Annals of Tourism Research Empirical Insights 3 (2022) 100079

Contents lists available at ScienceDirect

Annals of Tourism Research Empirical Insights

journal homepage: www.sciencedirect.com/journal/annals-of-tourism-research-empirical-insights

Harnessing social media to understand tourist travel patterns in muti-destinations

Jinyan Chen [a,b,*], Susanne Becken [b], Bela Stantic [c]

[a] School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China
[b] Griffith Institute for Tourism, Griffith University, Australia
[c] School of Information and Communication Technology, Griffith University, Australia

ARTICLE INFO

Editor: Dr Kirilova Ksenia

Keywords:
Travel patterns
Social network analysis
Social media
Core-periphery theory
Centrality

ABSTRACT

Understanding travel patterns is helpful for decision-makers to draw insights from consumers' perspectives. This work took advantage of social media and analysed tourists' travel patterns from the length of itineraries and duration of stay. Using Chinese tourists in Australia as a case study, results showed that most visitors prefer to stay in two core destinations, with an average duration of 8.5 days, while adding one destination increases the stay by around 2.5 days and caps at approximately 14 days. The travel patterns were further analysed by social network analysis and explained the network structure using core-periphery theory. The results were compared with official national survey data and demonstrated encouraging accuracy, which provides practical implications for destination planning and management.

1. Introduction

Management of tourist destinations can be significantly enhanced by understanding visitors' mobility patterns (Buhalis, 2000). One of the biggest challenges for the tourism industry is to balance the concentration and the dispersal of visitors across destinations. What distribution level is deemed suitable or preferable depends on many factors, such as the availability of infrastructure, transportation networks, and community attitudes towards tourists, and is also likely to differ for various stakeholders (Becken & Simmons, 2019). Most governments pursue regional dispersion of visitors' economic impact as a major policy goal (World Tourism Organization, 2019). Information on whether and how visitors travel among destinations is therefore of great relevance.

Tourists build travel itineraries based on their preferences with limited resources, such as time, budget, distance among destinations, or available transportation (Yang, Li, & Li, 2017), which leads to some destinations being visited more while others are less. The inequalities and uneven power among the destinations could shape tourists' travel patterns. The inequity in destination connections causes some destinations to be in favour, even might suffer from over-tourism (Shoval, Kahani, De Cantis, & Ferrante, 2020), while other destinations, such as regional areas, may lack tourists. This uneven power among the destinations and the dispersal of tourists can be explained by the core-periphery theory. Friedman introduced the concept of core-periphery when analysing the development of Venezuela (Friedmann, 1966). He believed that countries are made up of core and periphery regions. A core region could be a city or a cluster of locations near the city with developed industries, fast-growing economies, and high population density, while areas in the periphery would have limited development and low population. Adopting the core-periphery theory, destinations often visited can be recognised as core, while those not popular are recognised as periphery. Therefore, the core-periphery theory offers a framework that helps understand travel pattern structures (Lai & Li, 2012).

Research on travel patterns or tourist mobility relies on government statistics, surveys, blogs, Global Positioning Systems (GPS), or other mobile data (Hardy, Birenboim, & Wells, 2020a; Y. Li, Xie, Gao, & Law, 2021; Shoval & Isaacson, 2007). The government's official statistical data are often static and can not trace the travel patterns from one destination to another. For example, the report from Tourism Research Australia (2019a) could show the number of Chinese tourists arriving in Australia, but the report can not identify where tourists travelled from one place to another.

Another type of data used to trace movements is the call records collected by mobile operators, which provide good coverage of locations for individual-level mobility analysis (Raun, Ahas, & Tiru, 2016;

* Corresponding author at: School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China.
E-mail addresses: jinyan.chen@polyu.edu.hk (J. Chen), S.Becken@griffith.edu.au (S. Becken), B.Stantic@griffith.edu.au (B. Stantic).

https://doi.org/10.1016/j.annale.2022.100079
Received 30 October 2021; Received in revised form 14 November 2022; Accepted 18 November 2022
Available online 25 November 2022

Saluveer et al., 2020). However, it is challenging to collaborate with companies that are willing to share the calling history data. Alternatively, people's movements can also be tracked by a device that records their locations (Shoval, 2010) or by installing an app to collect movements (Hardy, Birenboim, & Wells, 2020b). But the data collection process can be costly to collect data that may be restricted to a relatively small number of participants.

On the other hand, social media and its associated metadata provide another new data source with several advantages. The social media platform offers a vast volume of data, and the geo-located posts could also show the high spatial and temporal resolution of where tourists went (Chen, Becken, & Stantic, 2021b), which could be applied to analyse visitors' spatial and temporal movements. The rich content includes text, pictures, and user-profile information that could also be used to understand the visitors' satisfaction or factors that drive travel behaviours. Furthermore, social media data is often free of charge. For example, the Application Program Interface is usually provided by the company to allow people to access data under certain conditions (e.g. Twitter). Besides, the data collection can go backward with certain months or even years. Since the data can be collected when the user posts, it may reduce the bias in recalling memories compared with direct surveys.

Therefore, this work took advantage of the social media data and targeted understanding of tourist movements in multiple destinations. Such as the length of the itineraries, duration of stay, and potential reasons shaping such travel patterns. The length of the itineraries was extracted from visitors' social media posts. Furthermore, the duration of the stay was analysed based on the length of the itinerary, which provides a more detailed understanding of the visitor's behaviour and movements in space and time. Additionally, the travel pattern was modelled by social network analysis, and the network structure was explained using core-periphery theory. Based on the core-periphery theory, a comparison between travel patterns and domestic flight connections was conducted to understand the potential factors that shape such travel behaviour.

The case study adopted Chinese tourists in Australia and selected ten destinations from Australia, demonstrating tourists' spatial and temporal movement patterns. The reason for choosing Chinese tourists in Australia as the case study was the practical implication. Before the COVID-19 pandemic, China was the top tourist market for Australia. In 2019, 1.3 million Chinese tourists visited Australia, contributing 12.4 billion dollars to the Australian economy (Tourism Research Australia, 2014). However, due to travel restrictions under COVID-19, Chinese arrivals into Australia decreased by 90% in February 2020, compared with the same period last year (Australian Bureau of Statistics, 2020). The impact on the national economy is substantial (Cheng, 2020), and discussions on how best to recover tourism are ongoing (Tourism Australia, 2020a).

Regardless of the speed of recovery, it is important to understand where people travel once at their destination and what they want to experience, which creates value for re-organisations and improves the resilience of tourism recovery. An intense crisis like COVID-19 suppresses travelling, but the demand may become a strong rebound once the situation is controlled (McKercher, 2021). Therefore, gaining insights into visitors' dispersal and travel preferences from the past can be useful in planning for the return of visitors. The results could provide practical implications for Australian tourism stakeholders to understand Chinese tourists' travel patterns in Australia and theoretical contributions to core-periphery theory in understanding tourists' travel patterns. The proposed method and findings could be applied to other regions.

2. Literature review

2.1. Tourist movement patterns

Travel patterns represent models of the spatial-temporal relationship among the destinations on a particular trip, which is essential information for planning and managing tourism destinations and attractions (Shoval, 2018). Since the early stage, understanding tourist movements have attracted significant research attention. For example, Oppermann (1995) identified different types of travel flows, distinguishing single and multiple destination patterns, whereby multiple-destination visitors tend to have a wider geographic dispersal. Leiper (1989) proposed that travel patterns included three factors in the trip: origin, destinations and links between them. The intra-destination movement patterns further consider a destination as an area that tourists visit within a day (Lew & McKercher, 2006) or where they spend the night. They usually used a city as the destination and tried to map tourist movements from one attraction to another to identify travel patterns. Chen, Becken, and Stantic (2021a) identified that nearly 50% of eightytwo reviewed papers focused on the city scale. Raun, Shoval, and Tiru (2020) further explained that studies regarding the intra-national level suffered from insufficient data, which caused limited resources to conduct the intra-national tourist movements.

Traditionally to identify travel behaviour, research focused on ticket-selling numbers, revenue generation, and the number of tourists of arrival from statistical data. Although such data could provide some vision of tourist movements, the data is typically obtained from a statistic report posted by authorities or about only one particular location, which suffers from time lag and may not reflect the current situation. The statistical data also lacks information to identify the movement of individuals. Therefore, they cannot fully reflect the dynamic nature of tourist movements nationwide. To understand tourist travel patterns from one place to another, particularly on a broad scale like nationwide, tracing tourist movements and identifying their travel patterns requires more extensive and detailed data (Kádár & Gede, 2021). Advances in recent technologies enabled new data sources that can be used to better track the movements of tourists. The new types of data can be grouped into three categories: Mobile phone data, Tracking devices, and Social media data. These data often have characteristics of being of a large volume, having rich content, and having flexibility regarding the data collection.

For the mobile phone data, Birenboim, Anton-Clavé, Russo, and Shoval (2013) tracked visitors' temporal activities in the park and explored the dispersal of the visitors. Interestingly, they found that the visitor's movement patterns were stable and possibly repeated themselves in different situations in the park. Lewis, Hardy, Wells, and Kerslake (2021) tracked wine tourists' travel patterns by adopting similar technology. They found that although being wine tourists, the main reasons driving wine tourists to visit the cellar are more related to other attractions in terms of extending the length of the stay. Similarly, Park and Zhong (2021) used the sensor data from mobile phones (e.g. call records and mobile signalling data). With the particular case study of locations, the authors identified 12 different travel patterns. For example, tourists who only stay in the hotel for the day or travel within a 200 m radius of the accommodation fit in the type of "no movement". If tourists travel more in a list of attractions but always return to the accommodation where they started, it is called a cycle pattern. Within the cycle patterns, based on the particular destinations, tourists may visit a chain of places and return to the surrounding attractions. Modelling the tourist travel patterns using network analysis could help to understand tourist travel behaviour and destination management. Another method of using mobile to collect mobility data is to install applications on participants' mobile and track their movements, which also gained momentum because it provides high resolution and details of full tourist trips. For example, using a GPS device, which produces high-resolution data on tourist trajectory, can help to model tourist flows on a national scale (Raun et al., 2020) or focus on a particular type of tourist movements, such as wine tourists or cruise tourists (Hardy, Vorobjovas-Pinta, Wells, Grimmer, & Grimmer, 2021; Lewis et al., 2021).

However, many constraints still exist in using a mobile phone, mobile applications, or tracking device data. Collaborating with companies willing to share the calling history data has been reported as challenging, and using the tracking device is costly and restricted to a relatively small number of participants. For example, Raun et al. (2016) claimed that to access company data, different parties (data providers, government regulations) have to give access. Using the tracking device requires volunteers to carry devices and then record their movements (Shoval, 2010) or install an app to collect location data (Hardy et al., 2020b), but those methods still suffer from the number of participants or the data often is collected within specific areas.

With social media becoming an important part of life, people use social media to post stories from their daily lives, and they are particularly likely to share their travel experiences (Liu, Wu, Li, & Robert)., 2019). For the tourism industry, it has become an important channel for understanding tourists through social media data. For example, Chen et al. (2021a) modelled travel flows globally, extracting travel sequences from social media data. Also, García-Palomares, Gutiérrez, and Mínguez (2015) identified tourists' hot spots in Europe using Panoramio data and found that visitors' dispersion has a higher spatial concentration. Koo, Lau, and Dwyer (2017) found that visitation concentrated at famous places, with the number of tourists decreasing in less famous regions, representing market-specific power-law exponents. The key finding was that, while tourists continue to visit core destinations as volumes grow, there is a need for expanding the number of smaller, peripheral destinations to accommodate demand.

Conceptualising tourism travel patterns as travel networks have led researchers to apply social network analysis to analyse travel flows (Asero, Gozzo, & Tomaselli, 2016; Baggio, Scott, & Cooper, 2010). The network consists of a set of nodes and the ties which link these nodes, and in a travel network, destinations are described as nodes, and the movement among the destinations can be recognised as travel flows (Baggio et al., 2010; Casanueva, Gallego, & García-Sánchez, 2016; Scott, Cooper, & Baggio, 2008). For example, a recent study on international travel in Shanghai used Flickr data to identify that the majority of tourists are concentrated around the area of interest in the city centre, with areas of interest (nodes) and tourist routes (edges) containing power-law features (Mou et al., 2020). Also, a social network analysis was conducted to identify the spatial structure of the tourist attraction system in Seoul (Kang, Lee, Kim, & Park, 2018). Networks can also be segmented into different sub-networks based on the travel modules to uncover travel patterns. For example, Bendle (2018) examined the seven sub-networks and found that travel scales influence travel networks.

Furthermore, some research tried to explain the factors that influence travel patterns. For example, Huang and Wu (2012) used the space-time concept to analyse tourist spatial-temporal behaviour in the Summer Palace in Beijing and found that tourists could be clustered in seven temporal-spatial patterns. They further explored the relationship between the travel patterns and length of stay and found that the length of stay decreased with the time limit, especially due to existing time constraints (ending the trip at the Gate), by investigating the dispersal of international visitors in Australia. Gao, Ye, Zhong, Wu, and Liu (2019) investigated the intercity travel flows in China through complex network analysis. They detected four main destination communities based on visitors' travel networks, and findings indicate that visitors prefer to travel within one community rather than between communities, especially tourists from similar cities who prefer a similar choice of destinations. However, the literature often focuses on analysing the tourist flow and identifying important destinations in the network (hotspots), ignoring the connection of the destinations in the travel patterns or networks, such as how to explain such tourist flow from one destination to another and what might be the reasons to shape such travel behaviour. Therefore, this paper adopted the core-periphery theory to explain the structure of the networks and try to understand travel behaviour.

2.2. Core-periphery theory and social network analysis

When the concept of core-periphery was introduced, scholars mentioned that in different content, society was constructed by core and periphery factors. A core region could be a city or a cluster of locations near other places with developed industries, fast-growing economies, and high population density, while periphery factors would have limited development and low population (Lai & Li, 2012). Traditionally the core-periphery perspective assumes the systems work through the core to the periphery (D. Weaver, Moyle, McLennan, & lee., 2021). However, it was also argued that the peripherality places might also constitute sufficient opportunities (D. B. Weaver, 2017). Therefore this paper adopted the core-periphery theory to analyse the travel patterns to understand tourist travel in muti-destinations and build the travel network among the selected destinations using the social network analysis.

When applying social network analysis, centrality is an important measurement. Different types of centrality measure the property of the nodes and the edges differently, and several metrics are commonly used to describe the structure of a network. For example, identifying popular locations by calculating the centrality, specifically degree centrality, was applied to examine tourist flows in Beijing during the Olympic games. Findings indicated that international tourists preferred famous traditional attractions and focused their activity in the central area of Beijing (Leung, Fang, Bihu, Billy, & Zhuhua, 2012). Similarly, Zeng (2018) calculated three types of centralities (degree, closeness, and betweenness centrality) and concluded that Chinese tourists in Japan are mainly concentrated in central Japan.

Using social network analysis to identify core-periphery structure has also been studied (J. Gao, Peng, Lu, & Claramunt, 2022; Su, Stepchenkova, & Dai, 2020; Z. Wang, Liu, Xu, & Fujiki, 2020). In the early stage, Borgatti and Everett (1999) already proposed that core nodes in the network should be a high degree of centrality (high number of connections) and are also strongly connected by nodes with a high degree of centrality. Therefore different types of centrality have been applied to identify the core nodes in the network. However, the definition between core and periphery nodes has not been clearly addressed with the eigenvector centrality. Wang, Li, and Lai (2018) and Su et al. (2020) mentioned taking the dominant node eigenvector centrality as a core and then computing the rest. Z. Wang et al. (2020) studied the destinations in the network using point centrality, closeness centrality, and betweenness centrality. Gao et al. (2022) described the top 20% of attractions in their study, with the most prominent strength as core nodes.

Among all the different types of centrality, eigenvector centrality is better at identifying the core nodes in the network. Eigenvector centrality is not only based on the high number of connections of the node but also on the high connection of the nodes that are adjacent with (Y. Wang et al., 2018). Therefore The work considered the dominant eigenvector centrality as a core in the network and then calculated the mean of the rest of the eigenvector centralities. If the destination had an eigenvector centrality higher than the mean, it was classified as a semi-core; otherwise, it was deemed a periphery. Another question in the literature is the vague definition of "dominant", as it is unclear what threshold to consider the node as dominant. Therefore, this work bridged the gap by introducing the "elbow method" (Becken, Stantic, Chen, & Connolly, 2022) to clarify the dominant nodes (core destinations) based on eigenvector centrality.

3. Methodology

3.1. The framework to retrieve social media data

The volume of Weibo posts made in Australia by Chinese travellers is significant, and it is important to filter out those posts related to travel to specific locations of interest. According to Tourism Research Australia, Chinese education visitors continue to spend more than any other market, and 88% of them visit capital cities (Tourism Research Australia, 2019a). Therefore, the names of eight capital cities of Australian states and territories were considered. In addition, the literature emphasised that *Uluru* (Ayers Rock) has a high reputation among Chinese visitors (Pan & Laws, 2003), and Cairns is seen as the gateway to the *Great Barrier Reef* (Cai, Hio, Bermingham, Lee, & Lee, 2014). The importance of these two destinations to Chinese tourists is also reflected in work presented by Tourism Australia (2020b). Eventually, ten Australian destinations were identified for data collection: *Adelaide, Brisbane, Cairns, Canberra, Darwin, Hobart, Melbourne, Perth, Sydney, and Uluru*.

Data were obtained from Chinese social media Sina Weibo by relying on the keyword search function from their 'Application Program Interface' (Weibo, 2012) enhanced by a dedicated build crawler in Python programming language. Data were initially stored in NoSQL (Not only Structured Query Language) MongoDB database, which is suitable for storing and manipulating data in JavaScript Object Notation format (Stantic & Pokorný, 2014). After processing the data, extracted patterns were stored in a Relational database MySQL, for easier processing and analytics by harnessing powerful Structured Query Language.

The timeframe for data collection was defined from January 2018 to the end of December 2018 to cover the whole calendar year since different seasons attract different visitors. Data from 2018 is appropriate to assess the number of Chinese tourists in Australia because, after 2018, the number of Chinese visitors was initially impacted by the Australian bushfires in 2019 and later by the COVID-19 pandemic from early 2020 (Tourism Research Australia, 2019b). Fig. 1 shows the data collection framework and data pre-processing. Data collection meets the requirements from the model of ethics of tracking framework (Hardy, 2020). Data were obtained with the Application Program Interface provided by Weibo(Weibo, 2012), allowing people to access data under their terms and conditions. Furthermore, all collected data were stored in a secured database that only people with permission could.

access (NoSQL MongoDB database and MySQL database are restricted with access with Univerity firewall and password required). Besides, posts were presented by userID in the database, which is a random string of numbers and does not reveal private information. We also followed (Chen et al., 2021a), and all results were generated at an aggregated level so that no individual could be identified from the posts to respect user privacy.

When using social media data, one of the key questions is how to

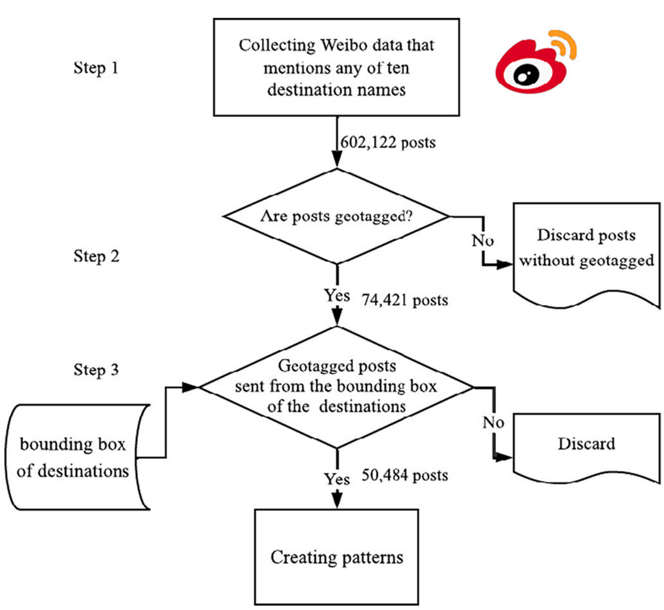


Fig. 1. Data collection framework and data pre-processing.

filter tourists from local users. There are different methods to conduct the process, such as based on the duration of the stay or user profile (Chen et al., 2021b), but there have no standard recognised methods to do so. To identify the tourists and differentiate them from locals, this paper proposed that tourists should visit more than one destination for a predefined duration of time. The reason behind this decision is related to the fact that locals mostly post only from one destination. The data also shows that a significant proportion of social media users posted only from one destination, which is most likely their place of residence. We acknowledge that some tourists who post only from one destination might be missed, which is a limitation that future research needs to address that further examine sub-groups within the sample of Weibo users.

Weibo posts contain latitude and longitude geotagged data, demonstrating wherefrom the posts were made, which is important for modelling travel patterns. Therefore this work first only considered Weibo posts attached with geotagged data. Then, a bounding box of each destination was recognised individually by using (Latlong, 2019), and only geotagged posts sent within the belonging bounding box of destinations were considered for further analysis. This step was essential as it discarded many irrelevant posts that mentioned the destination names by people who had never travelled to Australia and ensured that the Weibo posts were posted by users while being at the actual destination.

Table 1 presents the number of posts in different stages of data cleaning. Initially, there were 602,122 posts collected in step 1 from ten selected destinations; filtering and restricting to only geotagged posts resulted in 74,421 posts in step 2. Step 3 matched the geotagged posts with the bounding boxes of the destinations; eventually, it was found that about 68% of geotagged data were posted from destinations, which resulted in 50,484 posts used for further analysis and to retrieve the travel itineraries. The exact number of posts collected in each step at individual locations is listed in Table 1.

3.2. Detecting core-periphery destinations from travel network

The Eigenvector centrality is a measure of the influence of a node in a network. It considers not only how many connections a node has (i.e., its degree) but also the centrality of the vertices that it is connected to. A high eigenvector score means that a node is connected to many nodes, which themselves have high degrees. The Eigenvector centrality can be calculated by Eq. 1, where the Eigenvector centrality A_x of a node x is proportional to the sum of the centralities of the nodes to which it is connected, and n is the total number of nodes in the network.

$$A_x = \sum_{j=1}^{n} a_{ij}\, x_j, i = 1, \dots, n \tag{1}$$

In the literature, Eigenvector centrality was used to identify core destinations (Y. Wang et al., 2018). However, it has not clearly

Table 1
Number of Weibo posts after three steps of data pre-processing.

Destinations	Step 1 Collected posts by keywords	Step 2 Extracted geotagged posts	Step 3 Matched geotagged posts with destinations' geo-bounding box
ADELAIDE	22,416	2144	1208
BRISBANE	44,665	5388	3913
CAIRNS	25,205	3414	1883
CANBERRA	16,128	1616	950
DARWIN	55,474	2380	286
HOBART	4877	1151	865
MELBOURNE	180,602	26,241	18,944
PERTH	17,559	2475	1380
SYDNEY	232,449	29,147	20,881
ULURU	2747	465	174
Total	602,122	74,421	50,484

explained the threshold to separate the core and periphery nodes. This work bridges the gap and introduces the "elbow method" to identify the dominant nodes (core destinations) based on the eigenvector centrality and determine the threshold to recognize other nodes in the core and periphery theory. The Elbow method is widely accepted and often used to find the optimal number of clusters by calculating the corresponding errors consisting of plotting the values and identifying the elbow on the curve as the optimal value (an example shown in Fig. 3).

3.3. The length of itinerary and duration of stay

After collecting the data that meet the requirements from Fig. 1, the number of posts collected in each destination is shown in Fig. 2.

Modelling the travel patterns from the large unstructured social media data is another changeling step. To do so, the userID (which represents a 10-digit random number) and posting time are needed to recognize the travelling sequence. Data in each destination was considered as a separate dataset and stored in a structured format in a relational database (MySQL) for more efficient processing. A list of travel sequences was identified by finding the sequence of the posting time, locations, and userIDs. Since this work focused on travel patterns, users who only showed at one destination were omitted, which also reduced the possibility of posts being made by locals.

After extracting the travel sequences, the travels among the destinations were segmented based on the length of itineraries. The length of the itinerary represents the number of destinations that a user visited. The itineraries were directed; therefore travelling from destination A to B differs from destination B to A. The contribution to the core-periphery was analysed through each itinerary length to determine whether the length of the itinerary influences the core-periphery structure. Following the length of the itineraries, the number of days visitors travelled can also be identified. The duration of stay is an important indicator for understanding tourists' behaviour in the destination. A primary destination may attract more tourists, but a group of periphery attractions can attract tourists to stay longer than isolated places (Lue, Crompton, & Fesenmaier, 1993). Chen et al. (2021a) pointed out that the duration of stay is also an essential factor in identifying tourists from residents using social media data. If the trip length was within the same month or a maximum of two consecutive months, the person could be considered a tourist.

Current reports from survey data can calculate the average duration of stay. However, most studies did not consider the duration of stay based on the length of itineraries. The relationship between stay duration and itinerary length could reveal important information for destination management because tourists plan their travel based on limited time and budget. For example, in this research, if tourists would visit beyond the two core destinations and how long they might increase their stay. To achieve this, the time duration between the last and the first post from a Weibo user was used. This research considered the duration of stay based on the length of the itinerary. Therefore each itinerary was imported to the MySQL database, and we used the datediff() function to calculate the trip length.

4. Results

4.1. Identifying core-periphery destinations

The travel sequences were separated into pairs and a matrix to

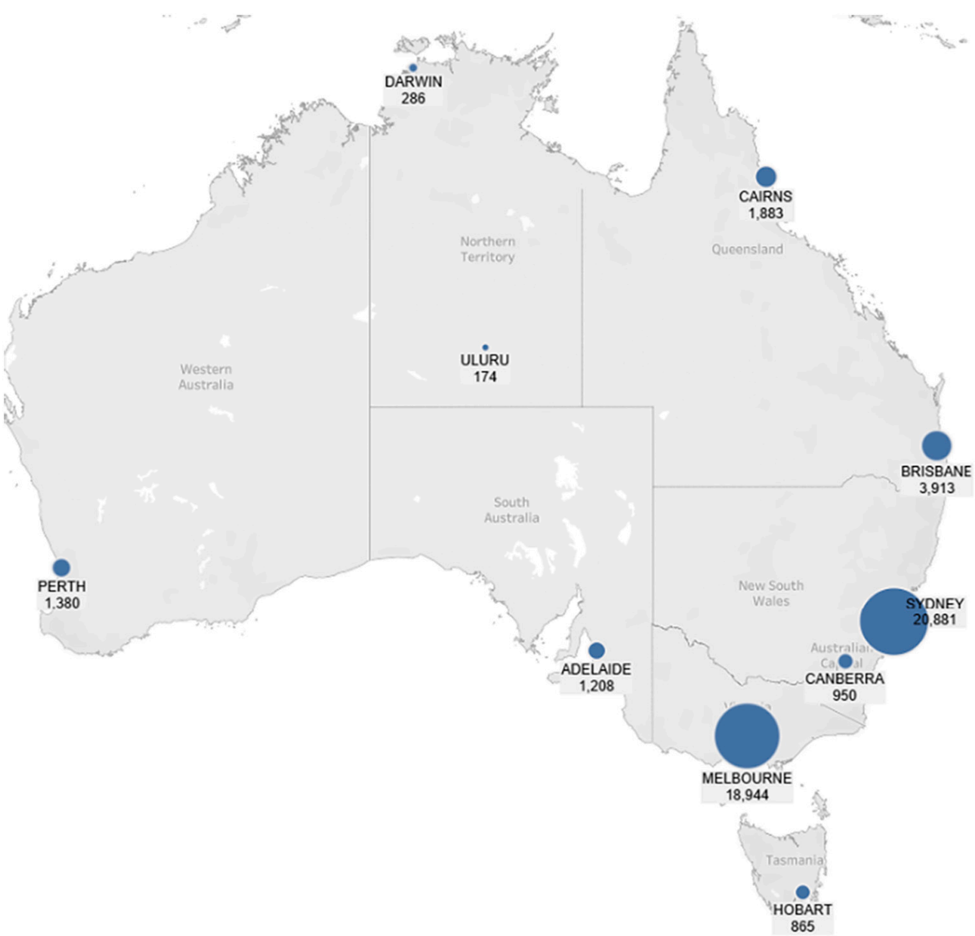


Fig. 2. Number of posts in the selected destinations.

analyse the travel patterns as a network. For example, if five tourists visited destinations A, B, and C, the travel sequences will be transferred from A to B with a volume of five, and B to C also with a value of five. Following this data matrix, a travel network was generated.

The value of eigenvector centrality was calculated using Eq. (1), shown in Table 2. Following the eigenvector centrality, the Elbow method was applied to identify the core-periphery destinations. For example, to identify core destinations, the distribution of the eigenvector centrality was plotted on a line graph, and the elbow is evident where a value dropped suddenly and is labelled in Fig. 3. Based on this method, two destinations were considered core as they appeared before the "elbow" (Sydney, Melbourne). Following the earlier work (Y. Wang et al., 2018), the mean value of the rest of the (none core) destinations was calculated (0.1903). If the destination's eigenvector centrality is higher or equal to the mean value, the destination is classified as a semi-core, in this case: Brisbane, Cairns, Hobart, and Canberra; Otherwise, the destination is considered a periphery: Adelaide, Perth, Uluru, and Darwin.

4.1.1. *Travel network characteristics*

Eigenvector centrality from Table 2 was applied to build the destination network, shown in Fig. 4. The node represents different destinations, and the node's size is proportional to the Eigenvector centrality. The node colour represents the core (Red), semi-core (Purple), and periphery destinations (Green). The visualization and network characteristics are analysed in the software Gephi (Gephi, 2009). The travel network shows the characteristics of the travel networks. For example, the average number of connections between destinations was 381.6, while the density of the overall travel network was 0.933 (out of 1). The density of the network density shows how well the whole network was connected effectively (A. Li et al., 2020). Therefore, those network characteristics implied that 93% of the destinations were directly connected by tourists' travel itineraries. Both parameters showed that the travel network among the selected destinations was closely connected.

The results confirmed the surveys from Tourism Research Australia (Tourism Research Australia, 2019a) that Sydney was the most connected destination considering that it hosts Australia's busiest airport, followed by another core destination, Melbourne. The network showed no connection between Darwin and Perth, which suggests that Chinese tourists did not prefer to travel to both periphery destinations on the same trip.

The highly connected travel network also showed a diverse core-periphery structure. Table 3 shows the detailed network structure based on core-periphery theory. Core destinations are represented by C, semi-core by S, and periphery destinations by P.

Table 3 shows all nine core-periphery structures. Most trips started from a core destination (63.50%), with more than 33% of users preferring to travel between the two core destinations. The high percentage between core destinations could be due to the international fame of the destinations and the international airports' connections with China, leading to a high level of people arriving and departing. Travelling from a core to a semi-core destination shared similar interests as semi-core to the core, with around 23% and 22%, respectively. Starting the trip from periphery destinations had the least interest, with only 8.20% in total, but travelling from a periphery to a core destination shared a similar percentage with travelling from core to periphery destinations, which could relate to the departing and arriving from the international airports.

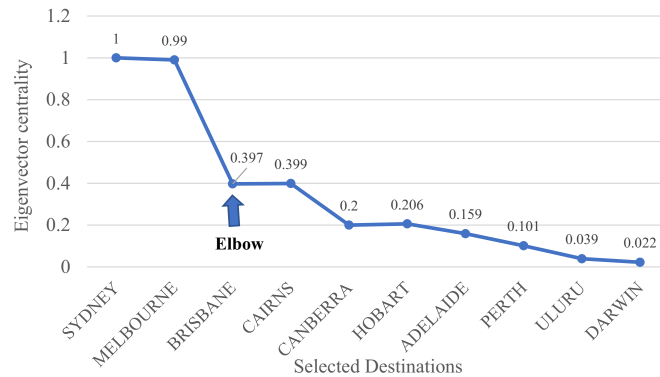


Fig. 3. Applying the Elbow method to identify the core destinations.

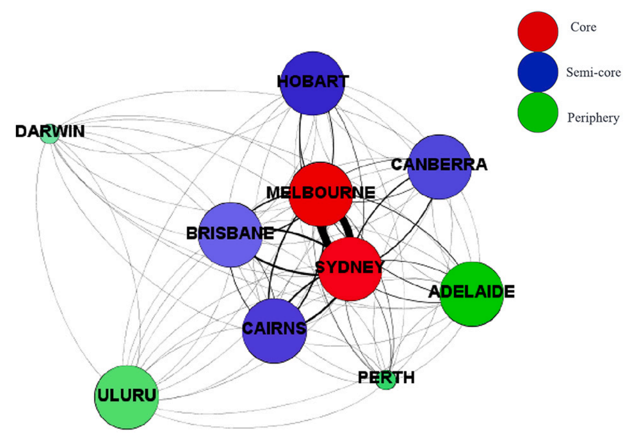


Fig. 4. Travel network among the ten selected destinations.

Table 3
The core-periphery structures.

Trip to:	C	S	P	Total
Trip Start				
C	33.75%	23.58%	6.16%	63.50%
S	22.12%	4.59%	1.60%	28.30%
P	5.97%	1.36%	0.86%	8.20%
Grand total	61.84%	29.53%	8.62%	100.00%

Tourism Research Australia identified some barriers preventing Chinese tourists from travelling to a regional destination, and one of the barriers is a lack of transport options, with little alternative to self-drive vehicles (Tourism Research Australia, 2019a). To investigate what role Australian domestic flight connections play in the core-periphery structure, the following sub-sector analysed the Australian domestic flight connections among the selected destinations and compared the flight network with the travel network.

4.1.2. *Comparing travel patterns between airline data and social media data*

The highly connected destinations could be related to the availability and transport in Australia. Many structural patterns were dictated by the availability of direct international flight connections, and some may be

Table 2
Eigenvector centrality and core-periphery destinations.

Destinations	Eigenvector centrality	Core-periphery destinations
Sydney	1.00	Core
Melbourne	0.99	Core
Brisbane	0.397	Semi-Core
Cairns	0.399	Semi-Core
Hobart	0.206	Semi-Core
Canberra	0.20	Semi-Core
Adelaide	0.159	Periphery
Perth	0.101	Periphery
Uluru	0.039	Periphery
Darwin	0.022	Periphery

related to the domestic transportation network. To test the results, data from Australia's domestic flights have been downloaded (The Department of Infrastructure, Transport, 2021) to analyse the flight network. Table 4 shows the connection of the destinations based on the flight and their role in core-periphery structure using eigenvector centrality calculated in R and the elbow method. Using the Elbow method, Sydney and Melbourne were still the core destinations based on flight connections. The mean value for the rest of the destinations is 0.27, so if the destination's eigenvector centrality is bigger than the mean, they are semi-core, otherwise periphery destinations. The core and periphery destinations from the flight network were compared with the social media travel pattern data, shown in Table 3.

The comparison shows that core destinations share the same role between Australian domestic flight connections and social media travel pattern data, indicating that flight connection is an important factor in forming core destinations in tourists' travel patterns. As for the semi-core, only Brisbane and Canberra remained in the same role as the semi-core from both flight connections and social media data. Although Perth and Adelaide had sufficient flights among the selected destinations, tourists did not often travel to them as they were labelled periphery from social media data. On the other hand, periphery destinations like Uluru and Darwin remained in the same role in both networks, but periphery destinations identified from travel itineraries showed that they were not impacted by domestic flights.

Table 5 selected four periphery destinations from the travel itineraries and compares their connections with domestic flight networks to validate the point that the domestic flight connections barely influenced periphery destinations. For example, sufficient direct flights between Darwin and Perth were found, but no travel was shown on the travel itineraries. Furthermore, Uluru's only directly connected destination was Sydney from the domestic flight, but Uluru was directly connected with all selected destinations in the social media travel patterns. The full comparison is provided in Appendix 1.

4.2. *Detecting core-periphery structure by the length of itineraries*

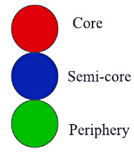
4.2.1. *Length of the itineraries*

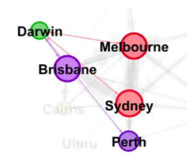
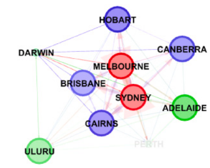
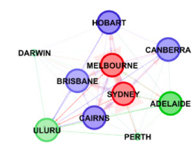
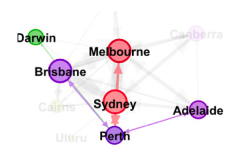
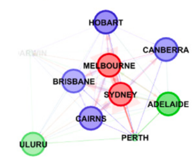
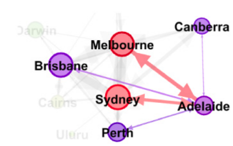
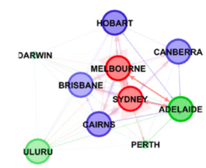
To investigate whether the length of the itineraries influences the core-periphery destination classification, tourists' itineraries were segmented based on the number of destinations and labelled as length 2, length 3 …, and length 10. The unique itineraries were also traced based on the travelling sequence (e.g. from destination A to B is different from B to A). Fig. 5 shows the number of visitors and unique itineraries in each itinerary; it can be seen that the total number of visitors captured in this analysis is 2666 (1852 + 585 + 149 + 41 + 24 + 9 + 2 + 1 + 3). Most Chinese visitors travelled to two selected destinations (length 2, with 1852 travellers). Within length 2, there were 75 unique itineraries. Travelling to three destinations has the most unique selection of destinations, with 134 different itineraries.

Table 4
Core - periphery destinations recognised by flight connections and social media travel patterns.

Destinations	Total flights among the ten destinations	Eigen centrality	Core-periphery from flight connection	Core-periphery from travel itineraries
Sydney	283,028	1	Core	Core
Melbourne	272,298	0.96	Core	Core
Brisbane	180,224	0.71	Semi-core	Semi-core
Canberra	70,698	0.35	Semi-core	Semi-core
Perth	64,782	0.28	Semi-core	Periphery
Adelaide	87,868	0.39	Semi-core	Periphery
Darwin	17,052	0.07	Periphery	Periphery
Cairns	42,882	0.20	Periphery	Semi-core
Hobart	36,588	0.18	Periphery	Semi-core
Uluru	2828	0.01	Periphery	Periphery

Table 5
Periphery destinations connections between flight networks and social media data (Periphery – green, Semi-core – purple, and Core – red colour)



Periphery destinations from domestic flights connections	Connections with the destination from direct flight connections	Connection with the destinations from social media travel itinerary
Darwin		
Uluru		
Perth		
Adelaide		

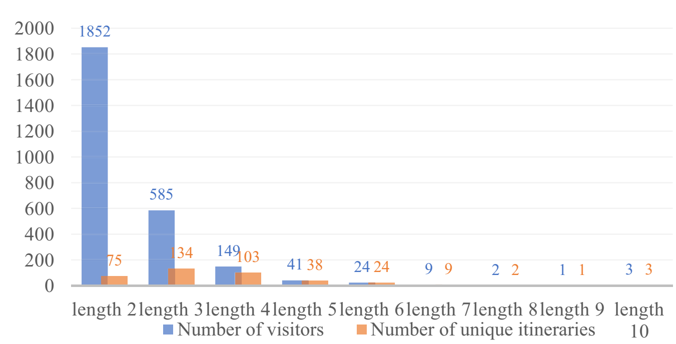


Fig. 5. The number of visitors and unique itineraries in each itinerary.

When travelling to more than five destinations, every tourist seems to have their own unique itinerary, as the number of visitors equals the number of unique itineraries. Since there were no common patterns, the following analysis, which investigated core-periphery structure based on the length of the itinerary, only considered the length of itineraries 2 to 5.

4.2.2. *Duration of stay*

According to Tourism Research Australia (2015), Free and Independent Travellers (FITs) were the dominant type of Chinese tourists in Australia. Among all types of Chinese FITs, visiting friends and relatives stay the longest, on average 45 nights (Tourism Research Australia, 2019a). Our findings are in line with this report. Since the data collection in this analysis was within a year and results showed that only a limited number of Weibo users travelled longer than 45 days so, Fig. 6 shows the duration of stay from length 2 to length 5, following the length of itineraries.

Fig. 6 shows valuable information on the relationship between the duration of stays and the length of the itineraries. Specifically, for people who travelled to two places, most people stayed one day (144), followed by two days (139 visitors) and three days (122 visitors). Nevertheless, on average, the duration of the stay for Length 2 was 8.53 days (within 45 days).

4.3. Core-periphery travel structure based on the length of itineraries

Since the travel itinerary has been transferred to a data matrix, as shown in Table 3, nine core-periphery structures have been identified based on the different lengths of itineraries, as shown in Fig. 7. Since patterns starting from the core destinations had a dominant percentage that suppressed other patterns' distribution. Fig. 7 separates into two figures (a) shows patterns starting a trip from core places, and (b) shows the rest of the patterns.

Findings showed that people mostly preferred to travel between two core destinations when their itinerary only had two destinations (71.9%), but with the increase in the itinerary length, the C - C pattern dropped the weight and stabilised around 29%. This result indicated that 29% of their itineraries went to core destinations when tourists travelled to more than two destinations. However, travelling between core and semi-core destinations (C - S) reduced the importance when the length of the itinerary increased. For example, the C - S pattern peaked in length 3 with 24.6% but slowly decreased with the itinerary length increase. Similarly, S - C peaked at length 3 with 24.7% and then gradually decreased to 20.7% in Length 5. On the other hand, the percentage of structures with C - P, S - P, and S - S increased with the length of the itinerary. Although only a small number of itineraries began from periphery destinations, all three structures increased the percentage with the length of the itinerary.

These discovered structures proved that the length of the itinerary plays an important role in increasing the opportunity for Chinese tourists to visit the periphery destinations, regardless of domestic flight connections. However, the increase in the itinerary length does not increase the importance of core destinations. These results could be great resources for tourism stakeholders to develop transportation connections or design travel itineraries and enhance destination management.

5. Discussion

5.1. Implications of the paper

Tourists' spatial and temporal movements have always been an

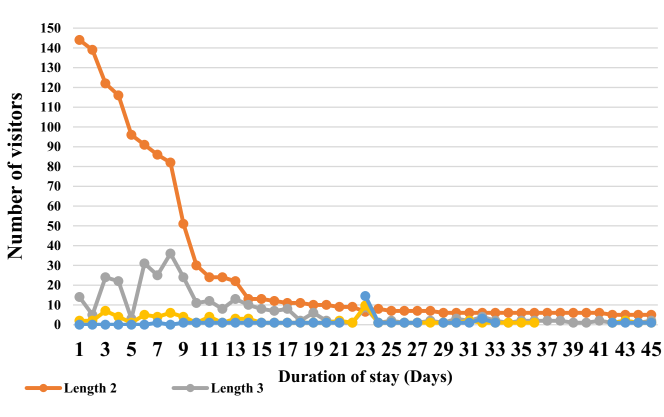


Fig. 6. Duration of stay for itinerary lengths 2 to 5.

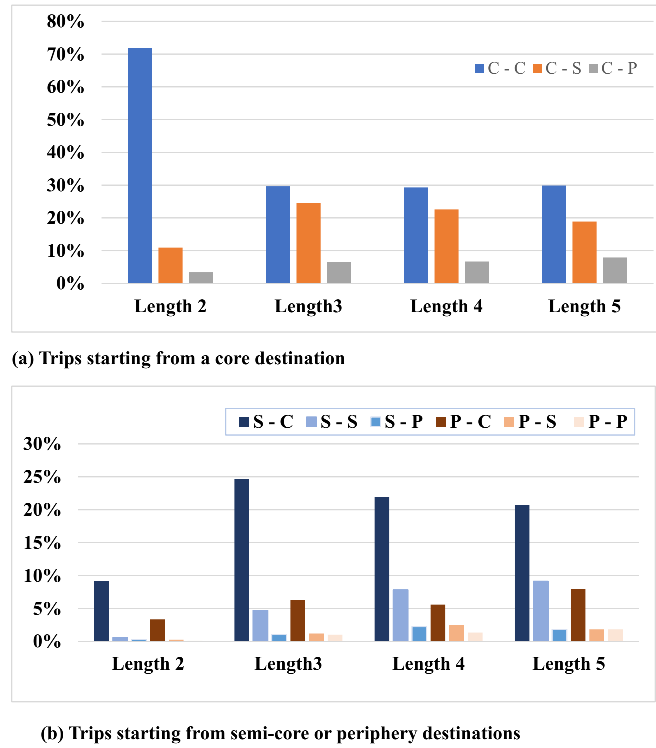


Fig. 7. The core-periphery structures for the itinerary lengths of two to five.

important perspective for understanding tourists' travel behaviour for destination management. Therefore, this work proposed to understand tourist mobility from the length of their itineraries and the duration of their stay. Furthermore, a social network analysis was applied to investigate the travel pattern structure in muti-destinations and adopted core-periphery to explain the tourist travel behaviours.

This work adopted Chinese tourists in Australia as a case study and selected ten destinations from Australia, demonstrating tourists' spatial and temporal movement patterns. The length of the itineraries was grouped by the number of places tourists visited, and their duration of stay was detected based on each itinerary. It was found that within the ten selected destinations, tourists mainly chose two of the core destinations in their trip, with an average duration of 8.53 days (within 45 days). Correlating the itinerary length and duration of stay, it was found that 2.5 to 2.9 days would increase when adding one more place to the itinerary, but the duration of the trip appeared to stay around 14 days when visiting four or more destinations. This information can help stakeholders allocate resources and better manage the trip design at the destination.

Following the spatial and temporal movements of the tourists, the potential factors that shaped such patterns were considered. In this work, we proposed correlating tourist travel patterns with domestic flight connections. It was found that core destinations (Sydney, Melbourne) and periphery destinations like Uluru and Darwin remained in the same role in tourists' itineraries and Australian domestic flight connections, indicating the importance of the gateway destinations that can receive international arrivals and departures. However, it is debatable how well the core destinations played their role in attracting tourists to more regional destinations. The results showed that tourists still prefer to travel to the two core destinations if they choose to travel to a small number of places, which confirmed the surveys from Tourism Research Australia (Tourism Research Australia, 2019a) that Sydney and Melbourne were the most popular destination. However, among all the travel patterns, the network showed no connection between Darwin and Perth, which suggests that Chinese tourists did not prefer to travel to both periphery destinations on the same trip. The reason could be that long distances between periphery destinations discourage tourists from travelling, although direct flights exist between the two destinations. Also, as the destinations offer similar attractions, tourists may not intend to visit them both on the same trip.

Analysing tourist mobility and identifying travel patterns in core-periphery structures helps destination stakeholders make data-driven recommendations for better management. Tourism Australia makes recommendations for various road trips within Australia, often focusing on particular themes (e.g. coastal or Australian Aboriginal experience) (Tourism Australia, 2022), but those recommendations are limited and not specifically targeted at Chinese travellers. It is unknown how well these suggestions fit into existing travel patterns within or beyond Australia as Chinese travellers put together increasingly complex itineraries. This research could help Tourism Australia compose evidence-based travel programs (i.e. reflecting Chinese consumer perspectives) and cover various destinations.

Destinations must have a strategy to attract international tourists from core places to more regional places. One of the indications was identifying the relationship between how long tourists would like to stay with how many places they prefer to visit. In this case, study, when planning itineraries, it is better to include at least one core destination in the itinerary and not encouraged to include more than one periphery destination in the same itinerary.

To truly elicit whether the current patterns reflect preferences, rather than transport or tour group schedules, it would be important to complement the spatial analysis presented in this research with further analysis of visitor experiences and satisfaction by using the content of social media posts. For example, considering sentiment and topics discussed in posts could serve as suitable information to complement the spatial analysis (Chen, Becken, & Stantic, 2022). The sequence of core-periphery destinations in itineraries is attractive, especially when trying to elicit regional development potential. Whilst this analysis is restricted to ten destinations, using eigenvalues to understand other structures could be extended to include more regional destinations.

5.2. Limitations of the study

There are several limitations inherent in the methodology. In terms of population, Weibo users in this sample include all people actively using the social media platform in Australia. The data does not allow for differentiation of Chinese visitors or Chinese who live in Australia, or even any other nationalities using Weibo for their communication. However, as we searched for keywords in the Chinese language, it is expected that people with a Chinese background send the posts, and according to a Weibo report, there are only 2– 3% of users from Hong Kong, Macao, or overseas, which share the confidence that the majority of posts we collected are users from mainland China (Weibo, 2017).

In order to select the users who tend to be tourists, this work processed data and only selected users who travelled to at least two selected destinations. As a result, only 2666 unique visitors satisfied this criterion. Future research could further examine sub-groups within the sample of Weibo users and increase the number of assessed visitors, but this would require more complicated methods. For example, literature mentioned that it is also possible to check a user's first and last post in the destinations, and tourists tend to post at the destination within a certain period (Chen et al., 2021b).

Another limitation is that users may not have posted from all places, or they did not mention the place whilst they were there. So, the patterns captured might be incomplete and only indicative. This omission could be biased towards destinations perceived as less important or where people stayed shorter. Furthermore, a limitation is that this study did not consider any political influences, rather than from a network science and mobility point of view to analyse tourists' mobility.

Since this paper focused on tourist mobility from the view of spatial and temporal movements, the work did not reveal why tourists travel in a specific pattern or what factors may influence the tourists' travel patterns. Future research could combine spatial and temporal analysis with content analysis to discover what would be the reason or factors that shape particular travel patterns.

6. Conclusion and future work

The study tried to understand tourist mobility by extracting the length of the itinerary and duration of stay. Network analysis further analysed travel patterns and explained them using core-periphery theory. Data were collected from the social media platform Weibo. A case study considered Chinese visitors in Australia and analysed their travel patterns in the muti-destinations of Australia.

This work proposed that when understanding tourist travel patterns in muti-destinations, destination stakeholders should consider the length of the itinerary, which significantly influences the duration of stay and other travel behaviour. For example, in this work, tourists who went to two destinations stayed an average of 8.53 days (within 45 days). When correlating the relationship between the length of itinerary and duration of stay, the patterns show that when tourists increased one destination in their itineraries, 2.5 to 2.9 days of stay would increase. Finally, a benchmark duration of stay was identified, around 14 days for tourists who wish to visit more than four destinations.

To further understand travel patterns, this paper adopted the core-periphery theory to investigate their travel structures. Proposing to use the elbow method, this paper classified the selected destinations as core, semi-core, and periphery destinations. The overall core-periphery travel patterns were associated with nine patterns. The results suggested that most tourists still focused on the core destinations (gateways) and left (travel among gateways and leave). Nevertheless, when adding one destination to the itinerary, the stay duration usually increases by two days if they travel beside the core destinations. Within the selected destinations in Australia, the core destination played a good role in connecting semi-core destinations, but not with the periphery destination.

These discovered patterns proved that the length of the itinerary plays an essential role in core-periphery structures. Furthermore, this paper investigated to what extent the length of an itinerary would influence the core-periphery structures. The results indicated that except for the high percentage (71.9%) in length two, core-to-core destinations account for around 29% of all travel patterns in any length of itinerary. However, travelling between core and semi-core destinations (C - C, C - S, S - C) showed decreasing trend with the increase of the itinerary lengths, while the pattern of (C - P, S - P, S - S) increased with the length of the itinerary. In particular, the trips started from periphery destinations (P - C, P - S, P - P) showed an increasing trend with the itinerary length.

The domestic flight connections were compared with the travel networks from social media data, and it was found that flight connection plays a significant role for core destinations but are less influential for periphery destinations.

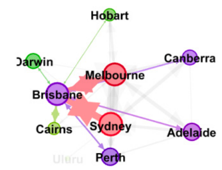
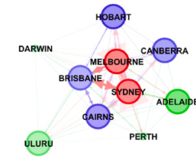
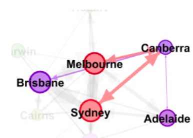
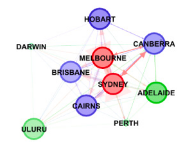
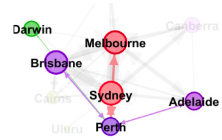
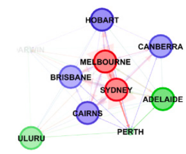
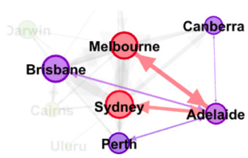
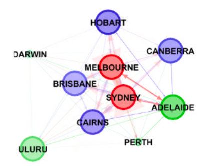
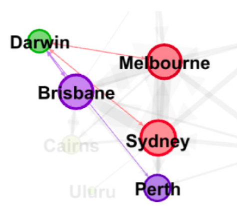
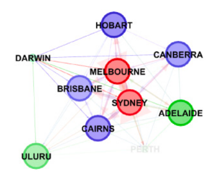
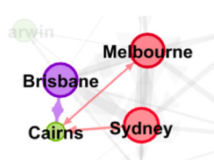
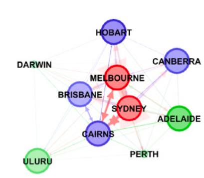
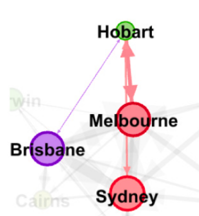
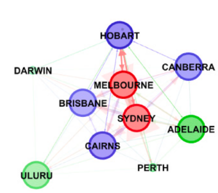
In future work, factors that influence travel patterns need further investigation. Recent research mentioned that long-term and close cooperation between public transportation and tourism companies (tour operators, railway companies, and hotels) has influenced how the trips can be organised (e.g. chained hotel, car rental) (Thao, von Arx, & Frölicher, 2020). This paper compared the travel patterns with domestic flight data, but an investigation of the relationships of visitors' travel patterns with other transportation methods, such as self-driving, could also provide some insights. Likewise, agent-based travel patterns can be analysed to see whether data mining from social media shows similar patterns. In addition, the text of the social media posts could be analysed to look into the topic content, which, together with travel patterns, can help understand the question of "why" people travel in a particular pattern.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.annale.2022.100079.

Declaration of Competing Interest

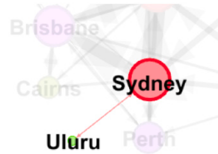
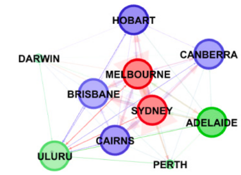
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix 1 Travel network compared between Domestic flights and social media data (semi-core and periphery destinations)

Semi-core and periphery destinations	Connection with the destinations from flight connection	Connection with the destinations from social media travel itinerary
Brisbane		
Canberra		
Perth		
Adelaide		
Darwin		
Cairns		
Hobart		
Uluru		

(continued on next page)

(*continued*)

Semi-core and periphery destinations	Connection with the destinations from flight connection	Connection with the destinations from social media travel itinerary
		

References

Asero, V., Gozzo, S., & Tomaselli, V. (2016). Building Tourism networks through tourist mobility. *Journal of Travel Research, 55*(6), 751–763. https://doi.org/10.1177/0047287515569777

Australian Bureau of Statistics. (2020). *Overseas arrivals and departures*. Australia.

Baggio, R., Scott, N., & Cooper, C. (2010). Network science. A review focused on tourism. *Annals of Tourism Research, 37*(3), 802–827. https://doi.org/10.1016/j.annals.2010.02.008

Becken, S., & Simmons, D. (2019). Stakeholder management: Different interests and different actions. In *Overtourism: Issues, Realities and Solutions* (p. 234).

Becken, S., Stantic, B., Chen, J., & Connolly, R. M. (2022). Twitter conversations reveal issue salience of aviation in the broader context of climate change. *Journal of Air Transport Management, 98*(October 2021), Article 102157. https://doi.org/10.1016/j.jairtraman.2021.102157

Bendle, L. J. (2018). The structures and flows of a large tourist itinerancy network. *Current Issues in Tourism, 21*(1), 103–122. https://doi.org/10.1080/13683500.2015.1092948

Birenboim, A., Anton-Clavé, S., Russo, A. P., & Shoval, N. (2013). Temporal activity patterns of theme Park visitors. *Tourism Geographies, 15*(4), 601–619. https://doi.org/10.1080/14616688.2012.762540

Borgatti, S. P., & Everett, M. G. (1999). Model of core / periphery structures. *Social Networks, 21*, 375–395.

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management, 21*(1), 97–116.

Cai, G., Hio, C., Bermingham, L., Lee, K., & Lee, I. (2014). Mining frequent trajectory patterns and regions-of-interest from flickr photos. In *Hawaii International Conference on System Science Mining* (pp. 1454–1463). https://doi.org/10.1109/HICSS.2014.188

Casanueva, C., Gallego, Á., & García-Sánchez, M. R. (2016). Social network analysis in tourism. *Current Issues in Tourism, 19*(12), 1190–1209. https://doi.org/10.1080/13683500.2014.990422

Chen, J., Becken, S., & Stantic, B. (2021a). Using Weibo to track global mobility of Chinese visitors. *Annals of Tourism Research, 89*(May), Article 103078. https://doi.org/10.1016/j.annals.2020.103078

Chen, J., Becken, S., & Stantic, B. (2021b). Harnessing social media to understand tourist mobility: The role of information technology and big data. *Tourism Review.*. https://doi.org/10.1108/TR-02-2021-0090

Chen, J., Becken, S., & Stantic, B. (2022). Assessing destination satisfaction by social media: An innovative approach using importance-performance analysis. *Annals of Tourism Research, 93*, Article 103371. https://doi.org/10.1016/j.annals.2022.103371

Cheng, M. (2020). *We Depend so much more on Chinese Travellers now. That Makes the Impact of this Coronavirus Novel*.

Friedmann, J. (1966). *Regional development policy: A case study of Venezuela*.

Gao, J., Peng, P., Lu, F., & Claramunt, C. (2022). A multi-scale comparison of tourism attraction networks across China. *Tourism Management, 90*(November 2021), Article 104489. https://doi.org/10.1016/j.tourman.2022.104489

Gao, Y., Ye, C., Zhong, X., Wu, L., & Liu, Y. (2019). Extracting spatial patterns of intercity tourist movements from online travel blogs. *Sustainability (Switzerland), 11*(13). https://doi.org/10.3390/su11133526

García-Palomares, J. C., Gutiérrez, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Applied Geography, 63*, 408–417. https://doi.org/10.1016/j.apgeog.2015.08.002

Gephi. (2009). *The Open Graph Viz Platform*.

Hardy, A. (2020). The ethics of tracking. In *Tracking tourists : Movement and mobility* (pp. 21–28). Goodfellow Publishers, Limited. https://doi.org/10.23912/9781911635383-4277.

Hardy, A., Birenboim, A., & Wells, M. (2020a). Using geoinformatics to assess tourist dispersal at the state level. *Annals of Tourism Research, 82*(March), Article 102903. https://doi.org/10.1016/j.annals.2020.102903

Hardy, A., Birenboim, A., & Wells, M. (2020b). Using geoinformatics to assess tourist dispersal at the state level. *Annals of Tourism Research, 82*. https://doi.org/10.1016/j.annals.2020.102903

Hardy, A., Vorobjovas-Pinta, O., Wells, M., Grimmer, L., & Grimmer, M. (2021). Measuring cruise passenger dispersal through technology. *Annals of Tourism Research, 93*(3), Article 103319. https://doi.org/10.1016/j.annals.2021.103319

Huang, X. T., & Wu, B. H. (2012). Intra-attraction tourist spatial-temporal behaviour patterns. *Tourism Geographies, 14*(4), 625–645. https://doi.org/10.1080/14616688.2012.647322

Kádár, B., & Gede, M. (2021). Tourism flows in large-scale destination systems. *Annals of Tourism Research, 87*, Article 103113. https://doi.org/10.1016/j.annals.2020.103113

Kang, S., Lee, G., Kim, J., & Park, D. (2018). Identifying the spatial structure of the tourist attraction system in South Korea using GIS and network analysis: An application of anchor-point theory. *Journal of Destination Marketing and Management, 9*(May), 358–370. https://doi.org/10.1016/j.jdmm.2018.04.001

Koo, T. T. R., Lau, P. L., & Dwyer, L. (2017). The geographic dispersal of visitors: Insights from the power law. *Journal of Travel Research, 56*(1), 108–121. https://doi.org/10.1177/0047287515625131

Lai, K., & Li, Y. (2012). Core-periphery structure of destination image. Concept, evidence and implication. *Annals of Tourism Research, 39*(3), 1359–1379. https://doi.org/10.1016/j.annals.2012.02.008

Latlong. (2019). *latlong*.

Leiper, N. (1989). Main destination ratios. Analyses of tourist flows. *Annals of Tourism Research, 16*(4), 530–541. https://doi.org/10.1016/0160-7383(89)90007-8

Leung, X., Fang, W., Bihu, W., Billy, B., & Zhuhua, X. (2012). A social network analysis of overseas tourist movement patterns in beijing: the impact of the olympic games. *International Journal of Tourism Research*, 101–113. https://doi.org/10.1002/jtr, 20 October 2011.

Lew, A., & McKercher, B. (2006). Modeling tourist movements: A local destination analysis. *Annals of Tourism Research, 33*(2), 403–423. https://doi.org/10.1016/j.annals.2005.12.002

Lewis, G. K., Hardy, A., Wells, M. P., & Kerslake, F. L. (2021). Using mobile technology to track wine tourists. *Annals of Tourism Research Empirical Insights, 2*(2), Article 100022. https://doi.org/10.1016/j.annale.2021.100022

Li, A., Mou, N., Zhang, L., Yang, T., Liu, W., & Liu, F. (2020). Tourism flow between major cities during China's National day Holiday: A social network analysis using Weibo check-in data. *IEEE Access, 8*, 225675–225691. https://doi.org/10.1109/ACCESS.2020.3044613

Li, Y., Xie, J., Gao, X., & Law, A. (2021). A method of selecting potential development regions based on GPS and social network models–from the perspective of tourist behavior. *Asia Pacific Journal of Tourism Research, 26*(2), 183–199. https://doi.org/10.1080/10941665.2018.1515092

Liu, H., Wu, L., Li, X., & (Robert).. (2019). Social media envy: How experience sharing on social networking sites drives Millennials' aspirational Tourism consumption. *Journal of Travel Research, 58*(3), 355–369. https://doi.org/10.1177/0047287518761615

Lue, C. C., Crompton, J. L., & Fesenmaier, D. R. (1993). Conceptualisation of multi-destination pleasure trips. *Annals of Tourism Research, 20*(2), 289–301. https://doi.org/10.1016/0160-7383(93)90056-9

McKercher, B. (2021). Can pent-up demand save international tourism? *Annals of Tourism Research Empirical Insights, 2*(2), Article 100020. https://doi.org/10.1016/j.annale.2021.100020

Mou, N., Yuan, R., Yang, T., Zhang, H., Tang, J., & Makkonen, T. (2020). Exploring spatio-temporal changes of city inbound tourism flow: The case of Shanghai, China. *Tourism Management, 76*(August 2019), Article 103955. https://doi.org/10.1016/j.tourman.2019.103955

Oppermann, M. (1995). A model of travel itineraries. *Journal of Travel Research, 33*(4), 57–61. https://doi.org/10.1177/004728759503300409

Pan, G. W., & Laws, E. (2003). Tourism development of Australia as a sustained preferred destination for Chinese tourists. *Asia Pacific Journal of Tourism Research, 8*(1), 37–47. https://doi.org/10.1080/10941660308725454

Park, S., & Zhong, R. R. (2021). Pattern recognition of travel mobility in a City destination: Application of network motif analytics. *Journal of Travel Research.*. https://doi.org/10.1177/00472875211024739

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management, 57*, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006

Raun, J., Shoval, N., & Tiru, M. (2020). Gateways for intra-national tourism flows: Measured using two types of tracking technologies. *International Journal of Tourism Cities, 6*(2), 261–278. https://doi.org/10.1108/IJTC-08-2019-0123

Saluveer, E., Raun, J., Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., … Silm, S. (2020). Methodological framework for producing national tourism statistics from mobile

positioning data. *Annals of Tourism Research, 81*. https://doi.org/10.1016/j.annals.2020.102895
Scott, N., Cooper, C., & Baggio, R. (2008). Destination networks. Four Australian cases. *Annals of Tourism Research, 35*(1), 169–188. https://doi.org/10.1016/j.annals.2007.07.004
Shoval, N. (2010). Monitoring and managing visitors flows in destinations using aggregative GPS data. In *, 2010. Information and Communication Technologies in Tourism* (pp. 171–183). https://doi.org/10.1007/978-3-211-99407-8_15
Shoval, N. (2018). Sensing tourists: Geoinformatics and the future of tourism geography research. *Tourism Geographies, 20*(5), 910–912. https://doi.org/10.1080/14616688.2018.1437768
Shoval, N., & Isaacson, M. (2007). Tracking tourists in the digital age. *Annals of Tourism Research, 34*(1), 141–159. https://doi.org/10.1016/j.annals.2006.07.007
Shoval, N., Kahani, A., De Cantis, S., & Ferrante, M. (2020). Impact of incentives on tourist activity in space-time. *Annals of Tourism Research, 80*. https://doi.org/10.1016/j.annals.2019.102846
Stantic, B., & Pokorný, J. (2014). Opportunities in big data management and processing. *Frontiers in Artificial Intelligence and Applications, Databases, 15–26*. https://doi.org/10.3233/978-1-61499-458-9-15
Su, L., Stepchenkova, S., & Dai, X. (2020). The core-periphery image of South Korea on the Chinese tourist market in the times of conflict over THAAD. *Journal of Destination Marketing and Management, 17*(June), Article 100457. https://doi.org/10.1016/j.jdmm.2020.100457
Thao, V. T., von Arx, W., & Frölicher, J. (2020). Swiss cooperation in the travel and Tourism sector: Long-term relationships and superior performance. *Journal of Travel Research, 59*(6), 1044–1060. https://doi.org/10.1177/0047287519871105
The Department of Infrastructure, Transport, R. D. and C. (2021). *Airline On Time Performance Annual Reports*.
Tourism Australia. (2022). *Trips and Itineraries*. https://www.australia.com/zh-hk/trips-and-itineraries.html.
Tourism Australia. (2020a). Coronavirus information and Recources for Australia's Tourism industry. https://www.tourism.australia.com/en/events-and-tools/industry-resources/resources-for-industry/coronavirus-information-and-resources-for-industry.html.
Tourism Australia. (2020b). *High value traveller fact sheet* (pp. 1–6). https://www.tourism.australia.com/content/dam/assets/document/1/7/5/e/w/2013224.pdf.
Tourism Research Australia. (2014). Chinese Visitor Satisfaction. https://www.tra.gov.au/ArticleDocuments/185/Chinese_Satisfaction_Survey_FULL_REPORT_FINAL_24JAN2014.pdf.aspx?Embed=Y.
Tourism Research Australia. (2015). Dispersal of Chinese free and Independet leisure VIsitors in Australia. March 2015 https://www.tra.gov.au/ArticleDocuments/185/Dispersal_of_Chinese_Free_and_Independent_Leisure_Visitors_in_Australia_FINAL_17032015.pdf.aspx.
Tourism Research Australia. (2019a). *Chinese Free and Independent Travellers – Their potential for regional Australia* (pp. 0–73). https://www.tra.gov.au/data-and-research/chinese-free-and-independent-travellers-their-potential-for-regional-australia.
Tourism Research Australia. (2019b). *International Visitors in Australia Year ending December 2019*. Issue December.
Wang, Y., Li, X., & Lai, K. (2018). A meeting of the minds: exploring the core–periphery structure and retrieval paths of destination image using social network analysis. *Journal of Travel Research, 57*(5), 612–626. https://doi.org/10.1177/0047287517706262
Wang, Z., Liu, Q., Xu, J., & Fujiki, Y. (2020). Evolution characteristics of the spatial network structure of tourism efficiency in China: A province-level analysis. *Journal of Destination Marketing and Management, 18*(October), Article 100509. https://doi.org/10.1016/j.jdmm.2020.100509
Weaver, D., Moyle, B. D., McLennan, C., & lee.. (2021). A core/periphery perspective on mega-event sustainability: Dystopic and utopic scenarios. *Tourism Management, 86* (January), Article 104340. https://doi.org/10.1016/j.tourman.2021.104340
Weaver, D. B. (2017). Core–periphery relationships and the sustainability paradox of small island tourism. *Tourism Recreation Research, 42*(1), 11–21. https://doi.org/10.1080/02508281.2016.1228559
Weibo. (2012). Weibo API. https://open.weibo.com/wiki/2/search/topics.
Weibo. (2017). *2016 Weibo User Development Report*.
World Tourism Organization (UNWTO). (2019). 'Overtourism'? – understanding and managing urban tourism growth beyond perceptions, volume 2: case studies. In *, Vol. 2. Overtourism'? Understanding and Managing Urban Tourism Growth beyond Perceptions Volume 2: Case Studies*. https://doi.org/10.18111/9789284420643
Yang, Y., Li, D., & Li, X. (2017). Public transport connectivity and intercity tourist flows. *Journal of Travel Research, 004728751774199*. https://doi.org/10.1177/0047287517741997
Zeng, B. (2018). Pattern of Chinese tourist flows in Japan: A social network analysis perspective. *Tourism Geographies, 20*(5), 810–832. https://doi.org/10.1080/14616688.2018.1496470

Jinyan Chen* is a Research Assistant Professor at Hong Kong Polytechnic University. Her research interests include analysing tourists' travel patterns, sentiment, and network analysis.

Susanne Becken is a Professor of Sustainable Tourism Griffith Institute for Tourism, and her research interest is related to sustainable tourism.

Bela Stantic is Professor and director of the Big Data and Smart Analytics Lab at Griffith University, with expertise in big data analytics.