Heliyon 10 (2024) e33260

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

The border tourism hotspots network based on travelogues

Siyue Zhang [a,c], Zhaoping Yang [a,b], Cuirong Wang [a,*]

[a] State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, 830011, China
[b] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China
[c] University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Keywords:
Border tourism
Association rule learning
Network analysis
Travelogues
Qinghai-tibet plateau

ABSTRACT

This study applies the Apriori algorithm and social network analysis to analyze travelogue data from the Qinghai-Tibet Plateau, effectively mapping significant border tourism hotspots and their interconnections within defined networks. Our findings distinctly partition the region into two principal sub-networks: Xinjiang and Tibet, highlighting the geographical segmentation that potentially impedes comprehensive regional tourism integration. The analysis underscores a pronounced reliance on transportation, reflecting the inherently multi-destination nature of border tourism in the area. Key nodes within these networks include Kashgar City and Lhasa City, serving as central hubs in their respective sub-networks, while Gar County and Hotan County act as pivotal connectors bridging the two distinct areas. Additionally, Gyirong County, Nyalam County, and Tashkurgan Tajik Autonomous County are identified as well-developed border tourism destinations, with Gyirong County and Tashkurgan Tajik Autonomous County possessing substantial potential to evolve into core border tourism hubs. Drawing on these insights, the study proposes targeted development strategies to enhance the structure and efficacy of border tourism on the Qinghai-Tibet Plateau.

1. Introduction

The natural environment and social development of the Qinghai-Tibet Plateau are both unique. It has a total border length of nearly 6000 km, connecting the provinces of Xinjiang, Tibet, Qinghai, Sichuan, Yunnan in China, and neighboring countries including Kyrgyzstan, Tajikistan, Pakistan, India, Nepal, Bhutan, Afghanistan, Myanmar, and Mongolia. It serves as a national security barrier for China [1]. In the context of globalization, border regions have become hotspots for national cooperation, communication, and openness. Border tourism has become an effective means to promote international cooperation, increase income for border residents, and stimulate economic development in these areas [2–5]. Developing border tourism in the Qinghai-Tibet Plateau can optimize the industrial structure of border regions, ensure the stability of the border population through tourism, enrich the variety of regional tourism products, and contribute to the construction of the “Qinghai-Tibet Plateau World Tourism Destination”. It is an important measure to promote social and economic development in border regions and ensure the stability and long-term prosperity of ethnic minority areas along the border. However, the research on China’s border tourism has mainly focused on national-level border tourism destinations or specific regional studies on border tourism spatial structures, tourist flows, development strategies, and regional systems in Northeast China, Southwest China, Xinjiang, and other regions. There has been no specific research on the Qinghai-Tibet

* Corresponding author.
E-mail address: wangcr@ms.xjb.ac.cn (C. Wang).

https://doi.org/10.1016/j.heliyon.2024.e33260
Received 26 January 2024; Received in revised form 12 June 2024; Accepted 18 June 2024
Available online 19 June 2024
2405-8440/

Plateau border region [6–13].

The border region of the Qinghai-Tibet Plateau is located on the periphery of China's economic development space, with relatively backward social and economic development, and situated at the end of domestic transportation routes. The infrastructure, service levels, tourism market size, and tourist consumption levels are all at a relatively low level. Compared to all other border regions in China, except for the Tashkurgan Tajik Autonomous County, the tourism competitiveness of other border counties in the Qinghai-Tibet Plateau is weak. The comprehensive development index for tourism in all border regions is low [8,11,14]. The characteristics of being far from central cities and having poor transportation accessibility make tourism in the border regions of the Qinghai-Tibet Plateau "multi-destination" in nature. In other words, the border regions are one of the destinations for tourists in their travel routes within the Qinghai-Tibet Plateau, and tourists have to pass through multiple regions to reach the border regions. At the same time, the multi-destination model can attract more tourists and extend their length of stay. Creating tourist clusters (or connections) can accelerate the enhancement of the destination's competitive advantages. Creating scenic clusters (or destinations) that combine primary and secondary attractions can enhance the overall attractiveness of the region [15]. So, what are the tourism destinations closely related to the border regions in the Qinghai-Tibet Plateau? What are the connections between these nodes? What are the characteristics of the border tourism spatial layout structure composed of the border regions and their associated destinations?

To address the aforementioned issue, this study focuses on the border regions of the Qinghai-Tibet Plateau as the research area. It is important to note that data on border tourism in the Qinghai-Tibet Plateau is scarce, making it challenging to directly obtain statistics on tourist numbers and tourism revenue. However, geospatial big data provides a rich source of information for analyzing patterns and behaviors of tourists. The utilization of these extensive datasets enables a comprehensive analysis of the spatial distribution and temporal variations in tourist activities [16]. And online travel narratives and tourism blogs are a form of geospatial big data; they reflect tourists' experiences at specific destinations. These textual sources are vital for comprehending consumer needs and expectations, improving and monitoring destination images, and adjusting competitive strategies [17]. These texts record visitors' travel routes, places of stay, and activities, containing a large amount of geotagged geographic data that more accurately reflects the behavioral characteristics of tourists than traditional statistical data [15,18]. By analyzing geographic data and extracting spatial interaction information of tourist destinations using machine learning algorithms, the spatial network structure, distribution patterns, and spatiotemporal evolution of destination networks can be analyzed, which provided new insights into the study of interactions between tourism destinations [19]. For example, by utilizing geotagged photos from Flickr, a dataset comprising photos taken by tourists entering Hong Kong was compiled. Density clustering and Markov chain techniques were applied to mine patterns of tourist behavior, identifying points of interest, travel routes, and times of activities, thus providing practical recommendations for tourism managers in Hong Kong [20]. By analyzing the visual content of tourist photos at Beijing travel destinations, the behaviors and perceptions of tourists were explored [21]. Based on the geographic information of the photos, tourist cognitive maps of different perceived themes were visualized using ArcGIS. Additionally, digital trajectories of urban bicycle travel were used to identify the main tourist and leisure hotspots in Krakow [22]. A spatial interaction network of tourists' activities in Huangshan City was constructed using geotagged data extracted from Weibo, a popular social media platform, and urban tourist areas were divided using community detection algorithms [23]. The strength of contacts between different cities was explored through regional segmentation, using data from social media platforms to construct a complex network [24,25].

Network analysis is a research method that focuses on the study of "actors and their relationships" to determine the positions of actors and the characteristics of their relationships within the network. In the study of the spatial structure in tourism, destinations can be regarded as nodes in a regional tourism spatial network, and the network is constructed based on the relationships between nodes. This research aims to uncover the characteristics of the network structure, classify nodes, and examine tourism connections between them by processing and analyzing network data [26,27].

This study addresses the challenges of researching and developing border tourism on the Qinghai-Tibet Plateau by utilizing a methodological framework called "node selection-network construction-network analysis" with network travelogue texts as the primary data source. It employs machine learning techniques to identify tourist "hotspots" in the region and assesses the relational dynamics among these nodes to construct a network of hotspot connections. The network analysis provides insights into the spatial structure and characteristics of the nodes. Based on these findings, the study offers strategic recommendations for border tourism development in the Qinghai-Tibet Plateau.

2. Literature review

2.1. Border tourism

Border tourism can be understood as the collection of all elements of tourism behavior and activities carried out in the "border area" or "borderline", which are related to or exist because of the "border" [28]. It is characterized by its high sensitivity to geopolitical relations and the inherent power barrier of national boundaries, which are referred to as its "political characteristics". Concurrently, it exhibits spatial characteristics of "marginality" and "intermittency". These two distinct features set border tourism apart from both domestic and international tourism, establishing it as a unique category [29].

"Borders" refer to the political boundaries between nations, serving three primary missions: bordering, ordering, and othering [30,31]. These boundaries restrict the flow of materials and people across neighboring territories, leading to both political and cultural isolation. Consequently, border regions exhibit greater "resource heterogeneity" than internal areas and more pronounced "policy heterogeneity" compared to adjacent countries, which are key drivers for border tourism. For national security, borders serve as a protective barrier; for the tourism market, they form an integral part of destination marketing; and for border communities, borders represent socio-economic assets, such as being a valuable resource for the tourism industry [32]. The development of border tourism is primarily influenced by international political relations and national policies due to the political nature of borders. The definition of border tourism varies depending on the specific policies in place, although there is common agreement on two key points: that border tourism involves departing a country via land travel and that the border area serves as the primary geographic locus for border tourism activities.

The primary purpose of border tourism is the "cross-border experience", wherein the ease of clearance and cultural differences play crucial roles in its development. Border tourism destinations can be classified into four groups based on the ease of crossing the border and the socio-cultural similarities between the regions. The first group comprises destinations with high border crossing difficulty and substantial cultural differences, such as Israel and its Arab neighbors. The second group consists of destinations with easy border crossing but significant cultural disparities, exemplified by Greece and Turkey. The third group pertains to areas with challenging border crossing but limited cultural variation, like Mongolia and China. Lastly, the fourth group includes regions with low cultural disparity and effortless border crossing, such as Austria and Germany [33].

In the field of tourism, "border tourism" refers not only to "cross-border tourism" involving travel across national boundaries, but also to domestic tourist activities that occur within border regions without necessarily crossing borders. Within this context, the economic development level of the border area, the attractiveness of its tourism resources, and the quality of its tourism infrastructure are critical factors that influence the growth of border tourism [3,33,34].

There are two primary categories of border tourist attractions. The first category includes the border as a tourist attraction, encompassing boundary markers, customs buildings, flags, and other border-associated symbols. Additionally, it involves the crossing experience across varied political, social, and economic domains. The second category includes attractions situated in the border area, such as natural wonders, world heritage sites, and cultural landscapes. These attractions, while not directly linked to the border, still hold tourism appeal [33,35]. The differences in social development levels and policy disparities due to political isolation lead to cross-border movements for purposes like medical treatment and shopping. Despite not primarily intending to engage in tourism, strategic marketing, promotion, infrastructure, and service provision enable their temporary transformation into border tourists [36].

The term "border tourism" can be more precisely defined as "touristic activities occurring in border regions," encompassing a spatial concept that includes not only the border itself but also the adjacent areas [37]. The assessment of potential tourism development in border areas and the formulation of corresponding strategies can be accomplished using indicators including the ease of border crossings, the resource endowment of the border area, service levels, and transportation accessibility. For example, the evaluation of tourism destination competitiveness encompasses factors such as tourism appeal, management systems, community composition, and the assessment of potential collaborative tourism destinations between Slovenia and Croatia. This assessment aims to investigate the potential for enhancing the competitiveness of tourism in the border regions of both countries through a strategic approach to planning and managing cross-border tourism destinations [38].

2.2. *Social network analysis and tourism destinations*

Network analysis serves as a research methodology for examining the relationships among the nodes within sophisticated systems. Tourism destinations are intricate systems comprising tourism resources, services, public facilities, and socio-cultural activities. Interactions arise when a destination's action prompts responses from others, forming a Network of Tourism Destinations (NTD), a geographical system of interconnected nodes and connections [39]. From a systemic perspective, tourism can be examined as a series of interconnected elements, with a focus on the relationships that depict interactions among diverse components of tourism activities, which are pivotal in tourism research. These relationships significantly influence the planning and management of tourism destinations. Consequently, the use of social network analysis is pivotal in examining diverse relationships within the tourism sector. This method can reveal the mathematical properties inherent in tourism relationships and shed light on the interconnectedness among products, destinations, tourist behaviors, and stakeholders in the tourism industry [40].

Research at the micro-level has specifically focused on the interrelationships among stakeholders within tourism destinations. Scott defined tourism as a complex network of organizations that interact to provide services. They conducted a study using the Gold Coast, Southern Downs, Wine and High Country, and Great Ocean Road in Australia as case studies to examine the relationships and structural dynamics of stakeholders within these tourism destinations. The findings suggest that network analysis provides a valuable tool for comprehending the structure and cohesion of tourism destinations. This method exposes variations in cooperative structures and inter-organizational cohesion across different types of destinations, while also pinpointing structural deficiencies. These insights can inform policy development and management practices [41]. Baggio examined both the static structural features of the tourism stakeholder network in Elba, Italy, and the dynamics pertaining to the dissemination of destination information. The findings established that stakeholder cohesion and adaptability served to facilitate the process of destination information dissemination [42].

When analysis expands beyond individual destinations, it reveals a wealth of intriguing and unique attributes. Tourist destinations can be envisioned as networks, with nodes symbolizing tourist attractions and edges indicating the flow of tourists between these attractions. By analyzing these network structures, it is possible to identify clusters of attractions that tourists frequently visit together. This insight facilitates the development of destinations tailored to the specific needs of tourists [43]. The examination of destination-specific features at the mesoscale encompasses multiple destinations within a given region. The findings from this analysis can be leveraged to improve destination management. It is recommended that governments and tourism organizations allocate appropriate tourism facilities to destinations based on the network characteristics of the tourism routes, thereby providing a comprehensive tourism experience [26]. Lee evaluated 43 Korean villages for centrality using GIS and network analysis. Then categorized the villages based on their centrality to formulate an integrated tourism strategy, considering the available local resources in each village [44]. After examining a destination network matrix derived from travel packages offered by different travel agencies, Hong utilized network analysis to investigate the characteristics of the destination network and the correlation between nodes and tourist flows. The findings suggest that establishing a hierarchical destination network can reduce competitive intensity among destinations in the tourism market, enhancing the competitive advantage of individual destinations within the network. Furthermore, the study illustrates a close relationship between node connections among destinations and tourist flows [39]. Wu constructed a spatially embedded network of popular tourist destinations in Beijing by examining the travel paths of tourists. This involved analyzing the structural traits of the network and the prominence of specific attractions, leading to the identification of pivotal nodes within the network. The study's results provide crucial decision-making support for the development of tourism resources, facilitate the prediction of tourist routes, and aid in the planning of tourist bus routes [45].

Tourists, being the most dynamic and critical actors within the destination system, connected various tourist attractions through their discretionary consumption behavior and mobility. Consequently, the movement and mobility of tourists constitutes one of the most significant factors in shaping the network of destination attractions [46]. Additionally, tourists serve as integral stakeholders in the destination system; destination studies that embrace a tourist perspective can aid in developing a more comprehensive and nuanced approach to managing destinations [47]. Data derived from tourist-generated travel streams, web content, and other sources offer new insights for analyzing tourist destination networks.

3. Materials and methods

3.1. Study area

Zhang proposed that the geographic range of the Qinghai-Tibet Plateau extends starting from the northern edge of the Pamir Plateau, and the northern foothills of the Western Kunlun and Qilian Mountains in the north, extending to the southern foothills of the Himalayas and other southern mountain ranges, and stretching from the Hindu Kush Mountains and the western edge of the Pamir Plateau in the west to the eastern extent of the Qilian Mountains and the eastern edge of the Hengduan Mountains in the east [48]. Considering the integrity of administrative divisions, the Qinghai-Tibet Plateau includes six provinces: Qinghai, Yunnan, Gansu, Xinjiang, and Sichuan, with a total of 216 county-level administrative units (Fig. 1). According to the China's county-level administrative unit classification standard, these 216 county-level units can be categorized into four types: city districts, county-level cities,

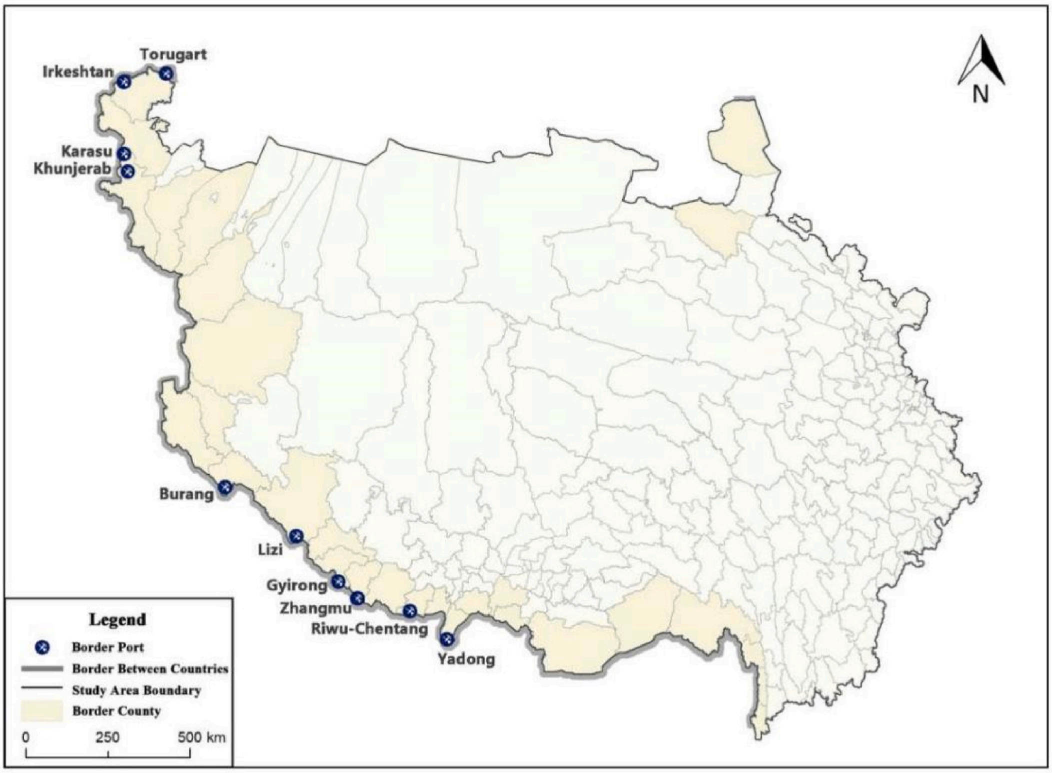

Fig. 1. Overview map of the study area (source: created by the author)
The figure represents 216 administrative units at the county level within the research area (Qinghai-Tibet Plateau), with a special emphasis on the border counties and land ports within the region.

counties, and autonomous counties. Based on the behavioral characteristics of tourists and the narrative conventions found in travelogue texts, some city districts have been consolidated into urban areas. The consolidation involves merging the "Durlongdeqing District" "Dazi District" "Chengguan District" into "Lhasa," as well as merging the "Huangzhong District" "Chengbei District" "Chengxi District" "Chengzhong District" and "Chengdong District" into "Xining". Moreover, the "Bayi District" is designated as "Nyingchi" Therefore, this study includes 210 counties. The term "county" is divided into four types: [1] for urban areas, [2] for county-level cities, [3] for counties, and [4] for autonomous counties [49]. Within the framework of this research, the analysis encompasses a total of 210 counties, of which 28 are border counties. These counties, due to their unique geographic locations, boast rich natural landscapes and a deep cultural heritage. In the border regions of Tibet, one can find world-renowned natural sites such as Mount Everest and the Yarlung Tsangpo Grand Canyon. The border counties in Xinjiang, situated on the Pamir Plateau, once served as crucial nodes on the ancient Silk Road. As such, this area is not only home to numerous peaks exceeding 8000 m but also preserves a wealth of cultural heritage. In Yunnan Province, the border counties form the core area of the Three Parallel Rivers World Natural Heritage site and serve as cultural hubs for ethnic minorities such as the Lisu, Derung, and Nu peoples. Gansu's border counties are known for their expansive grasslands and Mongolian cultural attributes. The border counties include ten land ports: Zhangmu, Gyirong, Burang, Lizi, Yadong, Riwu-Chentang, Torugart, Irkeshtam, Khunjerab, and Karasu (Table 1). A land port is defined as a facility equipped with border control, customs, maritime supervision, health quarantine, animal and plant quarantine, and commodity inspection agencies, along with other institutions empowered by the state [50].

In China, border tourism refers to organized tourism activities facilitated by approved travel agencies for citizens of China and its neighboring countries. The participants travel in groups, entering and exiting the country through specified border land ports, and engage in exploration within the agreed border area and duration, as stipulated by the respective governments [51]. The national government has introduced the concept of border tourism pilot zones and cross-border tourism cooperation zones to enhance the development of border tourism. These zones are established based on existing border land ports officially approved by the State Council for external access, and are authorized to manage the entry and exit of individuals. The People's Government of the border counties is designated as the primary authority responsible for the declaration [52]. Ports play a crucial role as the primary entry points for border tourism in China, with port counties emerging as significant destinations for such tourism [53]. Consequently, the objective of this research is to pinpoint county-level administrative units that encompass ports and function as destinations for border tourism. These include Nyalam[3], Gyirong[3], Burang[3], Zhongba[3], Dinggye[3], Yadong[3], Wuqia[3], and Tashkurgan[4].

3.2. Data source

3.2.1. Data collection and preprocessing

The travelogue text data used in this study primarily originates from two leading online travel service platforms in China: Ctrip (https://dst.ctrip.com) and Mafengwo (https://www.mafengwo.cn). As of 2020, these platforms have registered user counts of 300 million and 130 million, respectively [54,55]. Their substantial user base provides a rich dataset for this research, ensuring data breadth and diversity. Ctrip, as China's largest online travel agency (OTA), offers services such as hotel booking, transport ticketing, vacation packages, and a platform for sharing travel guides; Mafengwo is characterized by its user-generated content (UGC), fostering a community for travel sharing.

The active users of these platforms are predominantly between the ages of 30 and 39, accounting for more than 60 % of users, and are mostly from economically developed regions in China [56]. This demographic profile suggests that the data primarily reflects the travel behaviors and perceptions of middle-aged, economically capable urban residents. These travelers tend to prefer independent travel, and their travelogues not only document specific activities and experiences but also their deep impressions and evaluations of the destinations.

Using Octoparse web scraping software, travelogues were downloaded from the destination search pages of both websites by entering the names of eight border tourism destinations. For each travelogue, information was collected including travel duration, number of companions, expenses, travelogue text, and travel dates. A total of 3645 records were obtained; after excluding travelogues that lacked text, contained advertisements, irrelevant content, or were duplicates, 1296 valid records were compiled.

Table 1
Qinghai-Tibet plateau ports information.

Land crossing port	Administrative location	Crossing state	Neighboring countries
Zhangmu Port	Nyalam[3], Rikaze City, Tibet Autonomous Region	opening	Nepal
Gyirong Port	Gyirong[3], Rikaze City, Tibet Autonomous Region	opening	Nepal
Burang Prot	Burang[3], Ali Region, Tibet Autonomous Region	opening	Nepal
Lizi Port	Zhongba[3], Rikaze City, Tibet Autonomous Region	opening	Nepal
Riwu-Chentang Port	Dinggye[3], Rikaze City, Tibet Autonomous Region	to be opened	Nepal
Yadong Port	Yadong[3], Shigatse Prefecture, Tibet Autonomous Region	to be opened	India
Torugart Port	Wuqia[3], Kizilsu Kirgiz Autonomous Prefecture, Xinjiang Uygur Autonomous Region	opening	Kyrgyzstan
Irkeshtam Port	Wuqia[3], Kizilsu Kirgiz Autonomous Prefecture, Xinjiang Uygur Autonomous Region	opening	Kyrgyzstan
Khunjerab Port	Tashkurgan[4], Kashgar Region, Xinjiang Uygur Autonomous Region	opening	Pakistan
Karasu Port	Tashkurgan[4], Kashgar Region, Xinjiang Uygur Autonomous Region	opening	Tajikistan

Note: This table was created by the author.

3.2.2. Constructing itemsets from travelogues

In this study, content analysis was applied to extract keywords from the travelogues using ROST CM6 software, which produced a lexical frequency table that identified high-frequency administrative units. Notably, although "Kashgar[1]" is not located within the Qinghai-Tibetan Plateau study area, it emerged as the second most frequently mentioned location, indicating a strong association with border tourism on the Qinghai-Tibet Plateau. Together with the area's 210 counties, 10 land ports, and 9 neighboring countries, Kashgar[1] forms a critical dataset of key locations. A total of 230 key location names were extracted from the travelogues, with each travelogue contributing to an itemset comprising locations such as $\{Site_1, Site_2, \ldots, Site_n\}$. Data consisting solely of a single keyword was excluded. This process yielded a total of 875 itemsets, which formed the transaction database for Apriori analysis.

3.3. Research methods

3.3.1. Association rules mining

The Apriori algorithm is a widely applied method in machine learning for mining association rules [57], primarily utilized to reveal frequent co-occurrences among itemsets. This algorithm has found extensive application in the retail sector, aiding in key business operations such as product pricing, market promotion, and inventory management. Specifically, by identifying products frequently purchased together, businesses can better understand consumer behavior patterns and optimize their strategic approaches [58]. In recent years, scholars have begun to explore the integration of the Apriori algorithm with geospatial big data analytics to understand tourist behaviors and optimize destination management. For instance, a study utilized Wi-Fi tracking data to mine association rules between different tourist destinations, evaluating their attractiveness based on visitation frequency. This approach provides data-driven support for developing sustainable tourism management policies [59]. Moreover, by analyzing extensive tourist data, the Apriori algorithm helps identify potential knowledge and patterns within tourism product development and customer relationship management [60]. In a study focusing on inbound tourists to China, the algorithm was used to identify co-occurring preference attributes at scenic spots, and based on these attributes, market segmentation was conducted, offering profound insights into the tourism market [61]. And the Apriori algorithm demonstrates significant potential in uncovering relationships between tourism destinations and holds value in understanding and promoting the development of rural tourism, among other areas [62]. This method not only assists researchers in identifying key driving factors but also provides scientific decision-making support for the formulation and implementation of tourism policies.

By the Apriori algorithm, the data mining process primarily consists of two stages: discovery of frequent itemsets and extraction of association rules. Frequent itemsets refer to combinations of data items that commonly co-occur within the dataset. These are defined as the proportion of records in the dataset that contain a specific itemset, calculated as shown in (1) [63]. Association rules, on the other hand, reveal potential connections between different data items, measured in strength by their confidence, as calculated in (2) where $P \cup H$ represents elements occurring in both sets P and H. A higher confidence level indicates a stronger association between two itemsets [58].

The execution of the Apriori algorithm begins by generating a list of itemsets consisting of single elements, selecting those that meet a predefined minimum support threshold. Subsequently, the algorithm iteratively generates itemsets containing an increasing number of elements, again filtering based on the minimum support, and this process continues recursively until no new frequent itemsets can be generated. Based on these frequent itemsets, the algorithm then generates a list of potential association rules, filtering out strong association rules that meet the minimum confidence threshold by calculating each rule's confidence. Through this sequence of steps, the Apriori algorithm effectively extracts valuable patterns and associations from large datasets, providing insights for in-depth analysis and decision-making [58].

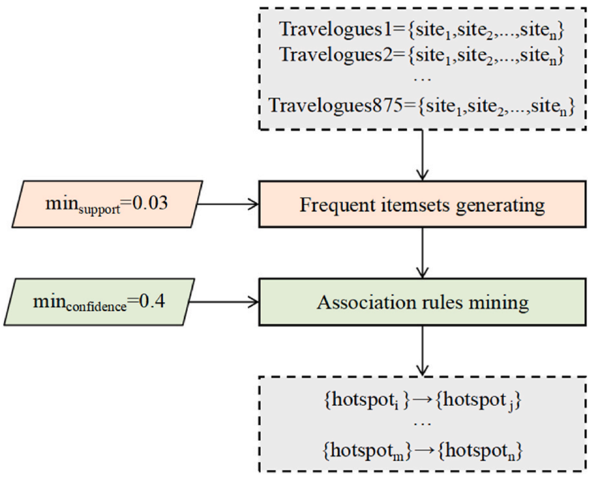

Fig. 2. The approach of association rule mining (source: created by the author)
This figure illustrates the workflow of the Apriori algorithm used in this study.

$$\text{Support(A)} = \frac{\text{Number of occurrences of itemset A}}{\text{Total number of records in the dataset}} \tag{1}$$

$$\text{Confidence}(P => H) = \frac{\text{Support}(P \cup H)}{\text{Support}(P)} \tag{2}$$

In this study, the Apriori algorithm was employed to conduct association rule learning on a collection of 875 travelogue datasets, with minimum support and confidence thresholds set at 0.03 and 0.4, respectively. The workflow is illustrated in Fig. 2.

3.3.2. Social network analysis

In this study, the Apriori algorithm was utilized to extract association rules from tourism textual data. These rules were subsequently converted into nodes and relationships for social network analysis. Specifically, the data derived from the association rule mining were binarized to indicate the presence or absence of connections between nodes. This transformation facilitated the construction of a tourism hotspots network, represented by a 35 × 35 association matrix. This matrix delineates the key nodes associated with border tourism in the Qinghai-Tibet Plateau, providing a structured framework for comprehensive social network analysis of border tourism.

With Ucinet 6, a software for social network analysis, to scrutinize the network structure and the extent of association among tourism nodes situated on the border of the Qinghai-Tibetan Plateau. The network's overall characteristics were evaluated using network density and cohesive subgroups. Additionally, the structural properties of nodes were assessed by nodal centrality [64,65].

(1) Network density

Density is considered the fundamental measure at the network level, reflecting the degree to which nodes in a network are interconnected. It is calculated by dividing the total number of all two-party relationships present in the network by the maximum possible number of two-party relationships in the network. The formula for density is expressed as(3). *X(i,j)* is the number of links between node *i* and other nodes; *N* is the number of network nodes [64].

$$D = \frac{\sum\sum X_{i,j}}{N(N\text{-}1)}(i \neq j) \tag{3}$$

(2) Cohesive subgroups

The aim of cohesive subgroup analysis is to identify "structural peer groups" within social networks, referring to sets of nodes exhibiting similar or substitutable connections with other nodes in the network. In this study, we utilize the CONCOR algorithm, which is based on the node correlation matrix, for conducting cohesive subgroup analysis. Notably, the algorithm de-emphasizes the attributes of individual nodes, instead highlighting the correlation relationships between nodes, allowing for the classification of individual nodes in the correlation network into clusters with similar structural positions [64].

(4) Nodal centrality

Degree centrality characterizes the number of connections a node has within a network; higher degree centrality signifies a greater number of associated nodes, reflecting the node's significance within the network. $X_{i,j}$ is the count of connections between node *i* and other nodes, *g* is the total number of nodes in the network. The formula is expressed as(4) [66]:

$$C_D(N_i) = \sum_{j=1}^{g} X_{i,j}(i \neq j) \tag{4}$$

The concept of betweenness centrality involves the ability to connect nodes that do not have a direct connection. A node with low degree centrality, but high mediator centrality, acts as a bridge between two separate networks and can potentially act as the sole conduit between them, thereby exerting significant influence. $g_{jk}(N_i)$ is the number of all paths between points *j* and *k* that pass through point *i*; g_{jk} is the number of paths between *j* and *k*. If the betweenness centrality is 0, the value signifies that none of the pathways between pairs of points in the network traverse through node *i*. The formula is expressed as(5) [67]:

$$C_B(N_i) = \sum_{j<k} \frac{g_{jk}(N_i)}{g_{jk}} \tag{5}$$

Closeness centrality represents the inverse of the sum of the shortest path lengths between node i and other nodes. It is used to gauge the distance between node i and other nodes in the network, thereby reflecting the challenge encountered by node i in reaching these other nodes. A smaller sum of the shortest paths for node i and other nodes corresponds to a larger proximity centrality for node i. There exists a negative correlation between closeness centrality and the geodesic distance of node i from other nodes. Higher closeness centrality indicates a shorter path between the node and other nodes, suggesting reduced difficulty in reaching other nodes and a closer connection with them. The formula is expressed as(6). $d(N_i, N_j)$ is the shortest path between points *i* and *j*; *g* is the total number of nodes in the network [68].

$$C_C(N_i) = \frac{1}{\sum_{j=1}^{g} d(N_i, N_j)} (i \neq j) \tag{6}$$

4. Results

4.1. Border tourism hotspots associated network

The frequency of co-occurrence of toponyms can be used to measure the intensity of spatial interaction [69,70]. The confidence determined by the Apriori algorithm indicates the frequency of co-occurrence of the two keywords in the itemset, which reflects the correlation between hotspots related to border tourism. As a result, 85 pairs of association relationships were discovered among 35 nodes.

4.1.1. Nodes identification

The 35 nodes identified can be categorized into three types: counties, ports, and countries, as shown in Fig. 3. These nodes represent hotspots in border tourism narratives, frequently mentioned and significantly associated with other nodes. They encompass 4 countries, 3 land ports, and 28 counties. Only Tajikistan, Pakistan, India, and Nepal are represented in the network, accounting for 44 % of the neighboring countries of the Qinghai-Tibetan Plateau. Bhutan, Afghanistan, Myanmar, and Mongolia are absent from the network, potentially due to the absence of ports. Although Kyrgyzstan features two major land crossing ports, Irkeshtam Port and Torugart Port, which rank among the highest in annual freight and passenger flows on the Qinghai-Tibetan Plateau [71], it is absent from the linkage network. This absence may be attributed to the slow development of Kyrgyzstan's tourism industry and a lack of core natural and cultural attractions. The three identified land ports, Khunjerab Port, Gyirong Port, and Zhangmu Port, account for 30 % of the total number of ports, representing a relatively low overall percentage.

The 28 county nodes are distributed across four provinces: Xinjiang, Tibet, Yunnan, and Qinghai, except for Xining[1] in Qinghai Province and Shangri-La[2] in Yunnan Province, which are located outside Xinjiang and Tibet. Despite their distance from port counties, Xining[1] and Shangri-La[2] remain hotspots for border tourism, possibly because these nodes serve as critical hubs in the tourism transportation network of the Qinghai-Tibet Plateau. When overlaying these 28 counties with the tourism corridors" of the Qinghai-Tibet Plateau [72], it becomes evident that many nodes are strategically positioned along these corridors, significantly influenced by

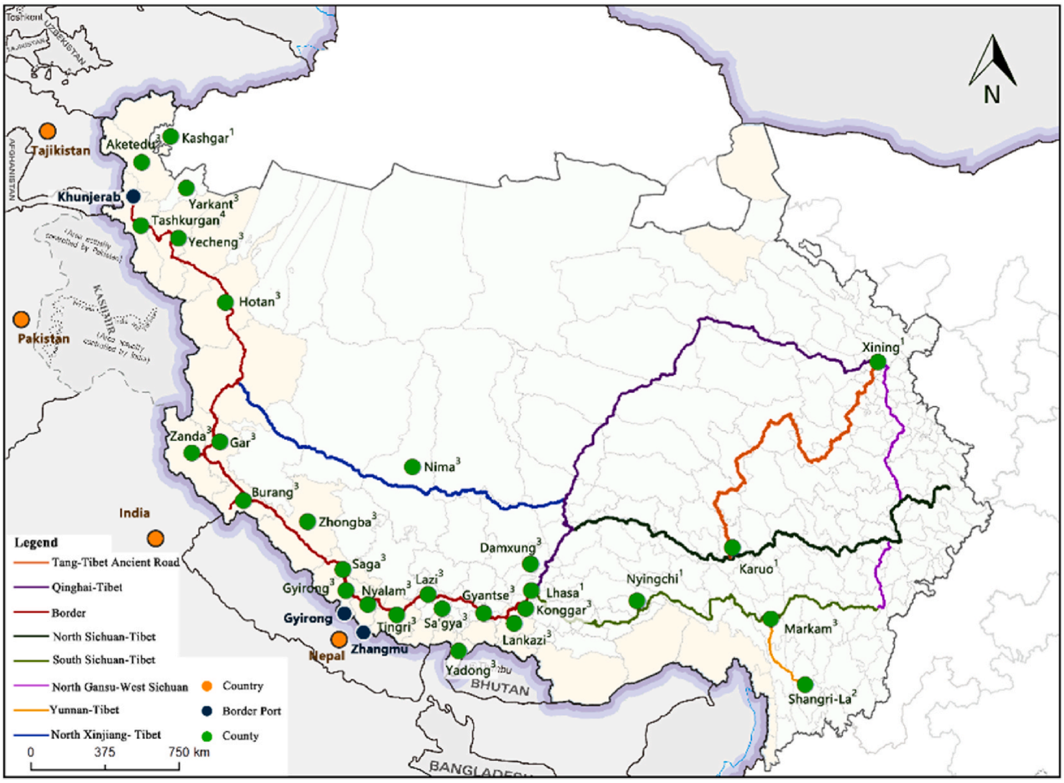

Fig. 3. Spatial distribution of border tourism network nodes (source: created by the author)
The identified 35 nodes are visually displayed using different legends to distinguish the types of nodes, and overlaid with the main "tourism transport corridors" of the Qinghai-Tibet Plateau.

its effects. Notably, 21 county nodes are situated along the border line, which relies heavily on China's longest border highway—the G219 National Road. This alignment substantiates the correlation between the identified nodes in our analysis and the actual geographical and infrastructural dynamics observed in the region. This validation underscores the reliability of our methodological approach in accurately reflecting the existing infrastructure and geographical realities.

4.1.2. Association rules mining

The degree of association between nodes was defined based on the confidence of their relationships, and a spatial visualization was subsequently performed (Fig. 4). The border tourism hotspots network of the Qinghai-Tibet Plateau is comprised of two distinct sub-networks: Xinjiang and Tibet. The Xinjiang sub-network includes ten nodes: Kashgar[1], Tashkurgan[4], Yarkant[3], Yecheng[3], Hotan[3], Aketedu[3], Gar[3], Tajikistan, Pakistan, and Khunjerab Port. Except for Gar[3], Kashgar[1] demonstrates a high level of connectivity with the remaining eight nodes. Gar[3], the only county node located in Tibet, shows a limited level of association with Kashgar[1], likely due to the presence of Shiquanhe town in Gar[3], which serves as a critical supply node at the intersection of the Xinjiang-Tibet Highway, the Anshi Highway, the "Hei-Ah" Highway, and the Azhi Highway. Additionally, Shiquanhe plays a vital role in connecting the provinces of Xinjiang and Tibet through its involvement in the Xinjiang-Tibet Border Tourism Route. Tajikistan and Pakistan display a strong correlation, suggesting a high likelihood of both countries being mentioned together in travelogues.

The Tibet sub-network consists of 21 counties, 2 ports (Gyirong Port and Zhangmu Port), and 2 countries (Nepal and India). Gyirong Port and Zhangmu Port have strong associations with Nepal, consistent with their roles as the main terrestrial border entry points for Nepal. Lhasa[1] is strongly associated with Lankazi[3], Saga[3], Damxung[3], Karuo[1], Lazi[3], Gyantse[3], Konggar[3], Markam[3], and Nyingchi[1], indicating that if a traveler is interested in any of the nine counties, there is more than a 74 % probability they will also be interested in Lhasa[1]. There are 18 pairs of association rules related to border tourism destinations (Table 2). Excluding Yadong[3], the other four destinations all have a moderate to strong association with Lhasa[1], suggesting that most tourists visiting border destinations pass through Lhasa[1]. Nyalam[3], Gyirong[3], and Zhongba[3] have associations with Nepal, while Burang[3] and Yadong[3] are associated with India. This suggests that over 50 % of tourists associate these border tourism destinations with the adjacent countries.

4.2. Network analysis

4.2.1. Network density

The connectivity strength of nodes within the tourism network of the Qinghai-Tibetan Plateau and its sub-networks was assessed by

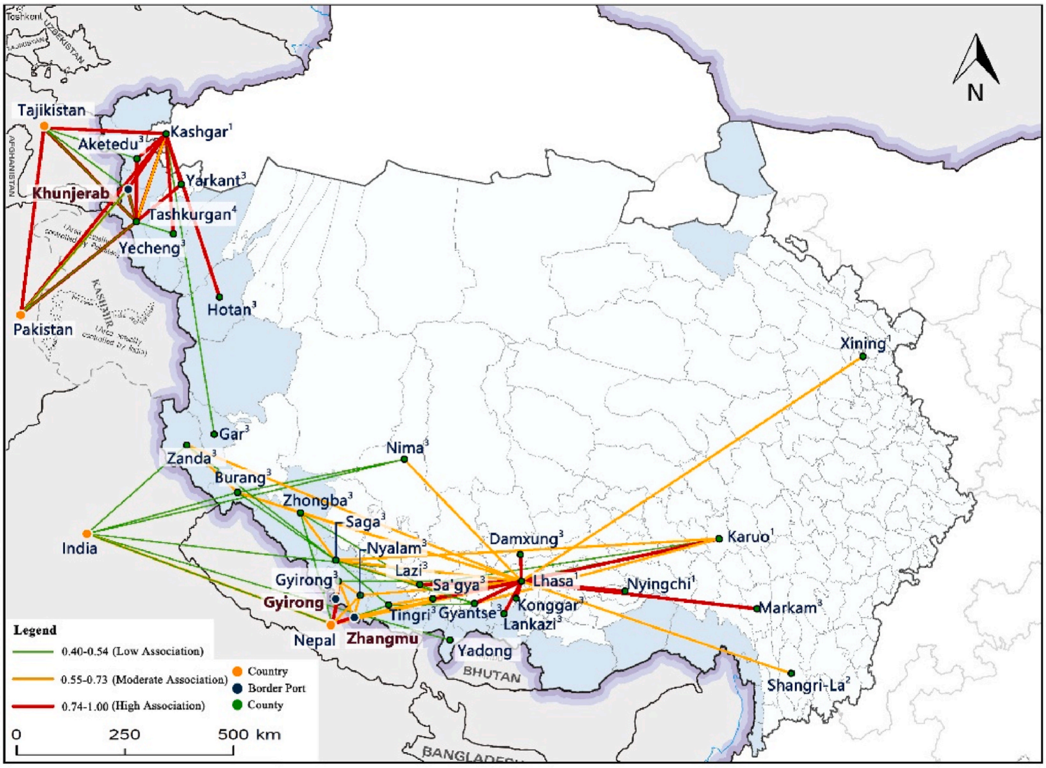

Fig. 4. Border tourism hotspots network (source: created by the author)
To visualize the 85 pairs of association relationships among 35 nodes identified through the Apriori algorithm, we categorize the nodes into three types: counties, ports, and neighboring countries. The association relationships are further classified into three levels: low, moderate, and high.

Table 2
Association rules for border tourism destinations in the Tibet sub-network.

Rule ID	Association rules	Confidence
1	Gyirong Port=> Gyirong^{3}	1
2	Gyirong^{3}=> Nepal	0.714
3	Nyalam^{3}=> Nepal	0.706
4	Zhongba^{3}=> Lhasa^{1}	0.686
5	Nyalam^{3}=> Lhasa^{1}	0.676
6	Zhongba^{3}=> Saga^{3}	0.657
7	Nyalam^{3}=>Gyirong^{3}	0.618
8	Zanda^{3} => Burang^{3}	0.615
9	Burang^{3}=> Lhasa^{1}	0.594
10	Nyalam^{3}=> Zhangmu Port	0.588
11	Zhongba^{3}=> Burang^{3}	0.571
12	Zhangmu Port => Gyirong^{3}	0.569
13	Gyirong^{3}=> Lhasa^{1}	0.548
14	Yadong^{3}=> India	0.541
15	Burang^{3}=> India	0.508
16	Lazi^{3}=> Zhongba^{3}	0.408
17	Nima^{3}=> Burang^{3}	0.407
18	Zhongba^{3}=> Nepal	0.4

Note: This table was created by the author.

calculating the network densities individually (Table 3). The sub-network densities for Xinjiang and Tibet were found to be 0.422 and 0.176, respectively, with a total network density of 0.121. It was observed that the nodes in the sub-network of Xinjiang are more tightly interconnected, whereas the nodes in the Tibet sub-network exhibit a lower degree of association. Consequently, it appears that the border tourism network possesses a loose structure and lacks equilibrium. These findings indicate that Xinjiang and Tibet constitute relatively autonomous networks and have not yet formed a regionalized cross-border tourism destination on a broader scale covering the entire Qinghai-Tibetan Plateau.

4.2.2. Cohesive subgroups

In a network, two nodes are considered structurally peer-to-peer if their connections with other nodes are identical, and such exact peer-to-peer relationships rarely occur in empirical data. The purpose of the cohesive subgroups analysis in this study is to assess the similarity model of pivotal node connections in Qinghai-Tibetan Plateau border tourism using the CONCOR algorithm, followed by the formation of structurally peer clusters. The outcomes of the analysis are determined by the node's interaction with the other remaining nodes within the network, irrespective of whether a direct connection exists between the two nodes.

Using UCINET to analyze the raw data for cohesive subgroups, eight cohesive subgroups were identified within the network, achieving a high model fit with a coefficient of determination (R-squared) of 0.609 when the segmentation depth was set at 3 (Table 4). The nodes in the Xinjiang sub-network were organized into subgroups 1 through 4, with both subgroup 3 and subgroup 4 exhibiting an internal connectivity density of 1.000 (Table 5), indicating dense direct connections within these subgroups. Notably, subgroup 4 is closely linked to the other three subgroups within the Xinjiang sub-network, suggesting its pivotal role in facilitating associations among these subgroups.

Within the Tibet sub-network, nodes coalesced into subgroups 5 to 8, all of which demonstrate closely interconnected internal links and a high level of reciprocity within their node structures. Subgroup 8 is particularly influential, showing strong connections with other subgroups and playing a significant role within the Tibet sub-network. Subgroup 7 consists of Nepal and its related nodes—Gyirong Port, Zhangmu Port, Nyalam^{3}, and Gyirong^{3}—forming a tightly knit cluster. Furthermore, Zhongba^{3} and Lhasa^{1} in subgroup 8 exhibit strong structural equivalence and are tightly linked to other clusters, benefiting from a network structure advantageous to tourism centers. Although the nodes in subgroup 7 maintain close internal links and are well-connected to Lhasa^{1} and Zhongba^{3}, their connections to the broader border tourism system are comparatively weak. This observation suggests that while these cities are core nodes within the Qinghai-Tibetan Plateau tourism industry, their direct links to border tourism development are not pronounced.

Table 3
Network and sub-networks density.

Network	Nodes	Density
Xinjiang sub-network	Aketedu^{3}, Pakistan, Gar^{3}, Hotan^{3}, Khunjerab Port, Kashgar^{1}, Yarkant^{3}, Tajikistan, Tashkurgan^{4}, Yecheng^{3}	0.422
Tibet sub-network	Damxung^{3}, Tingri^{3}, Konggar^{3}, Gyirong Port, Gyirong^{3}, Gyantse^{3}, Karuo^{1}, Lhasa^{1}, Lazi^{3}, Lankazi^{3}, Markam^{3}, Nepal, Nima^{3}, Nyalam^{3}, Burang^{3}, Saga^{3}, Sa'gya^{3}, Xining^{1}, Shangri-La^{2}, Yadong^{3}, India, Zanda^{3}, Zhangmu Port, Zhongba^{3}, Nyingchi^{1}	0.176
Total network	–	0.121

Note: This table was created by the author.

Table 4
Location of cohesive subgroups.

Sub-network	Cohesive subgroup	Typical counties
Xinjiang	1	Aketedu[3], Yarkant[3], Yecheng[3]
	2	Hotan[3], Gar[3]
	3	Pakistan, Khunjerab Port, Tajikistan
	4	Kashgar[1], Tashkurgan[4]
Tibet	5	Konggar[3], Xining[1], Damxung[3], India, Shangri-La[2], Burang[3], Zanda[3], Nyingchi[1], Nima[3], Lankazi[3], Yadong[3], Sa'gya[3], Markam[3]
	6	Saga[3], Gyantse[3], Karuo[1], Tingri[3], Lazi[3]
	7	Gyirong[3], Gyirong Port, Nepal, Nyalam[3], Zhangmu Port
	8	Zhongba[3], Lhasa[1]

Note: This table was created by the author.

Table 5
Density matrix of cohesive subgroups.

	1	2	3	4	5	6	7	8
1	0.000	0.000	0.111	1.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.500	0.000	0.000	0.000	0.000
3	0.111	0.000	1.000	1.000	0.000	0.000	0.000	0.000
4	1.000	0.500	1.000	1.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.077	0.031	0.015	0.462
6	0.000	0.000	0.000	0.000	0.031	0.900	0.040	0.700
7	0.000	0.000	0.000	0.000	0.015	0.040	0.800	0.600
8	0.000	0.000	0.000	0.000	0.462	0.700	0.600	1.000

Note: This table was created by the author.

4.2.3. *Nodal centrality*

The centrality level of a node directly correlates with its significance in accessing strategic resources within that network. Degree centrality measures a node's ability to form clusters and disseminate information, while closeness centrality evaluates how closely a node is connected to other nodes within the network. Nodes with high betweenness centrality are located in pivotal positions within the flow of resources in the network [44,46].

The centrality analysis of network nodes reflects their different functions, while centralization measures the balanced degree of centrality within the network. Centralization serves as a network analysis index based on node centrality, with larger values indicating a higher degree of network imbalance. In the context of border tourism, the sub-networks of Xinjiang and Tibet exhibit higher centralization for point degree centrality, closeness centrality, and betweenness centrality. This observation suggests that these two sub-networks are more centralized, with significant power disparities between nodes and notable advantages enjoyed by core nodes in

Table 6
Xinjiang sub-network node centrality.

Node	Degree centrality	Closeness centrality	Betweenness centrality	Node type
Kashgar[1]	9.00	0.111	20.17	County
Tashkurgan[4]	7.00	0.091	5.17	County
Tajikistan	5.00	0.077	0.67	Country
Pakistan	4.00	0.071	0.00	Country
Khunjerab Port	4.00	0.071	0.00	Port
Aketedu[3]	3.00	0.067	0.00	County
Yarkant[3]	2.00	0.063	0.00	County
Yecheng [3]	2.00	0.063	0.00	County
Hotan[3]	1.00	0.059	0.00	County
Gar[3]	1.00	0.059	0.00	County
Descriptive statistics				
Mean	3.80	0.070	2.60	-
Standard deviation	2.48	2.482	6.05	-
Network Centralization	72.22%	80.75%	54.22%	-

Note: This table was created by the author.

terms of network location.

The centrality measures, specifically closeness centrality, betweenness centrality, and degree centrality of nodes in the Xinjiang sub-network, are significantly correlated (Table 6). Kashgar[1], Tashkurgan[4], and Tajikistan are identified as relatively core nodes within this network. Both Kashgar[1] and Tashkurgan[4] exhibit substantially higher centrality measures compared to other nodes, positioning them as primary and secondary core nodes, respectively. In contrast, Gar[3] and Hotan[3] represent peripheral nodes characterized by lower centrality values. These nodes maintain minimal connections with other nodes within the sub-network. However, their strategic geographic locations at the intersection of Xinjiang and Tibet provide potential pathways for enhancing connectivity within the sub-network. By establishing links with the core nodes, Gar[3] and Hotan[3] can incrementally improve their positions within the network's structure.

Within the Tibetan sub-network (Table 7), Lhasa[1] exhibits the highest values for degree centrality, closeness centrality, and betweenness centrality, underscoring its pivotal role within the network. Nepal and India show significantly higher betweenness centrality compared to other nodes outside of Lhasa[1], with Gyirong[3] following closely. This highlights the critical function of these three nodes in facilitating connections among other nodes within the Tibetan sub-network. Saga[3] emerges as a secondary core within the network, attributed to its high degree centrality and betweenness centrality. This prominence is due to its strategic location at the intersection of two major self-driving tour routes in Tibet—G349 National Highway and G219 National Highway. This intersection

Table 7
Tibet sub-network nodal centrality.

Node	Degree centrality	Closeness centrality	Betweenness centrality	Node type
Lhasa[1]	22.00	0.038	209.75	County
Saga[3]	8.00	0.024	6.25	County
Nepal	8.00	0.025	21.75	Country
India	6.00	0.024	25.83	Country
Karuo[1]	5.00	0.022	0.25	County
Gyirong[3]	5.00	0.023	7.75	County
Tingri[3]	5.00	0.023	2.25	County
Burang[3]	5.00	0.023	1.50	County
Nyalam[3]	5.00	0.023	0.67	County
Gyantse[3]	5.00	0.022	0.25	County
Lazi[3]	5.00	0.022	0.67	County
Zhongba[3]	5.00	0.023	2.42	County
Zanda[3]	4.00	0.022	0.67	Port
Zhangmu Port	4.00	0.022	0.00	Port
Nima[3]	3.00	0.022	0.00	County
Gyirong Port	2.00	0.016	0.00	Port
Damxung[3]	1.00	0.020	0.00	County
Sa'gya[3]	1.00	0.020	0.00	County
Shangri-La[2]	1.00	0.020	0.00	County
Yadong[3]	1.00	0.015	0.00	County
Konggar[3]	1.00	0.020	0.00	County
Lankazi[3]	1.00	0.020	0.00	County
Markam[3]	1.00	0.020	0.00	County
Xining[1]	1.00	0.020	0.00	County
Nyingchi[1]	1.00	0.020	0.00	County
Descriptive statistics				
Mean	4.24	0.022	11.20	
Standard deviation	4.26	6.864	41.03	
Network Centralization	80.43%	83.62%	74.94%	

makes Saga[3] a crucial node for connectivity within the region [73].

5. Discussion

In this study, we employed the Apriori algorithm to conduct a comprehensive analysis of travelogue data from tourists visiting the Qinghai-Tibet Plateau border areas, successfully identifying a series of tourism hotspots and their interrelationships. Subsequently, using social network analysis, we structured an analysis of these tourism hotspots, aiming to evaluate the development of border tourism by examining network density, cohesive subgroups, and node centrality. Based on these analyses, targeted development strategies were proposed.

Despite the Qinghai-Tibet Plateau being a unified geographical entity, its border tourism network distinctly divides into two independent sub-networks: Xinjiang and Tibet. This segmentation evidently hinders the regional development of tourism on the plateau. Gar[3] and Hotan[3] serve as crucial linkage nodes between these two sub-networks, playing a significant role in the integrated development of border tourism across the Xinjiang and Tibet. Additionally, the network's tourism hotspots are primarily concentrated along the border transportation corridors, reflecting the strong dependence on transportation and the multi-destination nature of border tourism on the Qinghai-Tibet Plateau.

Six identified border tourism destinations have significant associations with neighboring countries. Ports in counties such as Gyirong[3], Nyalam[3], and Tashkurgan[4] have become focal points of border tourism, indicating that these counties' ports are not only geographical border passageways but also play crucial roles in actual border tourism activities, making them relatively mature tourism destinations.

Specifically, Gyirong[3], with its higher betweenness centrality compared to Nyalam[3], exhibits greater developmental potential. This could be attributed to the severe disruptions to the Zhangmu Port and its transportation facilities caused by the Nepal earthquake, which led to the temporary closure of the port. During the closure of the Zhangmu Port, the Gyirong Port became the main crossing to Nepal, thus attracting a significant number of tourists interested in Nepal. By enhancing tourism infrastructure, integrating tourism resources, and improving external transport accessibility, Gyirong[3] can effectively leverage its transit capacity to boost the regional tourism economy and enhance the overall efficacy of the tourism network.

Tashkurgan[4], already a secondary core within the Xinjiang sub-network, possesses good transport accessibility and a solid market size basis. The Khunjerab Port, famously featured in the Chinese movie "Visitors on the Icy Mountain", is well-known in China and was the starting point of the Karakoram Highway connecting China to Islamabad, Pakistan. In 2019, the port witnessed 30,332 cross-border movements, of which 22,809 were travelers, accounting for 75.20 % of the total movements [74]. Based on this foundation, Tashkurgan[4] could enhance its share in the border tourism market by diversifying border tourism products, increasing accessibility to Khunjerab Port, and improving customs clearance convenience.

Meanwhile, ports in counties like Zhongba[3], Burang[3], and Yadong[3] may still be underutilized and undeveloped due to limitations in transport connectivity, tourism infrastructure, or other factors. However, the reasons these three border counties have become hotspots in the network vary. Zhongba[3] and Burang[3] have higher centrality and moderate associations with Lhasa[1], the tourism center of Tibet, possibly because they are located along border highways and possess attractive tourism resources. For instance, Zhongba[3] contains the Jemayangzong Glacier, the source of the Yarlung Tsangpo River, while Burang[3] is at the heart of the Gang Rinpoche and Mapam Yumco sacred mountain and lake tourism area within the Kailash culture zone, recognized as "the center of the world" by multiple religions including Bon, Tibetan Buddhism, Hinduism, and Jainism. Despite Yadong[3]'s lower centrality and its peripheral position within the network, its geographical advantage being approximately 100 km from Gangtok, the capital of Sikkim, and its historical status as Tibet's most important trading port, continue to attract tourists interested in India. These three counties can leverage their existing tourist base to attract more visitors for border tourism through enhanced promotional activities and improved customs convenience. Thus, they could be developed as supplementary, rather than core, border tourism destinations on the Qinghai-Tibet Plateau.

For other supportive county and city nodes, Lhasa[1] and Kashgar[1], the tourism centers of Tibet and Xinjiang respectively, could increase accessibility to border tourism destinations and enhance promotion of border tourism at transportation hubs, guiding more tourists to border areas. Enhanced border tourism promotion in Nima[3] along the North Xinjiang- Tibet Line, Karuo[1] at the intersection of the North Sichuan-Tibet Line and the Tang-Tibet Ancient Road, Markam[3] at the junction of the South Sichuan-Tibet Line and the Yunnan-Tibet Line, Nyingchi[1] on the South Sichuan-Tibet Line, and major transportation hubs like Xining[1] in Qinghai and Shangri-La[2] in Yunnan, could attract more self-driving tourists to border tourism destinations.

The Qinghai-Tibet Plateau, with an average altitude exceeding 4000 m, presents harsh natural conditions characterized by thin air and frequent freeze-thaw disasters. These severe environmental conditions not only increase the difficulty of infrastructure construction but also pose significant challenges to the health of tourists from low-altitude regions. The border counties of this area are among the least developed in terms of transportation in China, with low accessibility. Tourists can only reach these areas by road, and most of the border counties have recently constructed roads with poor conditions. For instance, Motuo[3] was the last county in China to be accessible by road. The inadequacies in transportation infrastructure pose numerous challenges for tourists in planning and executing their trips, increasing the complexity and time costs of travel. Despite the unique natural and cultural landscapes that attract a large number of tourists to the border areas, the inadequate transportation facilities, lack of public transport, and the physical discomfort caused by the high-altitude environment may lead to negative evaluations from tourists, which could hinder the development of tourism in these regions. However, overcoming these challenges can also bring a sense of accomplishment and positive emotional experiences unique to the border areas. By gaining an in-depth understanding of tourists' experiences under high-altitude and poor transportation conditions, more effective tourism development strategies can be formulated. This, in turn, will promote the sustainable development of border tourism on the Qinghai-Tibet Plateau.

5.1. Theoretical implications

This study successfully captured the characteristics of border tourism from the tourist perspective by using travelogue texts as the data source. This approach not only deepens our understanding of tourists' motivations, behaviors, and experiences but also addresses the scarcity of tourism data in underdeveloped border areas.

Innovatively, this research combined Apriori association rule mining with social network analysis, using data mining techniques to identify the focal points of interest among border tourism tourists and the relationships between them. This facilitated the construction of a relational network for subsequent network analysis. The empirical validation conducted confirms the feasibility and effectiveness of this approach, providing a new methodological framework for handling and analyzing large volumes of unstructured tourism data.

Furthermore, an in-depth analysis of the network structure and characteristics of border tourism in the Qinghai-Tibet Plateau not only highlighted the multi-destination nature of border tourism but also provided theoretical support for tourism managers to optimize the spatial layout of tourism networks, develop regional tourism routes and products, and implement precise destination management strategies from a network science perspective.

5.2. Managerial implications

Firstly, by identifying highly focused border tourism destinations and their associated nodes, tourism managers can more accurately understand market demands and tourist behaviors, optimizing the spatial layout of the Qinghai-Tibetan Plateau border tourism network.

From a regional development perspective, tourism managers can leverage the findings of this study to introduce more targeted tourism routes and products. Integrating tourism resources from different destinations to create distinctive routes not only meets the diverse needs of tourists but also enhances the overall competitiveness of border tourism.

Additionally, by analyzing tourists' points of interest and behavioral patterns in-depth, managers can implement more targeted measures, such as strengthening infrastructure, enhancing service quality, and organizing unique activities. These efforts aim to improve the overall tourist experience and satisfaction levels.

This article focuses solely on the content of travel notes from Chinese travel social networking sites, overlooking the preferences of overseas cross-border tourists to the Qinghai-Tibet Plateau. Additionally, the analyzed travel notes are predominantly written by independent travelers, most of whom are young people from economically developed areas, hence the research results more reflect the travel preferences and experiences of this specific group. The hotspots network is an undirected symmetric network. Incorporating tourism flow data could establish a directed network, potentially leading to different analytical results. Selecting alternative research domains for comparative analysis could provide a more objective depiction of the development level of border tourism on the Qinghai-Tibet Plateau in China and facilitate the proposal of more rational development strategies at the national level. Future research should aim to address the unexplored aspects of this article, strengthen the theoretical foundation of border tourism research, and propose more pragmatic border tourism development strategies.

6. Conclusions

This study presents an innovative methodology by integrating Apriori association rule mining with social network analysis to examine border tourism through travelogue texts, particularly focusing on the Qinghai-Tibet Plateau. The approach not only provides a deeper understanding of tourist behaviors and motivations but also creates a theoretical framework for analyzing tourism text data effectively. Findings from this research are significant for tourism managers, offering insights into optimizing the spatial layout and enhancing the competitiveness of border tourism destinations.

(1) Network Structure Identification and Analysis. The border tourism network of the Qinghai-Tibetan Plateau is divided into two independent sub-networks: Xinjiang and Tibet. Gar[3] and Hotan[3] serve as key linkage nodes between these networks, playing a significant role in integrating tourism between Xinjiang and Tibet.
(2) Characteristics of Border Tourism Hotspots. The hotspots are primarily concentrated near transportation corridors, indicating that the development of border tourism heavily depends on transport hubs. Additionally, the distribution of these hotspots reveals the multi-destination nature of the tourism chain within border areas.
(3) Development Potential of Destinations. Gyirong[3] and Tashkurgan[4], as core destinations for border tourism on the Qinghai-Tibet Plateau, have excellent foundational strengths and significant development potential. Optimizing border management processes, simplifying visa and customs procedures, enhancing port accessibility and ease of passage, and establishing multilateral tourism cooperation agreements to jointly develop cross-border tourism products can effectively enhance the attractiveness of these areas and transform them into regional tourism hubs.
(4) Potential Tourism Destinations. Zhongba[3], Burang[3], and Yadong[3] have not yet fully realized their border tourism potential due to limitations such as the extent of port openness. These areas could gradually enhance their tourism appeal through policy support and infrastructure investment.

(5) Regional Development Strategies. Increasing the convenience of travel from Lhasa[1] and Kashgar[1] to border destinations and enhancing tourism promotion at major transportation nodes can attract more visitors to these locations, thereby boosting regional tourism flow and economic benefits.

This study has several limitations. Firstly, it focuses solely on travelogue content from Chinese tourism social media platforms, overlooking the preferences of international tourists towards the Qinghai-Tibet Plateau. Additionally, the analyzed travelogues are primarily written by independent travelers, most of whom are young individuals from economically developed regions, resulting in findings that predominantly reflect the preferences and experiences of this specific group. The hotspots network is an undirected and symmetric network; incorporating tourism flow data to establish a directed network could yield different analytical outcomes. Secondly, the study's scope is limited; selecting other research areas for comparative analysis with the Qinghai-Tibet Plateau could provide a more objective description of the development level of border tourism in the region, aiding in the formulation of more rational development strategies at the national level. Lastly, the study's content is limited as it focuses only on the place names mentioned in the travelogue texts. Further analysis of destination evaluations, tourists' feelings, and other related information in the travelogues would help clarify the motivations, needs, and experiences of border tourists. Future research should address these unexplored areas, strengthening the theoretical foundation of border tourism research and proposing more practical development strategies for border tourism.

Funding statement

This work was supported by the [the Second Tibetan Plateau Scientific Expedition and Research Program] under Grant [2019QZKK1004].

Data availability statement

Data included in article/supp. material/referenced in article.

CRediT authorship contribution statement

Siyue Zhang: Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Zhaoping Yang:** Supervision. **Cuirong Wang:** Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Fang, Special thinking and green development path of urbanization in Qinghai-Tibet Plateau, Acta Geograph. Sin. 77 (2022) 1907–1919, https://doi.org/10.11821/dlxb202208005.
[2] R.B. Bennike, M.R. Nielsen, Frontier tourism development and inequality in the Nepal Himalaya, J. Sustain. Tourism (2023), https://doi.org/10.1080/09669582.2023.2174129.
[3] A. Gelbman, D. Timothy, Differential tourism zones on the western Canada–US border, Curr. Issues Tourism 22 (2019) 682–704, https://doi.org/10.1080/13683500.2017.1304364.
[4] D.J. Timothy, J. Guia, N. Berthet, Tourism as a catalyst for changing boundaries and territorial sovereignty at an international border, Curr. Issues Tourism 17 (2014) 21–27, https://doi.org/10.1080/13683500.2012.712098.
[5] D. Newman, The lines that continue to separate us: borders in our ′borderless' world, Prog. Hum. Geogr. 30 (2006) 143–161, https://doi.org/10.1191/0309132506ph599xx.
[6] J. Wang, X. Chen, Z. Zhang, Spatial differences and drivers of tourism ecological security in China's border areas, Sustainability 15 (2023) 11811, https://doi.org/10.3390/su151511811.
[7] J. Gao, K. Cao, M. Xu, D. Kong, The evolution of Xinjiang's border tourism de- velopment and its affecting factors: a case study of Horgos, Journal of Chinese Ecotourism 12 (2022) 442–455, https://doi.org/10.12342/zgstly.20220016.
[8] T. Yu, B. Zuo, R. A, J. Gao, The spatial pattern and driving mechanism of tourism development in border regions of China, Econ. Geogr. 41 (2021) 203–213, https://doi.org/10.15957/j.cnki.jjdl.2021.02.022.
[9] J. Wang, Q. Zhang, X. Li, Spatial coupling and evolution of border tourism system: a case study of ruili city on the China-Myanmar border, Areal Research and Development 40 (2021) 86–92, https://doi.org/10.3969/j.issn.1003-2363.2021.03.015.
[10] J. Wang, L. Tian, X. Wu, Research on spatial structure and distributed model of border tourism system, Resource Development & Market 34 (2018) 123–127+138, https://doi.org/10.3969/j.issn.1005-8141.2018.01.023.
[11] Y. Shi, L. Zhong, T. Chen, Evaluation of border tourism competitiveness at the county level in China, Resour. Sci. 36 (2014) 1133–1141.
[12] L. Zhong, S. Zhang, Y. Shi, A. Zhang, Evaluation and utilization strategies of land border tourism resources in China, Resour. Sci. 36 (2014) 1117–1124.
[13] Q. Min, Research on the Spatial Structure and Pattern of Border Tourism Development: A Case Study of Yunnan Province, China Science Publishing & Media Ltd., Beijing, 2021.
[14] Y. Shi, The classification and characteristics of border county tourism competitiveness in China. Urban and Environmental Studies, 2019, pp. 60–75.
[15] S. Park, Y. Xu, L. Jiang, Z. Chen, S. Huang, Spatial structures of tourism destinations: a trajectory data mining approach leveraging mobile big data, Ann. Tourism Res. 84 (2020) 102973, https://doi.org/10.1016/j.annals.2020.102973.
[16] A. Yallop, H. Seraphin, Big data and analytics in tourism and hospitality: opportunities and risks, J. Tourism Futur. 6 (2020) 257–262, https://doi.org/10.1108/JTF-10-2019-0108.
[17] M. Banyai, T.D. Glover, Evaluating research methods on travel blogs, J. Trav. Res. 51 (2012) 267–277, https://doi.org/10.1177/0047287511410323.

[18] L. Zhong, S. Sun, R. Law, Movement patterns of tourists, Tourism Manag. 75 (2019) 318–322, https://doi.org/10.1016/j.tourman.2019.05.015.
[19] Y. Liu, X. Yao, Y. Gong, C. Kang, et al., Analytical methods and applications of spatial interactions in the era of big data, Acta Geograph. Sin. 75 (2020) 1523–1538, https://doi.org/10.11821/dlxb202007014.
[20] H.Q. Vu, G. Li, R. Law, B.H. Ye, Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos, Tourism Manag. 46 (2015) 222–232, https://doi.org/10.1016/j.tourman.2014.07.003.
[21] K. Zhang, Y. Chen, C. Li, Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: the case of beijing, Tourism Manag. 75 (2019) 595–608, https://doi.org/10.1016/j.tourman.2019.07.002.
[22] K. Banet, V. Naumov, R. Kucharski, Using city-bike stopovers to reveal spatial patterns of urban attractiveness, Curr. ISSUES Tour 25 (2022) 2887–2904, https://doi.org/10.1080/13683500.2021.2011841.
[23] H. Shao, Y. Zhang, W. Li, Extraction and analysis of city's tourism districts based on social media data, Comput. Environ. Urban Syst. 65 (2017) 66–78, https://doi.org/10.1016/j.compenvurbsys.2017.04.010.
[24] J. Lai, J. Pan, China's city network structural characteristics based on population flow during spring festival travel rush: empirical analysis of "tencent migration" big data, J. Urban Plann. Dev. 146 (2020) 04020018, https://doi.org/10.1061/(ASCE)UP.1943-5444.0000581.
[25] Y. Liu, Z. Sui, C. Kang, Y. Gao, Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data, PLoS One 9 (2014) e86026, https://doi.org/10.1371/journal.pone.0086026.
[26] H.-Y. Shih, Network characteristics of drive tourism destinations: an application of network analysis in tourism, Tourism Manag. 27 (2006) 1029–1039, https://doi.org/10.1016/j.tourman.2005.08.002.
[27] K. Pavlovich, The evolution and transformation of a tourism destination network: the Waitomo Caves, New Zealand, Tourism Manag. 24 (2003) 203–216, https://doi.org/10.1016/S0261-5177(02)00056-0.
[28] A. Weidenfeld, Tourism and cross border regional innovation systems, Ann. Tourism Res. 42 (2013) 191–213, https://doi.org/10.1016/j.annals.2013.01.003.
[29] A. Askew, M, C. Cohen, E. Erik, Pilgrimage and prostitution: contrasting modes of border tourism in lower South Thailand, Tour. Recreat. Res. 29 (2004) 89–104.
[30] A. Stoffelen, D.J. Timothy, Bordering, ordering and othering through tourism: the tourism geographies of borders, Tourism Geogr. (2023), https://doi.org/10.1080/14616688.2023.2291818.
[31] H. Van Houtum, T. Van Naerssen, Bordering, ordering and othering, Tijdschr. Econ. Soc. Geogr. 93 (2002) 125–136, https://doi.org/10.1111/1467-9663.00189.
[32] J. Gao, C. Ryan, J. Cave, C. Zhang, Tourism border-making: a political economy of China's border tourism, Ann. Tourism Res. 76 (2019) 1–13, https://doi.org/10.1016/j.annals.2019.02.010.
[33] D.J. Timothy, Political boundaries and tourism: borders as tourist attractions, Tourism Manag. 16 (1995) 525–532, https://doi.org/10.1016/0261-5177(95)00070-5.
[34] J. Yu, J. Yang, J. Li, L. Lin, Y. Ren, The role of tourism in promoting the urbanization of ethnic border areas: a case study of xishuangbanna, Sustainability 15 (2023) 15119, https://doi.org/10.3390/su152015119.
[35] D.J. Timothy, BORDERLANDS:An unlikely tourist destination? IBRU Bound. Secur. Bull. (2000) 57–65.
[36] B. Leick, T. Schewe, B.K. Kivedal, Tourism development and border asymmetries: an exploratory analysis of market-driven cross-border shopping tourism, Tour. Plan. Dev. 18 (2021) 673–698, https://doi.org/10.1080/21568316.2020.1837230.
[37] D.J. Timothy, J. Saarinen, A. Viken, Editorial: tourism issues and international borders in the nordic region, Scand. J. Hospit. Tourism 16 (2016) 1–13, https://doi.org/10.1080/15022250.2016.1244504.
[38] K. Vodeb, Cross-border regions as potential tourist destinations along the Slovene-Croatian frontier, Tourism Hospit. Manag. 16 (2010) 219–228, https://doi.org/10.20867/thm.16.2.7.
[39] T. Hong, T. Ma, T.-C. (T. C.) Huan, Network behavior as driving forces for tourism flows, J. Bus. Res. 68 (2015) 146–156, https://doi.org/10.1016/j.jbusres.2014.04.006.
[40] R. Merinero-Rodríguez, J.I. Pulido-Fernández, Analysing relationships in tourism: a review, Tourism Manag. 54 (2016) 122–135, https://doi.org/10.1016/j.tourman.2015.10.010.
[41] N. Scott, C. Cooper, R. Baggio, Destination networks: four Australian cases, Ann. Tourism Res. 35 (2008) 169–188, https://doi.org/10.1016/j.annals.2007.07.004.
[42] R. Baggio, N. Scott, C. Cooper, Network science: a review focused on tourism, Ann. Tourism Res. 37 (2010) 802–827, https://doi.org/10.1016/j.annals.2010.02.008.
[43] I. Paulino, S. Lozano, L. Prats, Identifying tourism destinations from tourists' travel patterns, J. Destin. Market. Manag. 19 (2021) 100508, https://doi.org/10.1016/j.jdmm.2020.100508.
[44] S.-H. Lee, J.-Y. Choi, S.-H. Yoo, Y.-G. Oh, Evaluating spatial centrality for integrated tourism management in rural areas using GIS and network analysis, Tourism Manag. 34 (2013) 14–24, https://doi.org/10.1016/j.tourman.2012.03.005.
[45] X. Wu, Z. Huang, X. Peng, Y. Chen, Y. Liu, Building a spatially-embedded network of tourism hotspots from geotagged social media data, IEEE Access 6 (2018) 21945–21955, https://doi.org/10.1109/ACCESS.2018.2828032.
[46] B. Liu, S. Sam Huang, H. Fu, An application of network analysis on tourist attractions: the case of Xinjiang, China, Tour, OR Manag. 58 (2017) 132–141, https://doi.org/10.1016/j.tourman.2016.10.009.
[47] D.G. Pearce, H.A. Schänzel, Destination management: the tourists' perspective, J. Destin. Market. Manag. 2 (2013) 137–145, https://doi.org/10.1016/j.jdmm.2013.06.001.
[48] Y. Zhang, B. Li, L. Liu, D. Zheng, Redetermine the region and boundaries of Tibetan Plateau, Geogr. Res. 40 (2021) 1543–1553, https://doi.org/10.11821/dlyj020210138.
[49] S. Shi, M. Liu, J. Xi, Hierarchical structure and organizational model of county tourism network of the Tibetan plateau, Land 11 (2022) 1880, https://doi.org/10.3390/land11111880.
[50] State Council, Several provisions of the state Council on the opening of ports. http://hangzhou.customs.gov.cn/customs/302249/302266/302267/3988494/index.html, 1985.
[51] China National Tourism Administration, Ministry of Foreign Affairs, Ministry of Public Security, General Administration of Customs, Provisional Measures for the Management of Border Tourism, 1997.
[52] Ministry of Culture And Tourism of the People' s Republic of China, Notice of the general office of the ministry of culture and tourism on the declaration of border tourism pilot zones and cross-border tourism cooperation zones. https://zwgk.mct.gov.cn/zfxxgkml/qt/202303/t20230329_941024.html, 2023.
[53] Q. Ge, J. Xi, S. Wang, Development Stages,Patterns and key strategic questions regarding border tourism in China, Resour. Sci. 36 (2014) 1099–1106.
[54] Ctrip Group Research Report, The OTA leader through the cycle, ansion international. https://www.vzkoo.com/read/20240219e110971e54343982db72d9c7.html, 2024. (Accessed 12 April 2024).
[55] How to let users "before traveling, first on the Mafengwo?, T. Wang. https://www.woshipm.com/operate/3291612.html. (Accessed 12 April 2024).
[56] PRODUCT ANALYSIS: Mafengwo - Can it truly bend the rules?, (n.d.). https://www.woshipm.com/evaluating/2714861.html (accessed April 12, 2024).
[57] J. Hipp, U. Güntzer, G. Nakhaeizadeh, Algorithms for association rule mining — a general survey and comparison, ACM SIGKDD Explor. Newsl. 2 (2000) 58–64, https://doi.org/10.1145/360402.360421.
[58] P. Harrington, Machine Learning in Action, People's Posts and Telecommunications, Publishing House, Beijing, 2013.
[59] T. Arreeras, M. Arimura, T. Asada, S. Arreeras, Association rule mining tourist-attractive destinations for the sustainable development of a large tourism area in hokkaido using wi-fi tracking data, Sustainability 11 (2019) 3967, https://doi.org/10.3390/su11143967.
[60] S. Liao, Y.-J. Chen, M. Deng, Mining customer knowledge for tourism new product development and customer relationship management, Expert Syst. Appl. 37 (2010) 4212–4223, https://doi.org/10.1016/j.eswa.2009.11.081.

[61] Y. Liu, L. Gao, C. Li, H. Li, Cluster analysis of China's inbound tourism market: a new multi-attribute approach based on association rule mining of tourist preferences at scenic spots, Asia Pac. J. Tourism Res. 26 (2021), https://doi.org/10.1080/10941665.2021.1887305.

[62] H. Xiao, Algorithm of apriori-based rural tourism driving factors and its system optimization, Mobile Inf. Syst. 2022 (2022) 3380609, https://doi.org/10.1155/2022/3380609.

[63] T. Liu, Y. Zhang, H. Zhang, X. Yang, A methodological workflow for deriving the association of tourist destinations based on online travel reviews: a case study of yunnan province, China, Sustainability 13 (2021) 4720, https://doi.org/10.3390/su13094720.

[64] R.A. Hanneman, M. Riddle, Introduction to Social Network Methods, Introduction to Social Network Methods, Intellectual Property Publishing House, Beijing, 2019.

[65] S. Yang, Franziska B. Keller, Social Network Analysis: Methods and Examples, Social Sciences Academic Press(China), Beijing, 2019.

[66] L.C. Freeman, Centrality in social networks conceptual clarification, Soc. Netw. 1 (1978) 215–239, https://doi.org/10.1016/0378-8733(78)90021-7.

[67] L.C. Freeman, A set of measures of centrality based on betweenness, Sociometry 40 (1977) 35–41, https://doi.org/10.2307/3033543.

[68] G. Sabidussi, The centrality index of a graph, Psychometrika 31 (1966) 581–603, https://doi.org/10.1007/BF02289527.

[69] X. Zhong, J. Liu, Y. Gao, L. Wu, Analysis of co-occurrence toponyms in web pages based on complex networks, Phys. Stat. Mech. Its Appl. 466 (2017) 462–475, https://doi.org/10.1016/j.physa.2016.09.024.

[70] Y. Hu, X. Ye, S.-L. Shaw, Extracting and analyzing semantic relatedness between cities using news articles, Int. J. Geogr. Inf. Sci. 31 (2017) 2427–2451, https://doi.org/10.1080/13658816.2017.1367797.

[71] China Association Of Port-Of-Entry, China's Ports-Of-Entry 2021 Yearbook, China Customs Press, Beijing, 2022.

[72] S. Ye, J. Xi, Identification and evaluation of tourism corridor: a case study of Qinghai-Tibet Plateau, Acta Geograph. Sin. 78 (2023) 2630–2644.

[73] Three national highways inject new vitality into Tibet's tourism development, Tibet Daily (newspaper). http://www.tibet.cn/cn/culture/wx/202401/t20240123_7562190.html (accessed April 12, 2024).

[74] China Association Of Port-Of-Entry, China's Ports-Of-Entry 2019 Yearbook, China Customs Press, Beijing, 2020.