Landscape and Urban Planning 204 (2020) 103934

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Co-visitation network in tourism-driven *peri*-urban area based on social media analytics: A case study in Shenzhen, China

Yao Sun[a], Yiwen Shao[b,*], Edwin H.W. Chan[c]

[a] *Shenzhen Tourism College, Jinan University, Nanshan District, Shenzhen, 518053 Guangdong Province, China*
[b] *School of Architecture and Urban Planning, Shenzhen University, Nanshan District, Shenzhen, 518061 Guangdong Province, China*
[c] *Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong, China*

ARTICLE INFO

Keywords:
Social media analytics
Co-visitation network
Social network analysis
Tourism-driven urbanization
Peri-urban area

ABSTRACT

The *peri*-urban area has undergone tremendous changes in various aspects. Tourism-driven development in this area, echoing the longing of urban residents for recreational and leisure space, provides an alternative to conventional urbanization approaches. Particularly, in view of heightened conflicts and fierce competitions that often occur in the area, developing the knowledge of the co-visitation network, or how different attractions and villages interact with each other from the visitors' perspective, is key to more informed planning strategies. With the rapid development of online social media platforms and mobile devices, researchers can collect massive real-time data on the basis of online user-generated content. While there are already many attempts on its application on tourism design, few studies to date have explored how *peri*-urban attractions and villages are organized in same trips. In this paper, we design a new methodology on the basis of social media analytics to construct and map a co-visitation network, and apply NetMiner-enabled social network analysis to explain its structure and nodal functions. The usefulness of this methodology is demonstrated through a detailed case study in Shenzhen, China. Research findings reveal types of existing co-visitation patterns, which help planners predict future trends, design new experiences and take specific measures in line with visitor demands.

1. Introduction

The *peri*-urban area, especially that of developing countries, is characterized by a mixture of urban and rural land uses and fragmented landscapes, resulting from fluid, dynamic and interconnected socio-economic process (Serra, Saurì & Salvati, 2017; Zlender & Thompson, 2017; Duvernoya, Zambonb, Saterianoc & Salvatid, 2018; Tian & Guo, 2019). While many natural and cultural landscapes in the *peri*-urban area are cleared away to make room for urban expansion, still a considerable number of landscapes, including country parks, mountains, seashores, and cultural heritages, are retained and transformed into scenic attractions for urban residents (Sanesi, Colangelo, Lafortezza, Calvo & Davies, 2017). Providing recreational pleasure and enhancing physical and mental health (Teixeira, Marques, Garabini, Cezar & Pereira, 2018; Kaltenborn, Kaltenborn & Linnell, 2019), these attractions, with high accessibility and distinct characteristics, have become increasingly popular short-distance tourism destinations (Komossaa, Zanden & Verburg, 2019). They are indispensable in maintaining tourism vitality through offering landscape services in both natural and human-modified habitats (Nahuelhual, Carmona, Lozada, Jaramillo & Aguayo, 2013; Bastian, Grunewald, Syrbe, Walz & Wende, 2014). Under this background, the trend of a new urbanization phenomenon, known as tourism-driven urbanization as opposed to urban sprawl and industry-driven urbanization, has been increasingly salient in the *peri*-urban area of many cities (Zlender & Thompson, 2017; Li, 2020).

Along with the development of scenic attractions in the *peri*-urban area, many adjacent *peri*-urban villages start to take advantage of the spillover effect of attractions to develop their economy and improve the livelihood of their villagers. While most of these villages are not endowed with enough attractive resources, they attempt to provide such supplementary and auxiliary services as parking, catering and lodging. It is through this manner that *peri*-urban villages are gradually engaged into the *peri*-urban recreation system, where exogenous factors outweigh endogenous pursuits. There are empirical studies focusing on the transitions in local industrial structure, employment profile and physical environment of these villages (Goossen & Langers, 2000; Guan, Gao & Zhang, 2019; Matatolu, 2019). However, more villages in this area are in a relatively disadvantaged and marginalized situation. They are either overlooked by relevant government policies and plans, or under threat of urban construction and development projects. Their fate

* Corresponding author.
E-mail addresses: sunyao@sz.jnu.edu.cn (Y. Sun), yiwenshao@szu.edu.cn (Y. Shao), bsedchan@polyu.edu.hk (E.H.W. Chan).

https://doi.org/10.1016/j.landurbplan.2020.103934
Received 30 January 2020; Received in revised form 25 May 2020; Accepted 26 August 2020
Available online 08 September 2020

are uncertain under conflicted interests around such controversial issues as cultural inheritance, social cohesion and ecosystem wellbeing (Leaf, 2002; Guo, Xiao & Yuan, 2017). The inclusive consideration of the entire area, together with attractions and villages, under the umbrella of tourism-driven urbanization is often suggested as a win–win solution.

Compared with conventional urbanization modes, tourism-driven urbanization, aimed at building co-created value between villages and attractions (Wilde & Cox, 2008; Kylänen & Rusko, 2011; Tolstad, 2014; Lorgnier, Su, Woratschek, Horbel & Popp, 2014), proves to be more cost-effective to assist transformation of *peri*-urban villages in this post-productivism era (Lang, Chen & Li, 2016). It is also agreed that the practice of such an urbanization mode demands a deeper understanding of the entire system and the interaction and relationship between its elements, both attractions and villages. This is because changes in any one part of a fully-fledged complex adaptive system would lead to changes in other parts and eventually exert influence on the stability of the whole supply chain (Yang, Yin, Xu & Lin, 2019).

In this paper, we propose and discuss on the concept of co-visitation as the proxy of the interplay between an attraction and a village, between attractions and between villages. Promoting co-visitation not only benefits the development of individual villages, but also contributes to tourism-driven *peri*-urban area at large. Taking a stand of visitors' demand, the value of co-visitation linkages is understood as the possibility that two locations, village or attraction, are organized in same trips by visitors. Although there exists qualitative research giving a picture of the interplays between different elements, the quantification and analysis of co-visitation in the context of tourism-driven *peri*-urban area is much of a research void which awaits more scholastic attention. A major cause of this lack of research is the tremendous difficulties associated with the collection of data. Fortunately alongside the development of technology-supported networks and accordingly the social media data analytics in the field of tourism design, we are more capable of conducting quantitative research. In this paper, for example, we design a new methodology on the basis of social media analytics to construct and map the co-visitation network, and apply social network analysis to explain its structure and nodal functions. The findings can inform us of co-visitation patterns, predict future trends and design new experiences, thus assisting decision-making and helping produce better plans in tourism-driven *peri*-urban area. We use a case in Shenzhen, China, to demonstrate the usefulness of this replicable methodology.

This paper is organized as follows. After this introduction, Part 2 provides a brief review on social media data analytics in the field of tourism design and proposes the three-fold contribution of our paper. Part 3 introduces our research preparation including data source and case study area selection, and constructs an innovative three-step methodology. Part 4 reports the results of our research in the case study area, describing and interpreting the co-visitation network. Part 5 further discusses how we understand co-visitation patterns in the tourism-driven *peri*-urban area and how we predict future trends and design new experiences based on the results. Part 6 concludes the entire paper through highlighting our contribution under the backdrop of tourism-driven urbanization in China's *peri*-urban areas.

2. Social media data analytics in tourism design: a brief review

The rapid development of technology-supported networks and popularization of the mobile devices have tremendously improved the tourism experience. Apart from travel search engines such as booking.com and travel company websites such as TripAdvisor.com, social media can also act as an effective platform for visitors to plan trip journeys, publicize travel logs, exchange information and communicate travelling experience (Leung, Law, Hoof & Buhalis, 2013). Many scholars have studied the wide usage of smartphones and social networking tools such as Facebook and Twitter as a new lifestyle (Xiang & Gretzel, 2010; Xiang & Fesenmaier, 2017). The result is the production of massive online user-generated content. The content, generated by numerous Internet users in the form of product reviews, blogs or messages on social media platforms, makes up an increasing share of the overall Internet information (Liu, 2011).

On the other hand, the aforementioned trend also opens a new window for interested investigators to collect big data in support of policy decision-making and thereby bring revolutionary changes to how we conduct research. As argued by Boyd and Crawford (2012), big data analytics, including social media analytics, can lead to an epistemological change in the constitution of knowledge and the process of research. The academic interests in social media data, containing volunteered geographic information, can date back to 2007 in the Specialist Meeting in Santa Barbara, California (Goodchild & Glennon, 2010). Elwood, Goodchild and Sui (2013) believe that the quality of volunteered geographic information, including social media data, can rival and even exceed that of traditional authoritative data from mapping agencies. Therefore the social media approach has the potential to impact society in a novel and important ways.

The application of big data especially social media data analytics in tourism design is emergent and increasingly popular, because successful tourism management is always reliant upon a deep understanding of ever-changing consumer behavior based on data inputs (Xiang & Fesenmaier, 2017). Tourism big data is more reliable, real-time and nowcasting, and contains more new information flows compared with conventional data (Song & Liu, 2017). It is generally agreed that the positive correlation between social media data and real-time statistics is sound and strong (Donahue, Keeler, Wood, Fisher, Hamstead, & Mcphearson, 2018; Tieskens, Zanten, Schulp, & Verburg, 2018). For example, Wood, Guerry, Silver and Lacayo (2013) find that higher density of geo-tagged social media posts is associated with higher actual visitations to a certain destination, and that higher frequency of co-occurrences in the same posts suggests closer relationship between different destinations. This matchup suggests that social media data can assist us to evaluate visitor behaviors and preferences, and accordingly monitor the performance of individual tourism destinations and their ties.

There are empirical research outputs based on social media analytics. For example, Li, Goodchild and Xu (2013) collect georeferenced data, tweets and photos from Twitter and Flickr, to explore the spatio-temporal distribution and socioeconomic characteristics of social media users across the United States. Ullah, Wan, Haidery, Khan, Ebrahimpour and Luo (2019) rely on 29,890 check-in data from Sina Weibo to analyze visitor distribution in 157 parks in Shanghai metropolitan area. Donahue et al (2018) use geotagged social media data from Flickr and Twitter to explore visitation patterns in 1581 urban parks and *peri*-urban green spaces in the Twin Cities Metropolitan Area, Minnesota. Current research using social media has moved beyond simple description to focus on predictive modeling. For instance, Stienmetz and Fesenmaier (2018) utilize geotagged photos from Flickr to build networks of describing visitor flow structures between 29 counties of Florida and match these networks with official visitor-related spending data to predict the economic impact of various visitation patterns.

In this paper, we make some progress to tailor social media analytics to the needs of studying tourism-driven urbanization in the *peri*-urban area. This is achieved mainly in the following three aspects. First, instead of focusing on the visitation of individual locations, such as urban parks, this study concentrates on the concept of co-visitation and extends the application of social media analytics to explore the interaction between attractions and villages (Tieskens et al., 2018; Donahue et al., 2018). Second, visitor-tracking data from social media platforms are usually performed in regional, national, or even global scales with relatively low spatio-temporal precision (Shoval & Ahas, 2016). We conduct research on a much smaller scale, the *peri*-urban area of a metropolis, through incorporating different tools into a methodology. That being said, the transformations that are taken place behind this specific investigation are not only far-reaching in this area, but also significant to reflect a nationwide trend. Third, we pay special attention to those locations with low online exposure, where solely relying on social media data is incapable of reflecting the complete co-visitation network. Accordingly, we incorporate online questionnaire investigation in our methodology in order to combine social media data with interviewing the users who created the data. The introduction of questionnaires opens up the possibility to fill in gaps that usually limit the scope of social media research.

3. Materials and methods

3.1. Finding a niche in research

3.1.1. Justifying data source selection

China has the largest population of netizens in the world. According to Chinese Academy of Cyberspace Studies (2020), there are as many as 904 million web users by March 2020, accounting for around 60% of the total population. Exchanging tourism information and recreation experiences are heated topics in the online community. On the one hand, online tourism has been a rapidly expanding industry and online tourism agencies, exemplified by Ctrip (ctrip.com), Mafengwo (mafengwo.cn) and Feizhu (fliggy.com), are increasingly popular on destination marketing and trip promotion. Ctrip, for instance, has over 135 million transaction users in 2018 and offers various services including reservation of accommodations, ticketing of transportation and arrangement of packaged tours (Zhang, 2019). Although these online platforms open a channel for their customers to comment on travelling experiences, these communications are limited to specific destinations which are the promotion partners of these agencies. Therefore, data from any of these agencies can hardly cover a full spectrum of destinations in a region.

On the other hand, social media offer an important platform on tourism-related information sharing. Among various social media services, Sina Weibo and Tencent Moments stand out. Sina Weibo is the top mass microblogging platform open to the public in China, containing huge volume of reviews generated by actual tourists. Tencent Moments is a built-in function of WeChat, the top short message service of China with over 1.15 billion monthly active users (Tencent Company, 2020). However, considering the fact that Tencent WeChat is a tool designed only for acquaintances applying certain user privacy protection protocols, it is quite unlikely to retrieve Moments data.

Our study selects Sina Weibo as the data source due to its wide coverage, openness to the public and rich information of visitor behaviors. First, Sina Weibo had 462 million monthly active users in 2018 and more than 93% of log-ins were achieved through mobile devices with geo-location function (Sina Weibo Data Center, 2019). Second, a majority of information and interaction between users are open to the public. Except for a small portion of private information, Sina Company has opened its microblog data via the Application Programming Interface. The open platform also makes it possible for researchers to conduct in-depth online investigations and collect supplementary information. By joining the Sina Webo community, users agree on the privacy protocols, therefore the consent from research subjects is guaranteed. Third, it is an ideal platform of communicating travelling experiences and commenting on tourism services, because people, especially the younger generations, tend to post microblogs related to their visits.

3.1.2. Profiling the case study area

This paper focuses on Dapeng Peninsula, a typical *peri*-urban area of Shenzhen, China (Fig. 1). The Peninsula is investigated for two major reasons. First, Shenzhen, commonly recognized as one of the four first-tier cities, is a fast growing metropolis confronted with mounting population and increasingly severe land shortage. The *peri*-urban area of Shenzhen is under the impact of urban sprawl with acute conflicts in economy, land use and governance (Sun and Shao, 2020). Second, different from the demolishing and rebuilding mode that is prevalent in many *peri*-urban areas of China, Dapeng is devoted to promoting a tourism-driven urbanization approach. Studying the co-visitation network is conducive to better evaluating the performance of this approach and making future plans.

The Peninsula has a population of roughly 200 thousand and consists of three towns, Kuichong in the north, Dapeng in the middle and Nan'ao in the south. It enjoys a mild subtropical marine climate and is endowed with rich cultural and natural landscapes. Apart from three town centers, the entire territory of the Peninsula is dotted with 26 nature-based attractions and 6 culture-based attractions as well as 42 *peri*-urbant villages (Fig. 2). Nature-based attractions include beaches, mountains, wetlands and reefs, and to a great extent keep original wilderness. The Peninsula is home to the finest beaches in the Pearl River Delta and the lush mountains are famed for hiking trails. Culture-based attractions include historic heritages, Buddhist temples and modern industrial landscapes. In particular, Dapeng Fort is the former military encampment and barracks in Qing Dynasty and designated as a national historic heritage. These attractions make the Peninsula an ideal tourism and recreation destination for urban residents.

Peri-urban villages on the Peninsula are mainly located near scenic attractions and enjoy the spillover from these attractions. These villages generally remain rustic landscapes, characterized by vernacular dwellings built by the Hakka, an ethnic sub-group of the Han Chinese. Traditional Hakka villages on the Peninsula feature one-story and two-story houses with white walls, small openings and black gable roofs. Houses are arranged around the ancestral halls of lineage and the general layout strictly follows *Fengshui* principles. Instead of traditional agriculture and fishing, the economy of villages is more reliant on tourism-related activities.

3.2. Introducing a three-step methodology

This study integrates different social media-based methods to describe and interpret co-visitation network of different locations (attractions and villages) in the case study area. The methodology consists of the following three steps (Fig. 3). Step 1 is to mine and process social media data. The processing of social media data involves two methods. First, a web crawler is used to mine a massive volume of Weibo posts conveying co-visitation information. With the help of data-cleaning tools such as Chinese semantic recognition, submission location recognition and visitor behavior recognition, we manage to select out those posts pertinent to targeted locations. Second, to explore the co-visitation information of those locations with low online exposure, we supplement with manually conducted online questionnaire surveys via Sina chatting sessions. Step 2 is to calculate the values of co-visitation linkages between pairs of locations considering different scenarios and constructing different formulae. On the basis of calculation results, we map the co-visitation network. Step 3 is to conduct social network analysis applying the NetMiner software. This includes 1) network structure analysis, which reveals underlying structure of the co-visitation network, 2) degree centrality and betweenness centrality analysis, which identifies the roles of different locations in the network.

3.3. Step 1: Mining and processing social media data

3.3.1. Co-visitation information from valid Weibo posts

Through connecting the web crawler with the Application Programming Interface on *open.weibo.com*, we managed to collect 2.4 million Weibo posts. Afterwards, we conducted data cleaning to delete noisy data and retain those valid datasets closely linked to the Peninsula.

To conduct data cleaning, we first applied the Python-based Chinese semantic recognition toolkit to select out those posts containing all the possible formal and informal names of 74 locations, which yielded 12,165 posts. In doing so, we paid special attention to those villages and

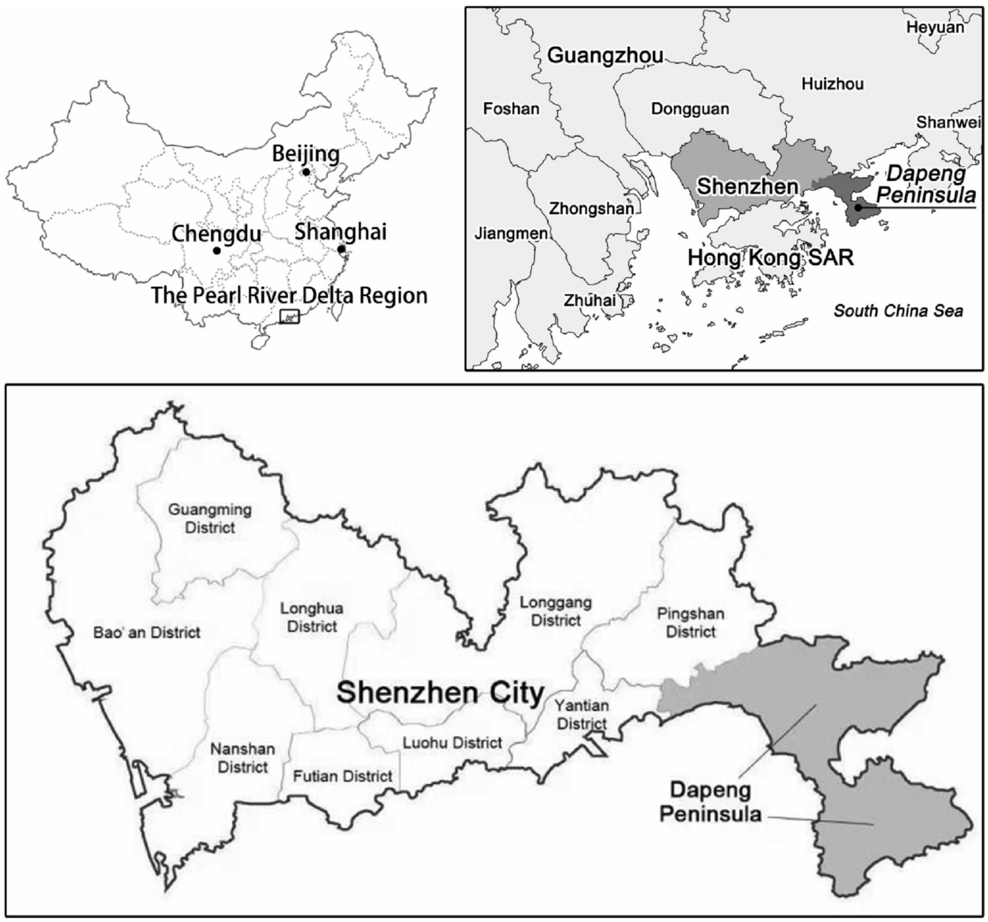

Fig. 1. Location of the case study area, Dapeng Peninsula. The upper left inset shows the location of the Pearl River Delta, one of the major economic engines in China. The upper right inset shows the Peninsula in the context of the Delta, and in relation to major cities such as Hong Kong and Guangzhou. The lower inset is the zoomed-in map of Shenzhen City, showing that the Peninsula constitutes the *peri*-urban area of the City.

attractions whose names could be easily mistaken. Next, we ran a Python-based location recognition toolkit and deleted 2574 invalid posts on locations that are with overlapping names but are actually not located on the Peninsula. We manually read the rest posts piece by piece and deleted 1562 items. This was to ensure that all remaining posts were submitted by real visitors. Following aforementioned procedure, we ended up with 8029 valid posts dated between January 2010 and January 2018. These posts were further grouped according to the locations they described. Generally, a complete Weibo post contains a series of useful information, including user's identity, submission time and location, main text and the numbers of forwards, likes, and comments.

Next, we extracted the posts which contained co-visitation information from those valid Weibo posts, and ended up with 1202 reliable pieces. In this step, two methods were used. First, with the help of Chinese semantic recognition, we selected out those posts containing two or more names of 74 locations. Second, we arranged Weibo posts submitted by same users in the chronological order in accordance with the release time. In this way, changes of release locations were a fair reflection of the user's visits in one trip. For example, if a traveler checks in his geographical location in Location A on a certain date, Location B in the next day, and Location C three days after, we can keep good track of his visits in the trip (A → B → C).

3.3.2. Co-visitation information from questionnaire data

In attaching the mined georeferenced Weibo posts to the places they describe, we found that those villages with low online exposure were not well articulated. This is because visitors tend to replace names of less known villages with names of well-known adjacent scenic attractions when they submit Weibo posts. As a result, co-visitation information from Section 3.3.1 could possibly overlook the role of some villages, while those villages are often quite important during trips through offering necessary supplementary services. To address this issue, we introduced online questionnaire investigation via Sina chatting sessions in 2018. Participants of questionnaire surveys were selected randomly according to the user's identity of valid Weibo posts. A total of 700 online questionnaires were disseminated, 382 of which were returned with ethical approval obtained. Questions on the profile of visitors and their trips were asked. We collected information on visitor age range, place of origin, and length of trip, and invited respondents to report their traveling routes through ticking those visited attractions and one or more co-visited villages, and connecting them with straight lines.

Through the questionnaire data, we can calculate the ratio of visitors of a particular village to the total visitors of the attraction, where this village is attached. This ratio (P_1) is an indicator of the co-visitation linkage between this village and this attraction. In the same vein, co-visitation linkage between two villages can be indicated by the ratio (P_2) of tourists of both villages to the total visitors of the attraction, where these villages are attached.

3.4. Step 2: Calculating values of co-visitation linkages

The calculation of co-visitation linkages considers the following two different scenarios and applies corresponding formulae. Before making choices on formulae, it is necessary to differentiate locations according to their online exposure. First, for any pair of locations with high online exposure, the value of their co-visitation linkage equals the total number of times that they are organized in the same trips, according to the co-visitation information from Section 3.3.1. Its calculation is presented as Formula (1), where C_a is the value of co-visitation linkage and H_a represents the total number of times that two locations are

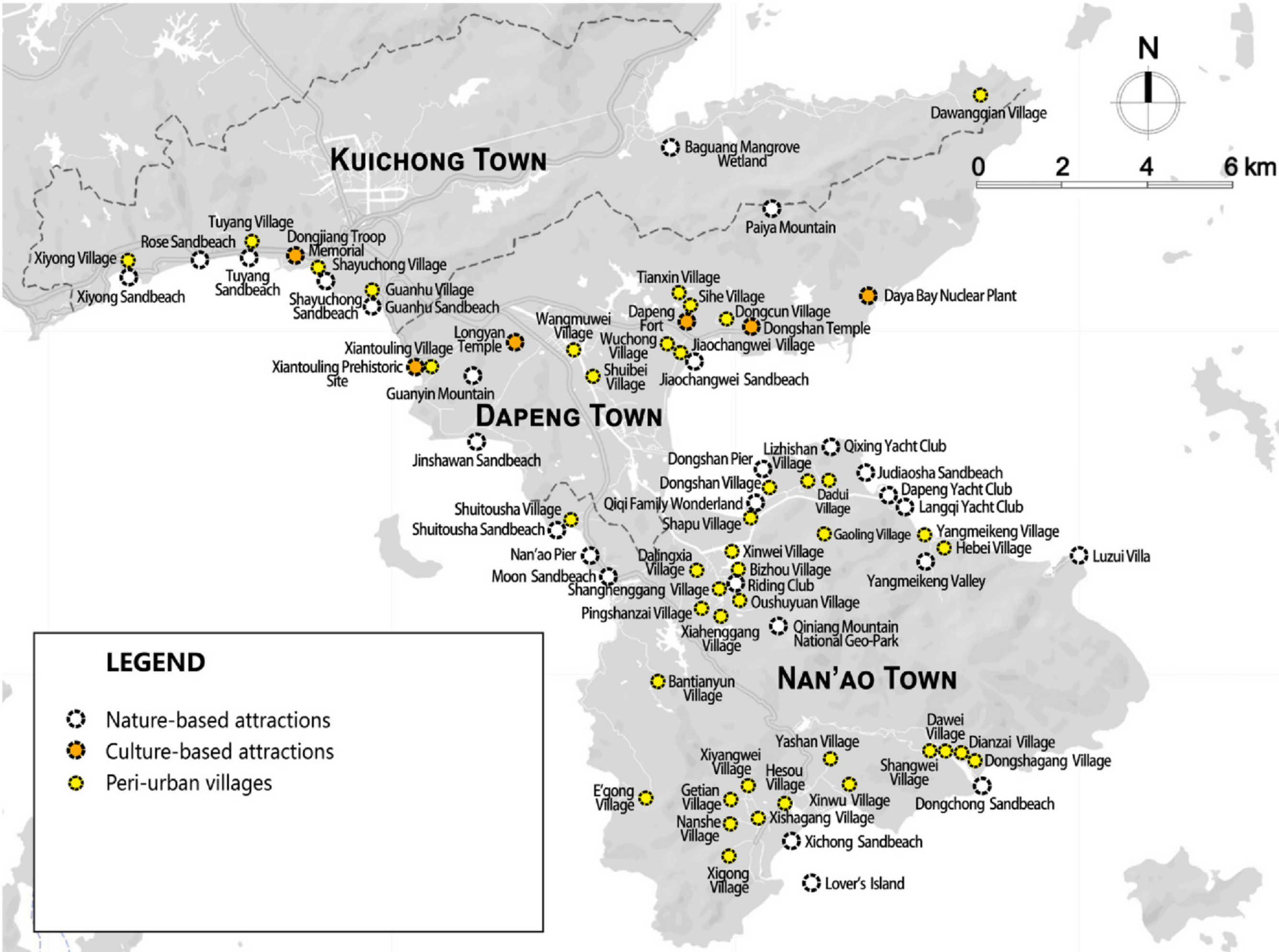

Fig. 2. The spatial distribution of scenic attractions and *peri*-urban villages of Dapeng Peninsula. There are 26 nature-based attractions, 6 culture-based attractions and 42 *peri*-urban villages, totaling 74 targeted locations.

incorporated in the same trips.

$$C_a = H_a \tag{1}$$

Second, the value of co-visitation linkage (C_{b1}) between a village with low online exposure and an attraction is presented as Formula (2), where $\sum_i^n H_b$ represents the total number of times that an attraction is

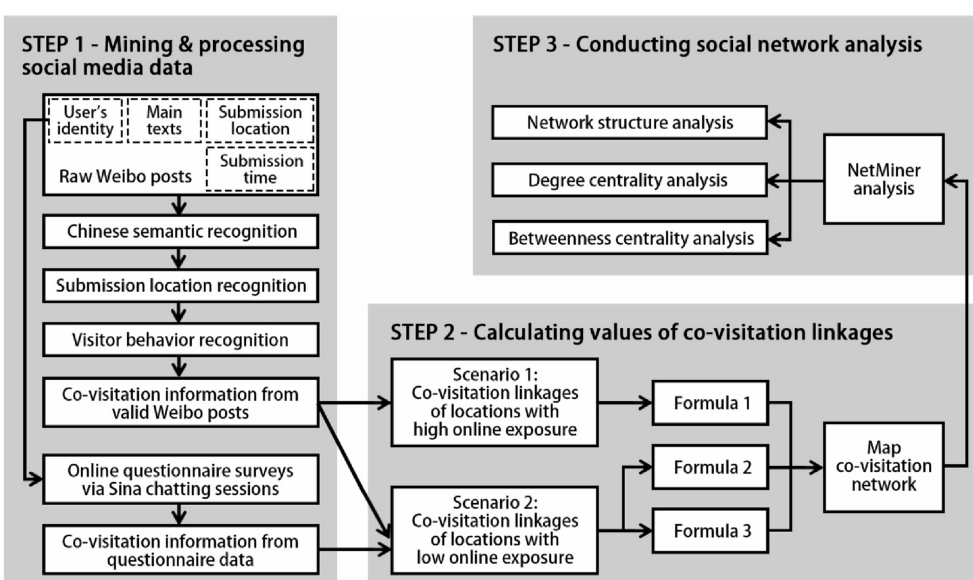

Fig. 3. The three-step methodology.

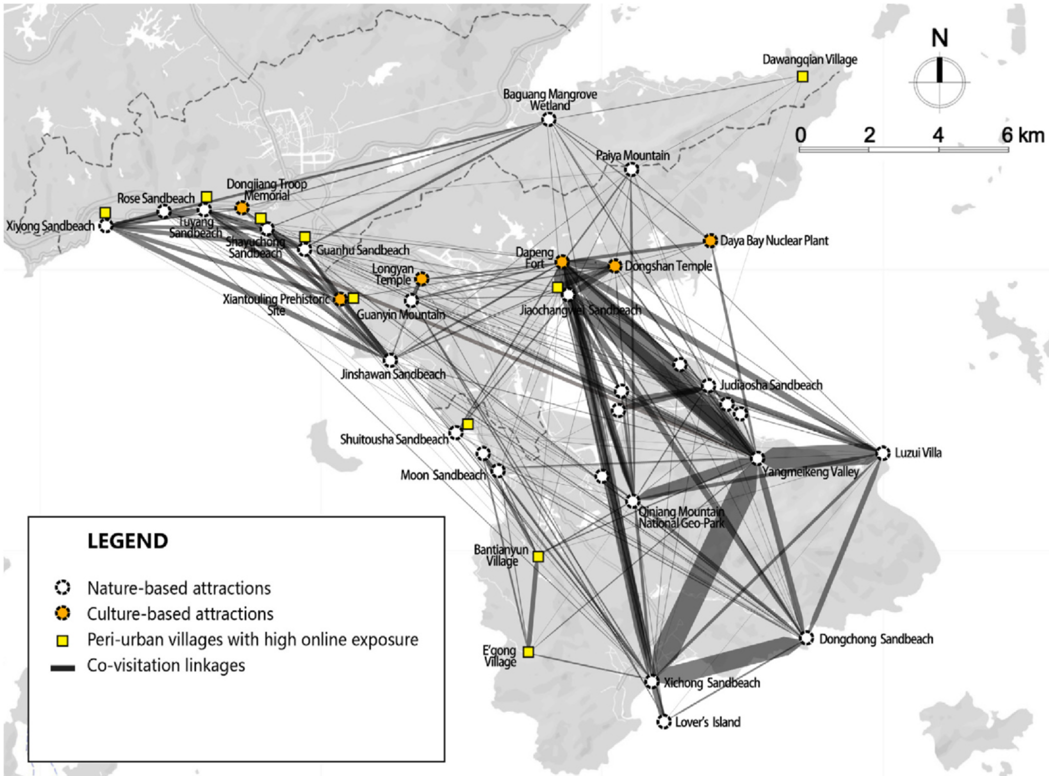

Fig. 4. The co-visitation network between 42 locations with high online exposure, including 26 nature-based attractions, 6 culture-based attractions and 10 *peri*-urban villages. The width of gray lines is proportional to the strength of corresponding co-visitation linkages according to calculation from Formula (1), wherein a pair of locations connected with thicker lines are more likely to be co-visited and vice versa.

involved in the trips according to the co-visitation information from Section 3.3.1, and the aforementioned P_1.

$$C_{b1} = \sum_{i}^{n} H_b \times P_1 \tag{2}$$

Likewise, the value of co-visitation linkage (C_{b2}) between two villages that are attached to the same attraction can be calculated by Formula (3).

$$C_{b2} = \sum_{i}^{n} H_b \times P_2 \tag{3}$$

3.5. Step 3: Conducting NetMiner-enabled social network analysis

A network consists of a set of nodes and the interconnected ties among these nodes. In our study, a co-visitation network is established with different locations as nodes and co-visitation linkages as ties. NetMiner, a software that synthesizes various algorithms based on the network theory (Furht, 2010; Newman, 2010), is applied here to conduct social network analysis on the co-visitation network. Specifically, it helps reveal the hidden features of networks and detect potential patterns and structures of networks. The software has been widely used in different disciplines, including urban planning (Yuan, Chen, Li, Ji, Wang, & Skibniewski, 2018) and tourism studies (Kim & Kim, 2017). In this context, two specific questions are explored. First, *what is the underlying structure of the co-visitation network*? Second, *what roles do different locations play in this co-visitation network*?

In response to the first question, we conduct network structure analysis, proposed by Kamada and Kawai (1989). This analysis lays out locations with stronger co-visitation linkages in the same cluster. That is to say, locations in the same cluster are more likely to be co-visited in the same trip. It should be noted that if all co-visitation linkage values of a certain location fall below the average, this location is no longer involved in the layout.

In response to the second question, we introduce degree centrality analysis and betweenness centrality analysis. A location of higher degree centrality has a larger number of linkages incident upon it, therefore it tends to play more important role in maintaining the structural integrity of the network (Sharma & Surolia, 2013). Through establishing co-visitation linkages with more locations, these locations demonstrate a higher degree of activeness and sociability. On the other hand, betweenness centrality identifies those nodes who play a bridge spanning role in a network and can be obtained by the number of times a node appears in the shortest path between other node pairs (Nooy, Mrvar, & Batagelj, 2011). In our context, a location of higher betweenness centrality is more important in facilitating transfers and promoting information diffusion between other locations. It should also be noted that while all nodes in a network have degree centrality, only a part of nodes show betweenness centrality.

4. Results

4.1. Profile of visitors and trips

Through analyzing results of the online questionnaire surveys, we can portray a profile of visitors. In terms of age structure, the Peninsula is highly popular among young generations. Visitors aged between 21 and 40 constitute the majority (71.7%) and those aged not more than 20 take up 10.5%. In contrast, senior people (over 60) accounts for 2.1%. In terms of place of origin, 76.7% of respondents come from other parts of Shenzhen and 18.1% of respondents come from adjacent Pearl River Delta cities, while only a minor 5.2% of visitors come from cities outside the Delta.

Questionnaire investigation also helps us get a picture of visitation pattern. In terms of length of trip, 38.2% of visitors take day trips and

57.6% of visitors choose overnight trips, while only 4.2% of visitors stay even longer. In terms of number of visited scenic attractions per trip, 16.5% of visitors visit a single attraction, 59.2% visit two scenic attractions, and 24.3% visit three or more attractions. It is clear that multi-destination trips take up a dominating majority.

4.2. Describing the co-visitation network

The 74 locations on the Peninsula can be divided into two groups according to their online exposure. The first group includes all 32 attractions and 10 villages with high online exposure, upon which we apply Formula (1). The second group includes the rest 32 villages with low online exposure, upon which we use Formula (2) and Formula (3). Co-visitation linkages are computed and mapped, as demonstrated below.

4.2.1. Co-visitation network of locations with high online exposure

Fig. 4 shows the co-visitation network of locations with high online exposure, wherein grey lines represent the co-visitation linkages between two locations and the width of these lines is proportional to the values of corresponding co-visitation linkages. Taking two typical nature-based attractions, Yangmeikeng Valley and Xichong Sandbeach, as an example, the total number of times that they are organized in same trips (H_a) is 176. According to Formula (1), the co-visitation linkage between them (C_a) is 176.

4.2.2. Complete co-visitation network considering locations with low online exposure

On the basis of Fig. 4, we add on the calculation results from Formula (2) and Formula (3) considering the rest 32 villages with low online exposure, and develop a complete co-visitation network, as shown in Fig. 5. Here we take Dawei Village, which is attached to Dongchong Sandbeach, as an example. We find through analyzing Weibo posts that the total number of trips relating to Dongchong Sandbeach is 173 ($\sum_i^n H_b$). Questionnaire investigation shows that 13.73% (P_1) of the Dongchong Sandbeach visitors would also pay a visit to Dawei Village. We can thereby calculate the value of co-visitation linkage between Dawei Village and Dongchong Sandbeach as 23.75 (C_{b1}), applying Formula (2). Likewise, only 0.58% (P_2) of Dongchong Sandbeach visitors choose to stop by Dawei Village and Shangwei Village in one trip. We can thereby calculate the co-visitation linkage between these two villages as 1.00 (C_{b2}), applying Formula (3).

Fig. 5 shows that there are a total of 1258 co-visitation linkages on the Peninsula with an average value of 14.62 and a network density of 46.58%. In this network, we can clearly read the linkages between any two locations and make comparison on the strength of linkages. For instance, Yangmeikeng Valley has the largest number of co-visitation partners (=51) and the highest value of co-visitation linkage appears between Xichong Sandbeach and Hesou Village (=219). We can also easily identify those backbone linkages, existing between Yangmeikeng Valley and Xichong Sandbeach, Yangmeikeng Valley and Luzui Villa, Yangmeikeng Valley and Dapeng Fort, and Dongchong Sandbeach and Xichong Sandbeach.

We can develop a complete inventory when focusing on any of the targeted locations, which lists its co-visitation partners and the strength of co-visitation linkages. For instance, Dapeng Fort, one of the most popular scenic attractions in the area, has 45 partners, with Yangmeikeng Valley as its closest partner (Fig. 6).

4.3. Interpreting the co-visitation network

4.3.1. Underlying structure of the co-visitation network

The network structure analysis result suggests that there exist two obvious co-visitation clusters on the Peninsula (Fig. 7). Cluster 1 is of smaller scale and consists of 14 locations, mostly situated along the northwestern coast. Cluster 2 is of lager scale and consists of 46

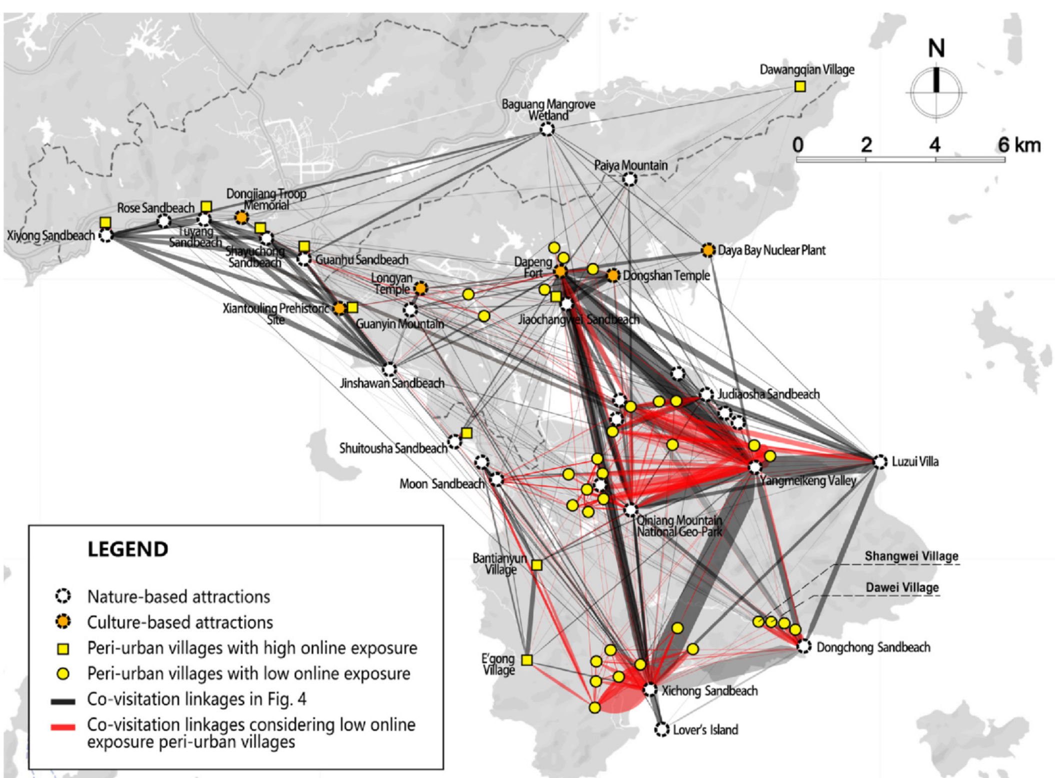

Fig. 5. The complete co-visitation network between all 74 targeted locations. Co-visitation linkages in Fig. 4 are supplemented through considering 32 *peri*-urban villages with low online exposure. In addition to grey lines in Fig. 4, new co-visitation linkages are mapped with red lines. The width of red lines is proportional to the strength of corresponding co-visitation linkages according to calculation from Formula (2) and Formula (3).

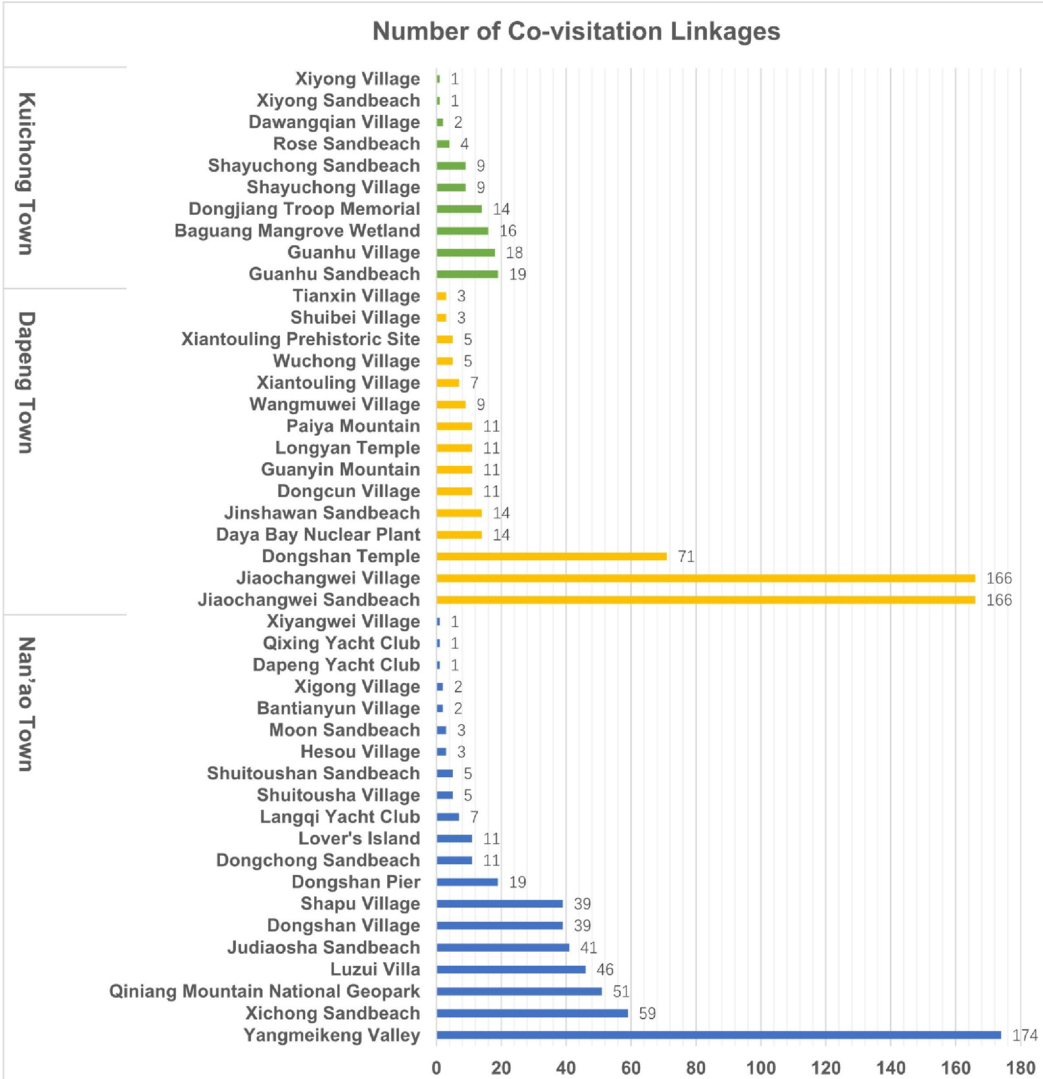

Fig. 6. The inventory showing the number of co-visitation linkages of Dapeng Fort.

locations. Rose Sandbeach and Baguang Mangrove Wetland simultaneously belong to both clusters. The rest 16 locations, mostly villages, are not included in either cluster.

4.3.2. Roles of different locations in the co-visitation network

Roles of different locations in the co-visitation network are explored through degree centrality analysis and betweenness centrality analysis. In terms of degree centrality, Yangmeikeng Valley has the highest value (= 0.753), implying that it is most active and sociable among all locations. The Valley is followed by Xichong Beach (= 0.644), Dapeng Fort (= 0.616), Jiaochangwei Sandbeach & Jiaochangwei Village (= 0.548) and Dongchong Sandbeach (= 0.534) (Fig. 8). These locations with high degree centrality are often those attractions with rich landscape resources. Locations with high degree centrality function as the backbones in maintaining the integrity of the network. Except the aforementioned locations, locations with values below 0.5 take up the majority. Particularly, degree centrality of *peri*-urban villages is relatively low.

On the other hand, only 18 out of 74 locations demonstrate betweenness centrality (Fig. 9). They function as bridges between other locations in the co-visitation network. Among these locations, Xichong Sandbeach (= 0.238), Yangmeikeng Valley (= 0.217) and Dapeng Fort (= 0.192) rank top three in terms of betweenness centrality and are much higher than other locations.

5. Discussion

5.1. Understanding co-visitation patterns in the tourism-driven peri-urban area

The analysis of co-visitation network sheds light on insights of interaction mechanisms between different locations as well as a general portrait of tourism-driven urbanization in *peri*-urban China. Through de-constructing the co-visitation network, we classify co-visitation linkages into those between attractions, those between an attraction and a village, and those between villages. Different co-visitation patterns, manifest from their co-visitation linkages, reflect interplays between different locations.

Co-visitation linkages between attractions demonstrate four patterns. The first is what we call “alliance between the powerful”. The powerful attractions, or core attractions, are those landscapes endowed with rich resources, high customer loyalty, and good online reputation. The combination of these attractions usually forms popular travelling routes. In our case, these core attractions include Dapeng Fort (the national-level historic and cultural heritage), Yangmeikeng Valley (the shooting location of several well-known movies) and Xichong

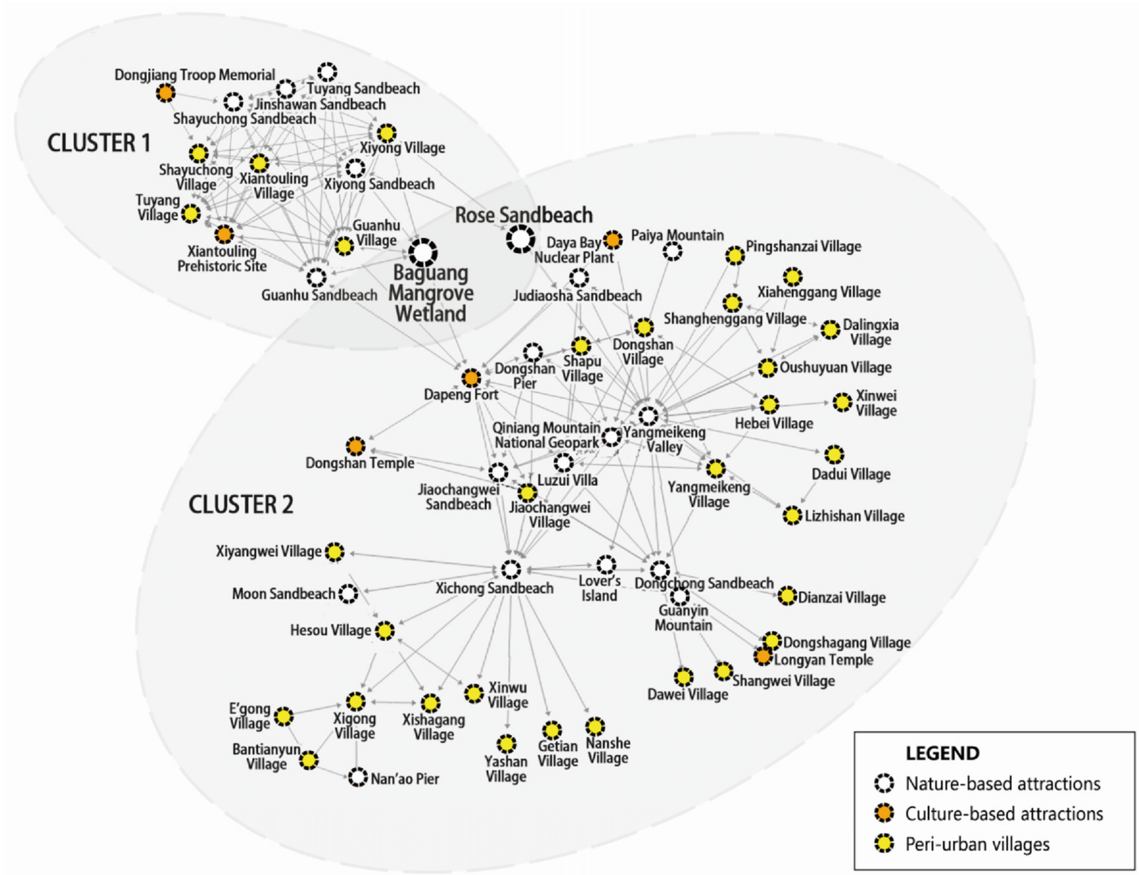

Fig. 7. The network structure analysis of targeted locations on the Peninsula run by NetMiner. The tie between any pair of locations is established as long as the value of co-visitation linkage is above the overall average of 14.62 (Kamada & Kawai, 1989). This figure shows two clusters consisting of 58 locations, after pruning off the ties with values below the average. Baguang Mangrove Wetland and Rose Sandbeach serve as the connecting nodes between two clusters.

Sandbeach (the high-quality sandbeach). The linkages between these attractions act as the backbone of the network. The second is "alliance in the vicinity". These attractions are connected mostly due to factors such as geographic proximity and accessibility. Nearby attractions usually cooperate with each other in order to solicit more visitors. The third is "heterogenous combo". This pattern reflects the needs of visitors who seek different experiences through co-visiting diverse attractions. For example, Qiniang Mountain National Geo-park and Dongchong Sandbeach are usually co-visited by those who are interested in both hiking and watersports. The fourth is "homogenous combo". This pattern reflects the needs of visitors who desire to deepen the understanding in particular landscapes through engaging in thematic trips. For instance, the co-visitation linkage between Dongchong Sandbeach and Xichong Sandbeach belongs to this pattern.

A salient feature of attraction-village co-visitation linkage is that the partners of a village are mostly its adjacent attractions. It is because, as we speculate, a majority of villages on the Peninsula are currently not capable of attracting visitors by themselves and have to share visitors with their adjacent attractions. This affiliation relationship can lead to a win–win situation, if handled properly. On one hand, attractions, faced with increasingly severe restrictions on further expansion for the purpose of natural conservation and preservation of original landscapes, are reliant on nearby villages to provide supporting facilities and amenities. On the other hand, through absorbing those visitors from attractions, villages can gain employment opportunities, increase economic viability, stimulate social regeneration and improve built environment. These incentives push *peri*-urban villages to form even closer relationships with scenic attractions.

In contrast, co-visitation linkages between villages are quite sparse. Visiting villages is an either-or choice for most visitors. Especially, villages attached to the same attraction usually compete fiercely for customers and cooperative interaction is rarely seen between these villages. As a result, a series of emergent issues, such as the redundant investment on tourism services and vicious competition between different villages, are yet to be tackled with in the future.

Moreover, studying the co-visitation network suggests that the *peri*-urban area has the potential to develop into an important component of urban recreational and leisure space through combining different landscape elements and offering diverse visitor experiences. Tourism-driven urbanization of the *peri*-urban area offers an alternative which not only caters the urgent needs of urban residents to stay close to nature and undertake recreational activities, but also effectively preserves endangering local natural and cultural landscapes.

5.2. Predicting future trends and designing new experiences

Compared with other areas, the future development of the *peri*-urban area is faced with more uncertainties. Practically, this area may easily become the target of the real estate market, because short-term economic gains from land deals are usually far greater than gains from tourism development. Nevertheless, we suggest taking long-term social benefits into account for the general public. The study of the co-visitation network can assist policy decision-making and planning practices on the basis of visitor demands. Possible measures include but are not limited to the following.

First, planners can better arrange traveling routes with the knowledge of co-visitation linkages. We suggest promoting bundle sales, improving accessibility, and sharing customers and resources between co-

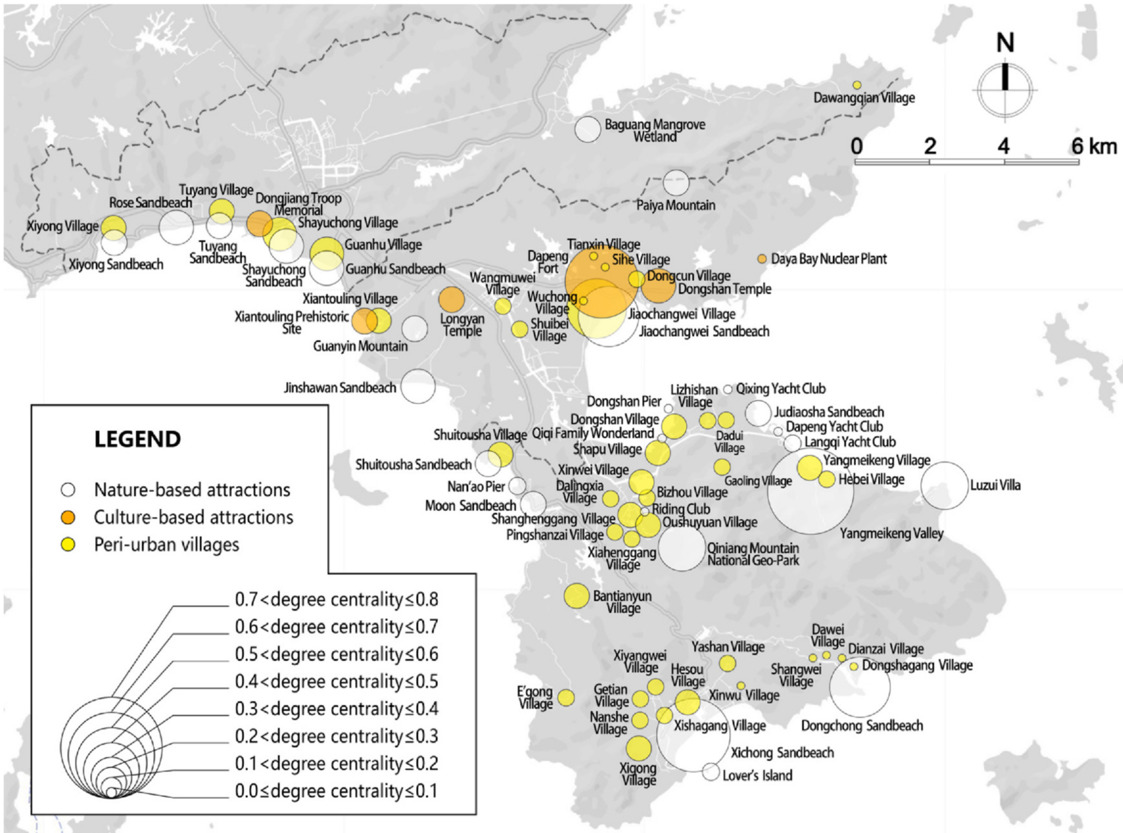

Fig. 8. The degree centrality analysis of locations on the Peninsula. 74 locations are graded according to the values of their degree centrality. A larger circle represents a location with higher degree centrality.

visitation partners, as a general effort of enriching overall travelling experience (Saraniemi & Komppula, 2019; Allred, Fawcett, & Wallin, 2019).

Second, planners need to pay attention to different roles of locations, as revealed by findings from centrality analysis, and apply differentiated measures. We suggest priorities in resource allocation and development be given to those locations with high degree centrality. These locations function as the indispensable catalysts to prosper regional tourism. Such facilities as transfer stations and reception centers are advised to be supplied in locations with high betweenness centrality (Shih, 2006; Lee, Choi, Yoo, & Oh, 2013). Especially, landscape services should be better maintained in core attractions with both high degree centrality and betweenness centrality.

Third, planners should seek to create new travelling experiences beyond the *status quo*. In the current network, most *peri*-urban villages are subordinate to attractions with their values largely underestimated, while many of these villages are in fact quite rich in tourism resources, including traditional folk dwellings, local customs and beliefs, and festivals and events. We suggest selecting some typical villages and developing experience-based tourism in the hope of exerting their resource potential (Oakes, 2006; Li, Cheng, & Wang, 2014).

Taking aforementioned measures, as we predict, will lead to new visible changes in the future co-visitation network. The number and value of co-visitation linkages will increase significantly and co-visitation clusters will be strengthened, thus enhancing the density of the network. Apart from core attractions, more locations will demonstrate higher degree centrality or betweenness centrality. Some *peri*-urban villages are likely to end their subordinate status and advance into new attractions.

6. Conclusion

This study measures co-visitation network in the *peri*-urban area of China under the backdrop of tourism-driven urbanization. To achieve this goal, we integrate different toolkits based on social media data and create a new and replicable methodology. Specifically, values of co-visitation linkages are calculated from massive Weibo posts, supplemented by online questionnaire investigation. Co-visitation linkages constitute an effective indicator of interactions between different locations. After establishing the co-visitation network, we conduct social network analysis in order to reveal the underlying co-visitation structure and identify the roles of different locations. We prove the usefulness of this methodology through a case study, which can be applied in similar situations as well. Nevertheless, two preconditions should be met before duplicating this methodology in other situations, namely, the availability of co-visitation information from social media platforms, and the targeted region containing a considerable number of locations with high online exposure.

As demonstrated throughout the paper, social media analytics assists us to establish the co-visitation network. The network contributes to better understanding spatial distribution of visitor demands, interpreting co-visitation patterns and thereby making more informed plans. Studying co-visitation patterns provides planners with valuable hints to take visitor-oriented measures, such as the arrangement of travelling routes and the creation of new experiences. Above all, the case study is the epitome of tourism-driven urbanization in the *peri*-urban area. Worth more scholastic attention, such a trend offers an alternative to the conventional urban sprawl or industry-driven urbanization and shows that the *peri*-urban area of China is capable of playing an increasingly important part as the recreational and leisure space for urban residents.

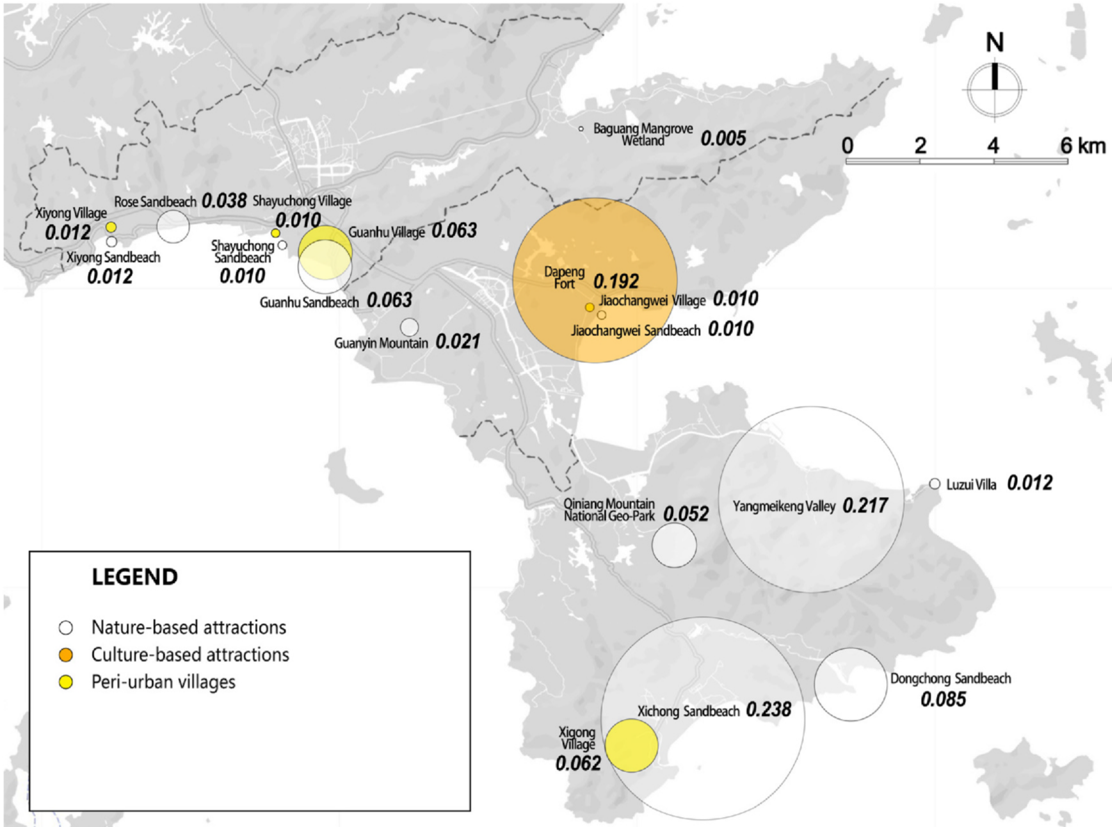

Fig. 9. The betweenness centrality analysis of locations on the Peninsula. The betweenness centrality of only 18 locations can be calculated. The diameter of a circle is proportional to the value of the corresponding location.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (51908243) and National Natural Science Foundation of China (51908362), and Fundamental Research Funds for the Central Universities in China (12619320). The authors would like to thank Professor Hang Ma of the Harbin Institute of Technology for her helpful suggestions.

References

Allred, R., Fawcett, E., & Wallin, C. (2019). A dynamic collaboration capability as a source of competitive advantage. *Decision Sciences, 42*(1), 129–161.

Bastian, O., Grunewald, K., Syrbe, U., Walz, U., & Wende, W. (2014). Landscape services: The concept and its practical relevance. *Landscape Ecology, 29*(9), 1463–1479.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679.

Chinese Academy of Cyberspace Studies (2020, April 28). 45th Statistical report on the development of Internet in China. Retrieved from: http://www.cnnic.cn/gywm/xwzx/rdxw/20172017_7057/202004/t20200427_70973.htm.

Donahue, L., Keeler, L., Wood, A., Fisher, M., Hamstead, A., & Mcphearson, T. (2018). Using social media to understand drivers of urban park visitation in the twin cities, MN. *Landscape and Urban Planning, 175*, 1–10.

Duvernoya, I., Zambonb, I., Saterianoc, A., & Salvatid, L. (2018). Pictures from the other side of the fringe: Urban growth and peri-urban agriculture in a post-industrial city (Toulouse, France). *Journal of Rural Studies, 57*, 25–35.

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2013). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers, 102*(3), 571–590.

Furht, B. (2010). *Handbook of social network technologies and applications*. Springer Press.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth, 3*(3), 231–241.

Goossen, M., & Langers, F. (2000). Assessing quality of rural areas in the Netherlands: Finding the most important indicators for recreation. *Landscape and Urban Planning, 46*(4), 241–251.

Guan, J., Gao, J., & Zhang, C. (2019). Food heritagization and sustainable rural tourism destination: The case of China's Yuanjia Village. *Sustainability, 11*(10), https://doi.org/10.3390/su11102858.

Guo, Y., Xiao, Y., & Yuan, Q. (2017). The redevelopment of peri-urban villages in the context of path-dependent land institution change and its impact on Chinese inclusive urbanization: The case of Nanhai, China. *Cities, 60*, 466–475.

Kaltenborn, B., Kaltenborn, E., & Linnell, J. (2019). "It's all about the scenery": Tourists' perceptions of cultural ecosystem services in the Lofoten Islands, Norway. *Journal of the Arctic Institute of North America, 72*(1), 38–51.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters, 31*(1), 7–15.

Kim, D., & Kim, S. (2017). The role of mobile technology in tourism: Patents, articles, news, and mobile tour app reviews. *Sustainability, 9*(11), https://doi.org/10.3390/su9112082.

Komossaa, F., Zanden, E., & Verburg, P. (2019). Characterizing outdoor recreation user groups: A typology of peri-urban recreationists in the Kromme Rijn area, the Netherlands. *Land Use Policy, 80*, 246–258.

Kylänen, M., & Rusko, R. (2011). Unintentional coopetition in the service industries: The case of Pyhä-Luosto tourism destination in the Finnish Lapland. *European Management Journal, 29*(3), 193–205.

Lang, W., Chen, T., & Li, X. (2016). A new style of urbanization in China: Transformation of urban rural communities. *Habitat International, 55*(1), 1–9.

Leaf, M. (2002). A tale of two villages: Globalization and peri-urban change in China and Vietnam. *Cities, 19*(1), 23–31.

Lee, S. H., Choi, J. Y., Yoo, S. H., & Oh, Y. G. (2013). Evaluating spatial centrality for integrated tourism management in rural areas using GIS and network analysis. *Tourism Management, 34*, 14–24.

Leung, D., Law, R., Hoof, H., & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. *Journal of Travel & Tourism Marketing, 30*(1–2), 3–22.

Li, J. (2020). Culture and tourism-led peri-urban transformation in China: The case of Shanghai. *Cities*. https://doi.org/10.1016/j.cities.2020.102628.

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science, 40*(2), 61–77.

Li, M., Cheng, H., & Wang, J. (2014). Making a cultural cluster in China: A study of Dafen Oil Painting Village, Shenzhen. *Habitat International, 41*, 156–164.

Liu, B. (2011). *Web data mining: Exploring hyperlinks, contents and usage data* (2nd ed.). Chicago: Springer.

Lorgnier, N., Su, C. J., Woratschek, H., Horbel, C., & Popp, B. (2014). Considering coopetition strategies in sport tourism networks: A look at the nonprofit nautical sports clubs on the northern coast of France. *European Sport Management Quarterly, 14*(1), 87–109.

Matatolu, I. (2019). A critical examination of indigenous people, tourism and quality of

life. *Journal of Heritage Tourism*. https://doi.org/10.1080/1743873X.2019.1597101.
Nahuelhual, L., Carmona, A., Lozada, P., Jaramillo, A., & Aguayo, M. (2013). Mapping recreation and ecotourism as a cultural ecosystem service: An application at the local level in southern Chile. *Applied Geography, 40*, 71–82.
Newman, M. (2010). *Networks: An introduction.* Oxford University Press.
Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with Pajek, Vol. 27*. Cambridge: Cambridge University Press.
Oakes, T. (2006). Cultural strategies of development: Implications for village governance in China. *The Pacific Review, 19*(1), 13–37.
Sanesi, G., Colangelo, G., Lafortezza, R., Calvo, E., & Davies, C. (2017). Urban green infrastructure and urban forests: A case study of the Metropolitan Area of Milan. *Landscape Research, 42*(2), 164–175.
Saraniemi, S., & Komppula, R. (2019). The development of a destination brand identity: A story of stakeholder collaboration. *Current Issues in Tourism, 22*, 1116–1132.
Serra, P., Saurì, D., & Salvati, L. (2017). Peri-urban agriculture in Barcelona: Outlining landscape dynamics vis à vis socio-environmental functions. *Landscape Research*. https://doi.org/10.1080/01426397.2017.1336758.
Sharma, D., & Surolia, A. (2013). *Degree centrality.* New York, NY: Springer.
Shih, H. Y. (2006). Network characteristics of drive tourism destinations: An application of network analysis in tourism. *Tourism Management, 27*(5), 1029–1039.
Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: The first decade. *Tourism Geographies, 18*(5), 587–606.
Sina Weibo Data Center (2019). Annual report of Sina Weibo users in 2018.
Song, H., & Liu, H. (2017). Predicting tourist demand using big data. In Z. Xiang, & D. Fesenmaier (Eds.). *Analytics in smart tourism design*. Springer.
Stienmetz, J. L., & Fesenmaier, D. R. (2018). Destination value systems: Modeling visitor flow structure and economic impact. *Journal of Travel Research*. https://doi.org/10.1177/0047287518815985.
Sun, Y., & Shao, Y. (2020). Measuring visitor satisfaction toward peri-urban green and open spaces based on social media data. *Urban Forestry & Urban Greening*. https://doi.org/10.1016/j.ufug.2020.126709.
Tencent Company (2020). WeChat Data Report in 2019. Retrieved from: https://mp.weixin.qq.com/s/vmhoiRzpBs7-JK_x2a7gZw.
Teixeira, G., Marques, P., Garabini, T., Cezar, M., & Pereira, A. (2018). The effects of landscape patterns on ecosystem services: Meta-analyses of landscape services. *Landscape Ecology*. https://doi.org/10.1007/s10980-018-0673-5.
Tian, L., & Guo, Y. (2019). *Peri-urban China: Land use, growth, and integrated urban–rural development.* Routledge.
Tieskens, F., Zanten, V., Schulp, E., & Verburg, H. (2018). Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landscape and Urban Planning, 117*, 128–137.
Tolstad, H. K. (2014). Development of rural-tourism experiences through networking: An example from Gudbrandsdalen, Norway. *Norwegian Journal of Geography, 68*(2), 111–120.
Ullah, H., Wan, W., Haidery, S. A., Khan, N. U., Ebrahimpour, Z., & Luo, T. (2019). Analyzing the spatiotemporal patterns in green spaces for urban studies using location-based social media data. *International Journal of Geo-Information*. https://doi.org/10.3390/ijgi8110506.
Wilde, S. J. & Cox, C. (2008). Linking destination competitiveness and destination development: findings from a mature Australian tourism destination. Proceedings of the Travel and Tourism Research Association (TTRA) European Chapter Conference-Competition in tourism: Business and destination perspectives.
Wood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports, 3*, 2976. https://doi.org/10.1038/srep02976.
Xiang, Z., & Fesenmaier, D. R. (2017). Analytics in tourism design. In Z. Xiang, & D. Fesenmaier (Eds.). *Analytics in smart tourism design.* Springer.
Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management, 31*(2), 179–188.
Yang, Z., Yin, M., Xu, J., & Lin, W. (2019). Spatial evolution model of tourist destinations based on complex adaptive system theory: A case study of Southern Anhui, China. *Journal of Geographical Sciences, 29*, 1411–1434.
Yuan, J., Chen, K., Li, W., Ji, C., Wang, Z., & Skibniewski, M. J. (2018). Social network analysis for social risks of construction projects in high-density urban areas in china. *Journal of Cleaner Production, 198*(2), 940–961.
Zhang, X. (2019, March 6). Ctrip: Total transaction users in 2018 surpassed 135 million. Economic Daily.
Zlender, V., & Thompson, C. (2017). Accessibility and use of peri-urban green space for inner-city dwellers: A comparative study. *Landscape and Urban Planning, 165*, 193–205.