introduction

1.1 Background and Significance

In tourism statistics, the scale of tourists and tourism revenue are the two most core indicators, with the latter being closely linked to the former. Therefore, understanding the scale of tourists in a timely and accurate manner becomes the most important issue in tourism statistics. Traditional manual statistical methods, based on statistical reporting and sampling surveys, suffer from issues such as inaccurate estimation of tourist scale and low statistical efficiency[1]. Tourism enterprises and management departments are eagerly looking forward to new statistical methods, especially the application of new technologies such as big data. Big data has the characteristic of objectivity, and its credibility is much higher than that of survey questionnaires. Incorporating it into tourism statistics is an inevitable trend of development. However, the data directly provided by the data sources may not effectively represent the complete research target group. For example, data from telecommunications operators contains a large amount of non-tourist and duplicate data. If not carefully selected and cleaned, or if the model used is inappropriate, it will often lead to incorrect statistical results. Therefore, it is necessary to distinguish between different types and characteristics of mobile phone users and adopt different methods and models for statistics.

Returning home tourism can be seen as a form or type of tourism under the category of visiting relatives and friends[2]. Visiting friends and relatives (VFR) tourism is an important but often overlooked segment of the travel industry. Prior to the late 1980s, there was scant research in this field. However, after the mid-1990s, a trend of studying VFR tourism arose in Western academia. Some research literature clearly demonstrates that the scale of the VFR tourism market is quite considerable[3], and for certain destinations, VFR tourism even constitutes the main part of their tourism market[4]. Jackson argues that due to the methods used in traditional tourism statistics, the size of the VFR tourism market has been greatly underestimated[5].

Among the tourists visiting relatives and friends, homecoming visitors—individuals who return to their hometowns or places of origin for a variety of reasons such as family gatherings, nostalgic revisits, or attendance at special events—constitute a unique segment that requires special consideration. In China, the impact of homecoming tourists is even more significant due to the economic development that has led to a large amount of people working away from home.

The activities and consumption behaviors of homecoming visitors are quite different from those of ordinary tourists[6]. Currently, tourism statistics do not distinguish this type of tourist but treat them the same as other tourists, and there is not enough understanding of their activity patterns and consumption characteristics. Especially during holidays, a large number of homecoming visitors will make the tourism statistical data show obvious abnormalities.

Mobile signaling data has emerged as a powerful tool for discriminating homecoming visitors from other tourist segments. This type of data, collected from cellular devices, provides rich insights into individuals' mobility patterns, allowing researchers to analyze travel routes, stay durations, and visit frequencies with high precision. By harnessing the potential of mobile signaling data, destination management organizations can gain a comprehensive understanding of homecoming visitors' travel behaviors, preferences, and needs.

In summary, the ability to accurately identify and distinguish homecoming visitors using mobile signaling data holds significant implications for destination management, marketing, and sustainable tourism development. By leveraging this data, destination managers can gain deeper insights into the travel behaviors and preferences of this unique segment, enabling them to develop more targeted strategies that enhance the visitor experience, drive economic growth, and foster community engagement.

1.2 Current Research Status

Visiting friends and relatives tourism, as a global phenomenon, has great potential for growth worldwide and has gradually become a new research topic in the Western academic community in recent years. However, China's research in this field lags far behind that of foreign countries. There is little research on tourism related to homecoming tourism. This has also led to the current lack of a clear definition for homecoming tourists—making it difficult to accurately distinguish and count them.

In fact, the lack of effective technological means to distinguish and count different types of tourists is one of the significant reasons why it is difficult to conduct segmented research on tourists. In traditional tourism classification research, the primary sources of data include surveys, transaction records, and social media content[7-12]. Although these methods reveal to some extent the behavior and preferences of tourists, they each have their limitations. For instance, surveys are often limited by sample size and may not fully reflect the overall situation[13, 14]; transaction records focus more on consumer behavior, neglecting other activities of tourists; social media content, while rich, may be influenced by users' tendencies to present themselves, and may not truly reflect the actual behavior of tourists. More importantly, these traditional methods are often coarse in spatial and temporal resolution, making it difficult to accurately capture the detailed activity trajectories of tourists.

Mobile signal data, with its high resolution, wide coverage, and real-time characteristics, provides researchers with a new perspective to observe and analyze tourist behavior. Such data can continuously track the movement of individuals, thereby more accurately revealing the activity patterns and preferences of tourists[15]. This data, with its high spatio-temporal resolution and wide coverage, presents an unprecedented opportunity to study tourist behavior at a fine-grained level.

Therefore, in our research, we attempt to overcome these limitations of traditional tourism classification research by combining mobile signaling data with advanced machine learning algorithms to dynamically and accurately identify different types of tourists. This paper integrates multi-dimensional data such as entry and exit timestamps, duration of stay, activity trajectories,

daily routines, and mobile device usage preferences to identify mobile phone users and determine whether they belong to the category of homecoming tourists.

Identifying tourists within a crowd is the primary issue that needs to be addressed in visitor classification research based on signaling data. The most basic method is to identify tourists through their mobile phone registration location[16, 17]. However, the registration location of a tourist may not necessarily coincide with their place of residence, and for privacy protection reasons, it is also inconvenient to obtain this information. In the absence of mobile registration data, the duration of stay and the scope of activities are often considered important criteria for judgment[18-20]. These methods are based on mobile phones and their signaling data for counting tourist numbers and analyzing tourist characteristics, but they do not achieve the distinction or identification of tourists returning home for family visits. Currently, few technology has been found that classifies, identifies, or analyzes the characteristics of tourists returning home for family visits based on mobile signaling data. The majority of studies have focused on other aspects of tourist behavior, such as movement patterns within a destination[21] or visitor profiling based on demographic characteristics[22]. This gap in the literature highlights the need for further exploration and innovation in utilizing mobile signaling data for the identification of homecoming visitors. Exploration in this field is still in its infancy, with many unresolved issues and challenges. For example, how to effectively extract useful tourism behavior characteristics from massive data? How to accurately distinguish returning tourists from other types of tourists? These questions are all in urgent need of further in-depth exploration by researchers.

2. Methodology

2.1 Data Collection and Preprocessing

Telecommunication operator data is primarily acquired through the collection of signaling data, with the key aspect being the development of a signaling collection system. This system is required to process data quickly, have a high collection recognition rate, and be capable of covering a full range of signaling data for 2G/3G/4G voice, data, and SMS, as well as to provide real-time parsing and application. Signaling data is mainly collected through the operator's A, Mc, C/D/E, and other interfaces, while also mining data from various sources such as LBS (Location-Based Services), BOSS (Business and Operations Support Systems), and the O, B, and M domains. After cleaning, transformation, and filtering, the data is structured into a formatted dataset.

The fields obtained from signaling data are shown in Table 1- Table 4:

Table 1. XDR signaling field obtained from Circuit Switch Call

Field	Type	Length	Example
Interface type	int	4 bytes	2G: 11 3G: 31
Start time	datetime	8 bytes	2016-01-18 09:59:39.1779610
Talk time	int	4 bytes	30 seconds
IMSI	bigint	8 bytes	460002678530469
IMEI	bigint	8 bytes	356395749230909
Calling number	bigint	8 bytes	13915568402
Called number	bigint	8 bytes	13387945678
Call type	int	4 bytes	The caller: 1; Called: 2
Business Start LAC	int	4 bytes	20813
Business Start CELL	int	4 bytes	17798
Business termination LAC	int	4 bytes	10834
Business termination CELL	int	4 bytes	25431

The start time and call duration indicate the time when the current XDR record occurs and the duration of the call, which are used to support data time statistics and big data analysis for inferring the changes in tourists' locations. The IMSI and user number are used to identify specific users to

avoid duplication, and the IMSI, as a unique identifier within the system, is used to associate and analyze all user records. Additionally, the called number supports the analysis of tourists' contacts, which plays a key role in the clustering analysis of tourists' profiles. The business start LAC-CELL and business end LAC-CELL provide effective user location information, offering necessary trajectory support for the analysis of tourists' profiles from the perspective of scenic spots.

Table 2. XDR signaling field obtained from location area update

Field	Type	Length	Example
Time	datetime	8 bytes	2016-01-18 09:59:39.1779610
IMSI	bigint	8 bytes	460002678530469
IMEI	bigint	8 bytes	356395749230909
User number	bigint	8 bytes	13915568402
Update type	int	4 bytes	0: Normal position update; 1: Periodic position update; 2: IMSI attachment; 3: IMSI Separation
LAC before update	int	4 bytes	20813
LAC after update	int	4 bytes	20852
CELL after update	int	4 bytes	25431

LAU (Location Area Update) serves as the foundational data for tracking the location of users in an idle state, continuously updating records of changes in the user's LAC (Location Area Code) position. Therefore, the above LAU fields are all crucial and necessary information fields for big data analysis oriented towards tourists.

Table 3. XDR signaling field obtained from Short Message Service

Field	Type	Length	Example
Time	datetime	8 bytes	2016-01-18 09:59:39.1779610
Calling number	bigint	8 bytes	13915568402
Called number	bigint	8 bytes	13387945678
IMSI	bigint	8 bytes	460002678530469
IMEI	bigint	8 bytes	356395749230909
MO or MT identification	int	4 bytes	0: MO(Sending process); 1: MT(Receiving process)
Business Start LAC	int	4 bytes	20813
Business Start CELL	int	4 bytes	17798

The SMS (Short Message Service) as a terminal service provides basic user information concurrently with the occurrence of the service, such as user identity, user location, and the time of service occurrence.

Table 4. XDR signaling field obtained from Packet Switching Domain

Field	Type	Length	Example
Start time	datetime	8 bytes	2013-05-07 11:53:03.3840000
End time	datetime	8 bytes	2013-05-07 11:58:03.3840000
IMSI	bigint	8 bytes	460002678530469
IMEI	bigint	8 bytes	356395749230909
User number	bigint	8 bytes	13915568402
Terminal manufacturer	int	4 bytes	
Terminal standard	int	4 bytes	2/3/4G terminal, TDD/FDD dual-mode support
Business type	int	4 bytes	
Sub business	int	4 bytes	
HOST	varchar	256 bytes	
Network identification			
LAC	int	4 bytes	20852
CELL	int	4 bytes	25431
User IP address			
Server IP address			

Upstream traffic	int	4 bytes	540
Downward traffic	int	4 bytes	220
URL address	varchar	256 bytes	md.openapi.360.cn/list/get?product=newsreader&version=1
REFER			
COOKIE			
USER AGENT			

The XDR data structure of user PS domain data service activity records is crucial for the application support of the entire tourism big data analysis. Currently, the frequency and diversity of users' data service usage are very high, which is essential for the analysis of tourist behavior.

2.2 Identification methods for homecoming visitors

To accurately identify homecoming visitors, the method adopted in this paper is as follows: targeting specific areas (with administrative counties and districts as the smallest units), using telecommunications operator signaling data as the data source, combined with base station engineering parameters, and supported by road network data. The method involves judging whether mobile phone users are returning home visitors to the target area based on their trajectories in the target area during specific periods (statutory holidays or artificially set periods).

Due to the massive volume of signaling data, HDFS is used for storage, and the model is calculated using Spark. The specific steps are as follows and represented in Fig. 1:

Step 1: Data Collection. Collect LTE signaling data from S1-MME and S1-U interfaces, CS domain signaling data from 2G/3G, and after filling in the number's registered location, save the user number, IMEI, IMSI, interface type, signaling time, and base station engineering parameters to HDFS for subsequent calculations on the Spark cluster.

Step 2: Identifying Regular Users in the Target Area. Use the accumulated signaling data from the three months prior to the specific period, combined with base station engineering parameters, to accumulate the monthly and daily stay time of mobile users in the target area. If a mobile user stays in the target area for more than 20 days per month and more than 6 hours each day, mark this mobile user as a regular user.

Step 3: Determine the Work and Residence Area of Regular Users in the Target Area. Among the mobile users marked as regular users in Step 2, use the working time (Monday to Friday, 8:00-18:00, or according to the common working and off hours in the target area) and rest days (Saturday and Sunday) as time scales, and use K-Means and DBSCAN clustering algorithms for spatial clustering of mobile users' geographic locations, combined with base station engineering parameters. Due to the large number of signaling data, if full data is used directly for spatial clustering, it requires high computational resources and low computational efficiency. Therefore, a two-stage clustering method is adopted in this step: First, use the K-Means clustering algorithm for individual regular users to obtain the work and residence area range of each regular user; Second, use the DBSCAN algorithm to cluster the work and residence area ranges of all regular users to obtain the work and residence area range of regular users in the target area.

Step 4: Determination of Passing Users. Match the latitude and longitude of the base stations in the base station engineering parameters with the geographic coordinates of the road network lines (including highways, secondary roads, high-speed railway lines) in the target area to obtain a road network base station data set covering the road network lines, which includes data such as province, city, county, road, base station number, latitude, and longitude. If a mobile user switches between base stations in the road network base station data set continuously during the specific period and the total stay time in the target area is less than 6 hours, mark it as a passing user.

Step 5: Determination of Tourists. Mobile users who enter the target area during the specific period and stay for more than 6 hours, if they are not regular users marked in Step 2, and are not passing users marked in Step 4, mark them as tourists.

Step 6: Establish a Pending Data Set for Returning Visitors. Take one year as a period, count the users marked as tourists in each specific period of the year, and gather their data together to establish a basic tourist data set (including the IMEI, IMSI, specific period, and stay days of the mobile users marked as tourists). Then filter out the tourists who have not appeared in the residence area range of the regular users determined in the target area in Step 3 from the basic tourist data set, to obtain a filtered tourist data set (including the IMEI, IMSI, specific period, stay days, and residence area stay days of the mobile users marked as tourists). On the basis of the filtered tourist data set, match the activity trajectory of the mobile user before and after the corresponding specific period in each record with the geographical coordinates of the transportation hub and the road network lines to obtain the mode of transportation and the time of entering and leaving the target area for the mobile user, and add it to the corresponding record to obtain a pending data set for returning visitors (including the IMEI, IMSI, specific period, stay days, residence area stay days, mode of transportation, and entering and leaving time of the mobile users marked as tourists).

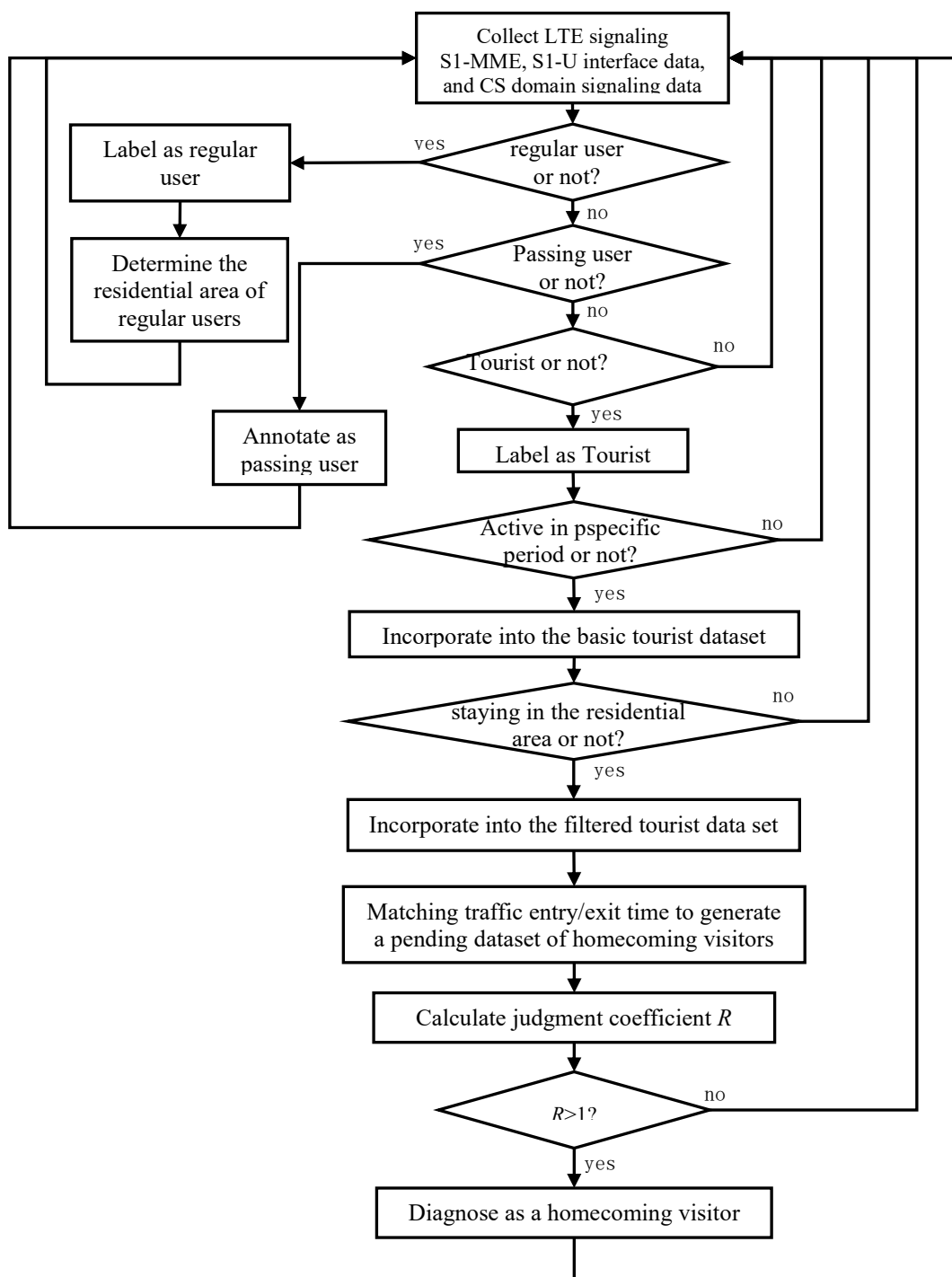


Fig. 1 Flowchart for determining homecoming visitors

Step 7: Calculate the Judgment Coefficient. In the pending data set for returning visitors obtained in Step 6, determine each mobile user marked as a tourist. The judgment coefficient R is calculated as follows:

$$R = \sum_{i=1}^n J_i + T_i + U_i + L_i \quad (1)$$

Where R is the judgment coefficient for determining whether the tourist is a returning visitor, n is the number of records containing the tourist in the pending data set for returning visitors, J_i , T_i , U_i and L_i are the specific period weight coefficient, stay time indicator coefficient, residence area stay time indicator coefficient, and traffic entering and leaving indicator coefficient for the i -th record, respectively.

The specific period weight coefficient J_i is set according to the proportion of returning visitors in the previous year's specific period compared to the total number of tourists (The weights are assigned in descending order based on historical experience: Spring Festival > National Day > May Day > Qingming > Mid-Autumn Festival > others).

The stay time indicator coefficient T_i is set as

$$T_i = \frac{D_{stay}}{D_{total}} \times J_i, \quad (2)$$

where D_{stay} is the number of days the visitor stayed in the target area; D_{total} is the total number of days of the specific period.

The residence area stays time indicator coefficient U_i is set as

$$U_i = \frac{D_{resi}}{D_{total}} \times J_i, \quad (3)$$

where D_{resi} is the number of days the visitor stayed in the residence area.

The traffic entering and leaving indicator coefficient L_i is set as

$$L_i = N_{io} \times W_{tra} \times J_i, \quad (4)$$

where the interval value N_{io} is the number of times the mobile user enters and leaves the target area within $\pm x$ days of the specific period, and x is calculated based on the number of days in the specific period and the specific period weight coefficient J_i ; The weights for the mode of travel W_{tra} is set according to the proportion of tourists who used that particular mode of transportation (such as airplanes, trains, self-driving, etc.) to return home in the previous year, relative to the total number of tourists.

Step 8: Determination of Homecoming Visitors. If the R value is greater than 1, it is determined to be a homecoming visitor; otherwise, it is another type of tourist.

3. Empirical Analysis

3.1 Study Area and Data

To validate the aforementioned method, we selected tourists in the Guangxi Zhuang Autonomous Region of China, along with the cities under its jurisdiction—Nanning, Guilin, and Yulin—during the National Day Golden Week of 2019 as the subjects of our study. We identified homecoming visitors and conducted a statistical comparison of visitor numbers. The statistical period was from October 1st to 7th, 2019. The geographical scope of the statistics covered the entire Guangxi region as well as the administrative boundaries of Nanning, Guilin, and Yulin.

3.2 Results and Discussion

Table 5 presents a comparison of the statistical data of homecoming visitors and other types of tourists identified from mobile signaling data across the Guangxi region.

It should be specifically noted that, to facilitate the examination of tourists' contribution to tourism consumption, this paper adopts person-days as the unit of statistics, which means the number of people multiplied by the number of days stayed, rather than the number of visitors commonly used in official statistics.

The total number of homecoming visitors is 1.9946 million person-days, accounting for 8.6% of the total tourist population. Among them, the total number of homecoming visitors from outside the Guangxi region is 536,239 person-days, representing 2.3% of the total tourist population, while the total number of homecoming visitors from within the region is 1,458,426 person-days, constituting 6.3% of the total tourist population.

Table 5. Comparison of the statistical data of homecoming visitors and other types of tourists identified from mobile signaling data across the Guangxi region (ten thousand person-days)

Statistical scope	homecoming visitors	Roaming resident users	Passing visitors	Dual Sim visitors
Outside Guangxi	53.62	61.91	56.06	14.62
Within Guangxi	145.84	343.38	101.54	24.61
Total	199.46	405.29	157.6	39.33

Table 6 presents a comparison of the statistical data of homecoming visitors and other types of tourists in the three cities of Nanning, Guilin, and Yulin, identified from mobile signaling data. It is noteworthy that the number of homecoming visitors in Yulin is particularly large, being 2.7 times that of the homecoming visitors in Nanning and Guilin, respectively. Yulin is an area where a significant number of people work outside their hometown, and there tends to be a substantial influx of returning people during holidays, which greatly impacts the tourist headcount statistics. This indicates that without distinguishing homecoming visitors, the statistical results may present a certain illusion and may not truly reflect the development status of a region's tourism industry.

Table 6. Comparison of the statistical data of homecoming visitors and other types of tourists in the three cities of Nanning, Guilin, and Yulin (ten thousand person-days)

Statistical scope	homecoming visitors	Roaming resident users	Passing visitors	Dual Sim visitors	Total visitors
Nanning	16.84	102.8	24	5.78	340.74
Guilin	16.3	51.34	18.8	3.24	244.26
Yulin	44.74	49.39	37.68	2.69	195.21

4. Summary

Homecoming visitors have certain particularities, and their consumption behaviors differ from those of general tourists. Special groups with unique regional and temporal characteristics should be analyzed and treated differently; otherwise, it can lead to anomalies in statistical data, failing to reflect the true state of the tourism industry. Moreover, this phenomenon also indicates that when conducting tourism statistics from the perspective of big data, the characteristics of big data—having a large sample size, being objective, and being multidimensional—should be fully utilized. Statistics should be conducted from multiple dimensions, perspectives, and with finer granularity to better serve economic development and enhance the level of social management.

In addressing the gap identified in the current research status, our methodology leverages mobile signaling data to devise a framework capable of distinguishing homecoming visitors from other tourist types. This study provides a simple and practical technical solution that can accurately determine whether mobile phone users belong to the homecoming visitors using big data analysis and machine learning methods, providing a basis for further statistics, research, and analysis of the behavioral characteristics and consumption habits of such tourists, and can be widely used in tourism statistics, tourism management, tourism product development, and precise marketing.

Acknowledgements

The work was supported by the Key R&D Program of Guangxi under Grant(AB21220023) and the Guangxi Key Laboratory of Culture and Tourism Intelligent Technology.

References

- [1] B. Jigang, W. Yajuan, T. Yonggang, M. Mingjian, W. Xianzhang, and C. Huiyan, "Research on Tourist Count Statistics from the Perspective of Big Data," *Tourism Forum*, vol. 13, no. 01, pp. 1-12, 2020, doi: 10.15962/j.cnki.tourismforum.202001001.
- [2] D. T. Duval, "When Hosts Become Guests: Return Visits and Diasporic Identities in a Commonwealth Eastern Caribbean Community," *Current Issues in Tourism*, vol. 6, no. 4, 2003, doi: 10.1080/13683500308667957.
- [3] B. McKercher, "Host Involvement in VFR Travel," *Annals of Tourism Research*, vol. 23, no. 3, pp. 701-703, 1996, doi: 10.1016/0160-7383(95)00100-x.
- [4] A. Seaton and C. Palmer, "Understanding VFR tourism behaviour: the first five years of the United Kingdom tourism survey," *Tourism Management*, vol. 18, no. 6.0, pp. 345-355, 1997, doi: 10.1016/s0261-5177(97)00033-2.

- [5] E. Backer, "VFR travel: It is underestimated," *Tourism Management*, vol. 33, no. 1, pp. 74-79, 2012, doi: 10.1016/j.tourman.2011.01.027.
- [6] L. Li, "A review of research on visiting relatives and friends abroad," *Tourism Tribune*, no. 02, pp. 90-96, 2008.
- [7] S. Yavuz, B. Kabadayı, N. Savaş, and A. C. Doker, "Tourism Demand of Turkey : A Gravity Modelling Approach," *Ekonomik ve Sosyal Araştırmalar Dergisi*, 2014.
- [8] Z. Xiaojun, "The Study on Regulatory Policy and Countermeasures about Tourist Flow in Guangzhou Baiyun Lake," Master, South China University of Technology, 2020.
- [9] V. V. Can, "Estimation of travel mode choice for domestic tourists to Nha Trang using the multinomial probit model," *Transportation research. Part A, Policy and practice*, vol. 49, pp. 149-159, 2013, doi: 10.1016/j.tra.2013.01.025.
- [10] T.-S. Lee, J.-N. Han, and H.-W. Kim, "A Study on the Consumption of Mobile Payment in Domestic Tourism Industry: Focused on Inbound Chinese Tourists," *Journal of MICE & Tourism Research*, vol. 21, no. 2, 2021, doi: 10.35176/jmtr.21.2.1.
- [11] C. J. León, C. Suárez-Rojas, J. M. Cazorla-Artiles, and M. M. G. Hernández, "Satisfaction and sustainability concerns in whale-watching tourism: A user-generated content model," *Tourism Management*, vol. 106, 2025, doi: 10.1016/j.tourman.2024.105019.
- [12] S. L. Rahjasa, G. Prasasti, and K. M. Apriliani, "How Travel Influencer and Social Media Influence Tourist Travel Decision to Ubud," *TRJ Tourism Research Journal*, vol. 8, no. 1, 2024, doi: 10.30647/trj.v8i1.243.
- [13] W. Ying, W. Peng, and C. Shiyu, "Passenger Characteristics and Transportation System Construction Analysis of Whole Regional Tourism Cities," *wf*, vol. 38, no. 5, 2022, doi: 10.3969/j.issn.1671-3400.2022.05.014.
- [14] C. Yanyan, Z. Ye, and S. Haodong, "Analysis of Tourist Flow Characteristics Based on Mobile Phone Signaling Data," *Journal of Beijing University of Technology*, vol. 48, no. 8, pp. 842-850, 2022, doi: 10.11936/bjutxb2021060001.
- [15] C. QIAN, "Analysis on Urban Visitor Activity Patterns Based on Mobile Phone Data," Doctor of Philosophy, College of Transportation Engineering, Tongji University, 2022.
- [16] R. Ahas, A. Aasa, Ü. Mark, T. Pae, and A. Kull, "Seasonal tourism spaces in Estonia: Case study with mobile positioning data," *Tourism Management*, vol. 28, no. 3, pp. 898-910, 2007, doi: 10.1016/j.tourman.2006.05.010.
- [17] K. Nilbe, R. Ahas, and S. Silm, "Evaluating the Travel Distances of Events Visitors and Regular Visitors Using Mobile Positioning Data: The Case of Estonia," *Journal of Urban Technology*, vol. 21.0, no. SP2, pp. 91-107, 2014, doi: 10.1080/10630732.2014.888218.
- [18] F Jia, X. Cheng, and Z. Duan, "Analyzing The Activity Areas Of Non-Resident Tourists Of Shanghai Expo Using Cellular Phone Data," *Procedia - Social and Behavioral Sciences*, vol. 96, 2013, doi: 10.1016/j.sbspro.2013.08.130.
- [19] Y. Wu, L. Wang, L. Fan, M. Yang, Y. Zhang, and Y. Feng, "Comparison of the spatiotemporal mobility patterns among typical subgroups of the actual population with mobile phone data: A case study of Beijing," *Cities*, vol. 100, 2020, doi: 10.1016/j.cities.2020.102670.
- [20] Z. Shihui and S. Hang, "A Study on the Types of Tourists in Religious Scenic Spots Based on Mobile Phone Signaling Data -- A Case Study of Mount Wutai Scenic Spot," *Tourism Overview*, no. 12, pp. 132-136, 2023.
- [21] J. Jiang, M. Chen, and J. Zhang, "Analyses Of Elderly Visitors' Behaviors To Community Parks In Shanghai And The Impact Factors," *Landscape Architecture Frontiers*, vol. 8, no. 5, pp. 94-109, 2020, doi: 10.15302/j-laf-1-020038.
- [22] D. Yang, J.-g. Han, P. Wu, and X. Zhao, "A Method of Tourist Identification and Travel Trajectory Matching Based on Mobile Signaling Data," *Transport Research*, vol. 5, no. 6, pp. 20-27, 2019, doi: 10.16503/j.cnki.2095-9931.2019.06.003.