Journal of Transport Geography 96 (2021) 103178

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Using cellular data to analyze the tourists' trajectories for tourism destination attributes: A case study in Hualien, Taiwan

Chih-Peng Chu, Yu-Hsin Chou *

Department of Business Administration, National Dong Hwa University, No. 1, Sec. 2, Da Hsueh Rd. Shoufeng, Hualien 97401, Taiwan

ARTICLE INFO

Keywords:
Multi-destination tourism
Brokerage
Betweenness centrality
Brokerage role analysis
Mobile phone positioning data
Social network

ABSTRACT

This study adopts social network analysis (SNA) to identity the spatial role of nodes in a tourism region. The trajectory of multi-destination tourists will provide great information for tourism management, but tracking tourists' movements is a big challenge. Thanks to the popularity of mobile communication devices (mobile phones) nowadays, these can passively or actively provide information on owners' movement. This study uses call detail records (CDR), which are mobile phone positioning data, to construct a trajectory network of tourists in Hualien, Taiwan. Three types of centralities, as well as brokerage analysis and role analysis, are used to measure the relations in a complicated network. Degree centrality (DC) reveals popular nodes that have good relationships with others directly. Reach closeness centrality (RCC) shows indirect relationships with others. Betweenness centrality (BC) indicates the most important mediator nodes that help others to connect. BC and brokerage analysis are used to classify destinations into five brokers: the coordinator, consultant, gatekeeper, representative, and liaison. We also identify social positions of each destination by role analysis. This study identifies a total of 78 nodes (origins and destinations) and their functions in the tourism network. We find some nodes have more than one mission, while some only have one. The functions of nodes can be different on weekdays and weekends. We distinguish the tourism region into different size of districts. The results of this study can help managers consolidate strategies for tourism marketing.

1. Introduction

With the promotion of two-day weekly rest life style, it stimulates the demand for more in-depth vacations. Hualien is an eastern city in Taiwan. According to Survey of Travel by R.O.C. Citizens (Tourism Bureau, 2019), over 50% of travelers stayed Hualien for two or three days. The report also showed the travelers came from residents and visitors that are lived in neighboring cities. This domestic travel style involves short distances and fewer days than before. Tourism information is also easier to obtain, so more and more tourists choose to arrange their own travels instead of going through travel agents. The movement patterns are needed not only to understand tourist behaviors but also to classify the destination attractions. Most people visit multiple destinations on their journeys. When a planned destination cannot meet travelers' desires, they might skip it and go to the next destinations on their vacation. So, it is difficult for a tourism destination to operate independently in a region.

Tourism resource management not only faces the problem that the administration of distinct destinations is different, but also considers the geographical restrictions that limit tourists' movement. It is difficult to establish a strategy of cooperation. Currently, social network analysis (SNA) has developed as a tool for providing the relationship among social units, including people, organizations, events, and even geographic locations, and can concretize their intangible relations (eg. Carrasco and Miller, 2009; Jones and Lucas, 2012; Pike and Lubell, 2016).

This paper integrates the logic of social networks with the concept of the region design model which Dredge (1999) proposed to construct spatial tourism networks. Previous articles have shown that centrality indexes are a basic method to measure the individual nodes of such networks (Freeman, 1978). Centrality indexes reveal which nodes have social power. Degree centrality, reach closeness centrality, and betweenness centrality (DC, RCC and BC) are employed to investigate centers of tourism destinations. DC reveals popular nodes that have good relationships with others directly. RCC shows indirect relationships with others. BC indicates the most important mediator nodes that help others to connect. Although BC can evaluate the ability of a destination to be an intermediate one, that is to say, BC examines the extent to which a node

* Corresponding author.
E-mail address: 610532001@gms.ndhu.edu.tw (Y.-H. Chou).

https://doi.org/10.1016/j.jtrangeo.2021.103178
Received 9 June 2020; Received in revised form 19 July 2021; Accepted 22 August 2021
Available online 10 September 2021

occupies a location linking node pairs, BC cannot reveal more about the types of intermediators. This paper combine BC and brokerage analysis (Gould and Fernandez, 1989) to distinguish the bridge types of tourism destinations. No matter centrality or brokerage indexes, these mothed discuss features for particular nodes. Actually, all nodes can be distinguished into different roles in a social network. Role analysis can find the relationship between each role. By using structural equivalence, the research provides a process to discover the tourist mobility between tourism destinations of each role. If some nodes connect tightly within a social network, they are called social group. As members of a social group in a tourism network, they are connected by actual movements and even are affected by geographical features. The research compute n-clan analysis and use geographic information systems (GIS) to design the routes for different social groups.

Questionnaires for the pre-analysis, post-analysis, or en-route surveys are the traditional approaches to collected information. The pre-analysis does not fully present the actual journal. It is also difficult to record very details movements in the post-analysis. En-route surveys provide the cross-section when interviewing the traveler sample during the trip. The data come from automatics traffic counts and or some personal interviews. However, there are many new technologies or assistant tools to help in data collection, like GPS log data, smart card data, mobile phone data, and social media (Yue et al., 2014). Since people now use mobile phones everywhere, they leave their footprints whenever they connect with the base stations to request telecommunication service, creating data called cellular data or mobile phone positioning data. From the demand point of view, we attempt to use such data, measuring actual tourist mobility, to build tourism networks.

Based on the research motivation, there are three objectives as below:

1. This research integrates the logic of social network with the concept of region design model. Tourist mobility (tourist flow) establishes the relationship of the nodes, tourism destinations. Trajectories of tourist mobility link pairs of nodes as the ties.
2. After transforming the tourism OD matrix into the spatial network, some indexes (degree centrality, betweenness centrality, and closeness centrality) are employed to investigate centers of tourism destinations. By using the brokerage analysis, this research distinguishes the bridge types of tourism destinations and the roles of roles destination.
3. When members of a social group tie together by geographical features, they might turn into the tourism routes. We compute n-clan analysis and use geographic information systems (GIS) to design the routes for different social groups.

2. Literature review

2.1. Multi-destination trip and tourism region design

People engage in a multi-destination tour as a rational behavior (Buchanan, 1983; Lue et al., 1993; Lue et al., 1996). First, multiple destinations can fulfill multiple desires for each group member. Second, it is hard for tourists to seek a series of benefits in a single destination. Third, multi-destination tours can reduce uncertainty and the risk of travel. Fourth, diverse destinations are needed to accommodate the constraints of the actual spatial, temporal, physical or psychological conditions. Multi-pronged tours also reduce the overall tourism cost. However, the tourism industry and research studies still not pay enough attentions on elements and effects of the travel chain behavior on the characteristics of destinations (Schiefelbusch et al., 2007).

Dredge (1999) identified six elements for destination region design. The framework consists of the following:

1. Destination Regions: Tourists can go everywhere in the region. Most management practices segment regions via administrative boundaries. It may be restricted and make an inappropriate plan.
2. Tourist Generating Markets: Markets and destinations are inseparable. The markets are usual places where the potential tourists live. Each destination has its diverse and limited markets. The characteristics of generating markets influence the choices of tourism destinations.
3. Nodes: Attraction complex contains the attractions, sights or objects that tourists visit. They have the cooperative and competitive relations. However, complementarity between them greatly enhances overall tourism attractiveness. Service components are supported facilities, like accommodation, restaurants, retail stores and so on. They bring economic value to a region. Two types of nodes are interdependent, and the definitions of them are gradually blurred. Primary nodes are the motive power that tourists visit the region. When they still are in their home, they already know these nodes. Tourists know the secondary nodes that are not the reason for visiting, but they can increase the overall attractiveness. Tertiary nodes are strange to the tourists when they are not going to the region. These nodes do not influence the arranged itinerary but extend the length of stay.
4. Districts: A region is a cluster of districts. Some nodes with a given cohesiveness or an internal consistency organize a district. These nodes group together as a result of a short distance or similar style.
5. Circulation Route: Paths permit tourists to move among nodes. Although everyone has the own spatial behavior, tourists may visit the same nodes during various trips. The quality of alternative routes may affect the decision-making of tourists. Not all tourists will use the same routes to their accommodation. Stewart and Vogt (1997) pointed out five trip patterns including single destination, en-route, base camp, regional tour, trip chaining (Fig. 1). Lew and McKercher (2006) pointed out three patterns of linear path models including point to point, circular pattern, and complex patterns (Table 1). Pazzaglia and De Beni (2001) summed up in three way-finding patterns. Landmark-centered-strategy is that people often refer to visible landmarks to identify the directions and then construct the routs without the fixed routes. Route-survey is that people move in a consecutive linear path between two or more landmarks that are origin or destination. And in the process, they will pass multiple anchor points. Orientation-strategy is that people have spatial knowledge at their fingertips. Their movement is elastic by mental cognition map.

6. Gateway: Gateways are along the roads. They are the entrances to arrive in the regions or the exits to link next region. The generating markets cause the region possesses multiple gateways. Gateways hold a power to adjust tourists' perception for regional tourism.

We know the tourism resources are scattering and immovable. They are the driven for the distribution of tourism flow. Because tourists usually stay more than one node on a trip, the destinations have to coordinate, cooperate and co-marketing based on their benefits. The planners have to dispatch tasks effectively. The background of cooperation needs to build on the spatial structure. Regarding the spatial interaction, regional planning considers two parts: nodes and generating markets. However, tourist mobility is a form of spatial interaction. This is useful for planners to identify tourism partnerships.

2.2. Social network analysis in tourism

2.2.1. Social network theory

In the beginning, the Social relationship is abstract, concealed and complex. Radcliffe-Brown (1940) identified a social relation that is when two or more individual organisms have some interactions with a purposive behavior. Because combing with Sociometry and Graph

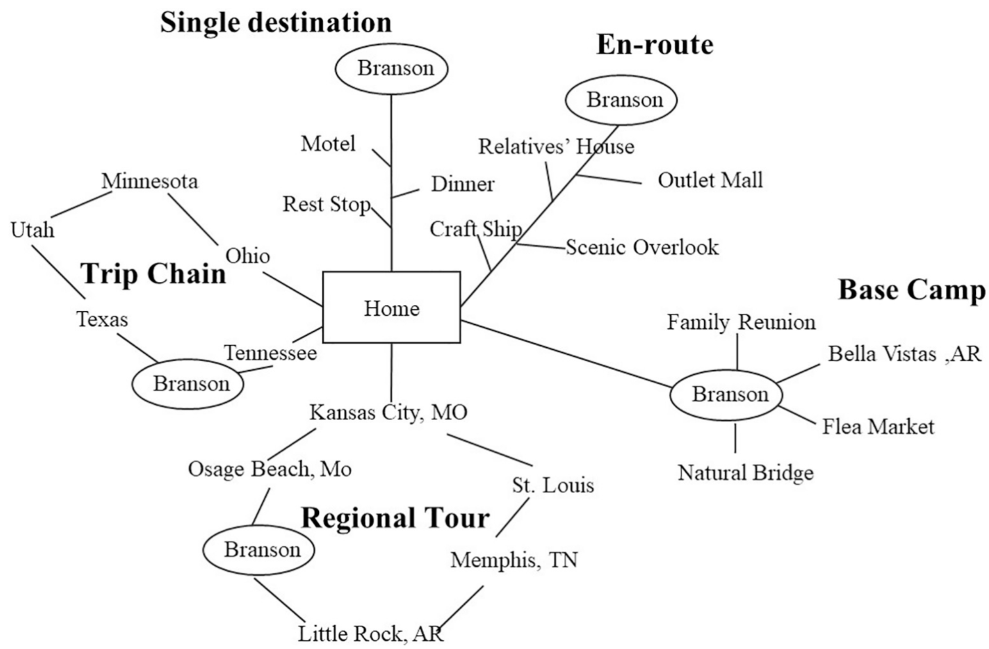
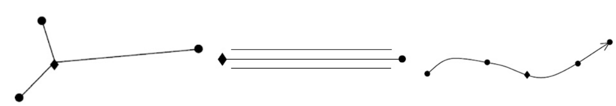
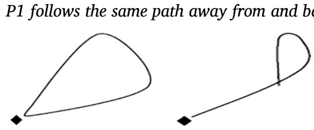
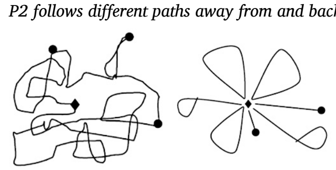


Fig. 1. Branson trip patterns (Stewart and Vogt, 1997).

Table 1
Linear path models of tourist behavior.

Linear path models	
Point-to-Point	 *P1 follows the same path away from and back to the place of accommodation.*
Circular	 *P2 follows different paths away from and back to the placeof accommodation.*
Complex	 *P3 consists of complex and undifferentiated combinations of the point-to-point and circular patterns.*

Data Source: Lew and McKercher (2006).

theory is a breakthrough application, the social relationship can be observed and analyzed. Moreno (1937) mentioned Sociometry is the quantitative method to study interpersonal relations among group members. At that time, it was still an idea. But he invented sociogram that presented social configuration used points as individuals and lines as relations. Mayo (1933) found the informal relationships are concealed in a formal organization. Warner and Srole (1945) pointed out a social configuration is made of the sub-groups. Clique is one type of sub-groups. We compare this phase like to introduction of the product lifecycle.

After 1950s, the scholars have paid attention to social network with a systematic development. Barnes (1954) separated territorial, industrial, and boundaryless types of social fields in a social system. He claimed a total network can display the whole social life and a partial network is one part of informal relations in a network which include the kinship, the friendship, and the neighborliness. Mitchell (1969) defined the total, partial and ego-centric network. He thought a partial network is any extract of the total network based on some criterion applicable. He also definitely referred to several morphological and interactional characteristics, such as reachability, density, range, directedness, intensity.

After the 1970s, we called maturity phase. There were lots sub-theories produced from two perspectives. One is the mathematical analysis by graph theory, especially applying algebra and statistic. Many measurements are proposed for the social structure (White et al., 1976; Freeman, 1978; Wasserman and Faust, 1994)

The others are conceptualization theories. Such as, the strength of weak ties. Granovetter (1973) show the importance of the bridge and weak ties. Granovetter found the linkage with the outside could make people find the job faster than asking their social circle. The weak ties enhance the opportunity of information mobility between sub-groups that link with strong ties independently. So weak ties are resources affecting social cohesion. Granovetter (1985) demonstrated social embeddedness by showing the rational behavior (ex: economic action) are embedded in social relationships. The structural embeddedness comes from an official network. The relational embeddedness comes from an informal personal network. Burt (1992) presented the concept pf structural holes by describing relationships of nonredundancy between pairs of nodes. A best network followed the efficiency and effectiveness principles. It must focus on keeping the connection among center nodes and primary nodes. The primary nodes are like ports of access and can extend the network. They bring the same information that is from the center node to their respective sub-groups, but they return different benefit information for the center. If a node has more structural holes, it holds more advantages.

2.2.2. Elements of SNA and applying SNA in Tourism

Social network analysis (SNA) is a method with a set of indictors to analyze the framework of relations constituted by different social units (individual, group & community). Wellman (1983) pointed out SNA searches the deeper structure, and excavates the regular network patterns hidden in the complicated social system. Therefore, we can learn how network structures constrain social behavior and social change. Wasserman and Faust (1994) defined a social network is composed of one or more groups of actors and their given relations. There are three elements to form a social network.

1. Actor: An actor is not just a concrete person and is a social unit. An organization, a group, a country, or an event can be called an actor, the position of an actor in the network is node.
2. Relation: A social network includes many types of relationships. Diversified relations generate different network patterns such as kinship, managers and staff, exchange (material, information).
3. Relational tie: If two nodes connect with a tie, it means the particular relation is tenable. It is notable that the object of SNA is the relation between actors, not the actor attributes. In common statistical approach, the data sets record rows as samples and columns as the attributes in a rectangular matrix. The relational data sets of social network are presented in a squared matrix. The rows and columns are the actors. The observation is the binary or strength, directed or undirected between each pair nodes.

There are five research directions in tourism using SNA, including the spatial structure, policy network, enterprise development and management, economic ties, and tourism research analysis. Some researchers have pointed out that SNA is an appropriate method to analyze the spatial intersection among destinations. For example, D'Agata, Gozzo, and Tomaselli (2013) computed density, centralities, components, cliques, bridges and so on in Sicily. Shih (2006) surveyed the flow of domestic tourists in Nantou, Taiwan by telephone interview. He described the structural characteristics of centralities and structural holes. Chung, Chung and Nam (2017) collected data on South Korean tourist mobility in Europe through a social media called NAVER blog (a S. Korean website and app), and used centralities to understand the connections between tourist cities and compare the difference across years. They discovered that transportation affected a city's degree centrality and betweenness centrality, because cities with a high DC score have flights between South Korea and these cities, and the Eurostar high-speed rail system affected cities' BC. Zee and Bertocchi (2018) applied user-generated content (UGC) data from TripAdvisor reviews to analyze review networks which were either international, domestic or local in scale, by a degree index. They showed that the geographic distances (in space) traveled by international tourists seem to be similar to their relational distances (in the review networks). They also pointed out that popular attractions do not necessarily occupy core positions, and may instead merely connect with core nodes. Sugimoto, Ota and Suzuk (2019) investigated tourist trajectories in Ueno district of Tokyo, Japan by using GPS logs. They not only applied strength, betweenness and page rank centralities, but also explained the influence between centralities and tourist duration time.The tourist flow or tourist mobility is the process that the destinations interact within the region and between the generating market. (Casanueva et al., 2016; Zeng, 2018; Kang et al., 2018). Tourism flow is the sets of relation types, for example, information flow, logistics flow, cost flow and so on. But we introduce the spatial stricture because it accords with our study.

2.3. Mobile phone positioning data

Mobile phone data can show the trajectories of users from Wi-Fi, Bluetooth, GPS, and base stations. Each positioning system of mobile phones has specific advantages and disadvantages, and none can entirely fill the needs of the present research inquiry. GPS positioning calculates the exact location (latitude, longitude, and altitude) of a GPS receiver from at least three satellites. The accuracy is nearly 10 m. Because GPS positioning needs an open sky environment, it might be suspended in an urban canyon area (Schiller and Voisard, 2004). Wi-Fi signals are embedded in selected indoor and outdoor spaces. The high AP (access point) density in urban areas allows nearly complete tracking data, but APs may be lower in number or even absent in a rural village (Zandbergen, 2009). Bluetooth allows mobile phones to exchange and communicate without wires, but through beacons (Bekkelien et al., 2012). If there is no access point or beacon detector, one cannot capture the individual tracking data collected by Wi-Fi or Bluetooth positioning.

The basic functions of mobile phones are to make a call and to surf the internet. A broad geographic space is partitioned into lots of cells, which sets up with a base station. When a signal connects with a base station to request telecommunication service, the triangulation technique is applied to determine the user location (Wang et al., 2017). The telephone systems generate call detail records (CDRs) that capture the telecommunication transactions. The second source is Internet usage by IP detail records (IPDRs). These two are collectively called CDR in this study. The accuracy is 200–500 m in urban areas and 300–1000 m or more in rural areas (Ahas et al., 2010). Iqbal et al. (2014) pointed out that the space-time stamps are lost when users do not use phones, so some individual trips may be missed. Acquired data include transient origins and destinations, not the actual OD. Sometimes, the users do not move, but the data show the change between locations. This is because the base station maybe carries exceed users, the telecom shifts users to the neighboring station to avoid the traffic (Table 2).

Most of the fundamental applications of cellular data in human behavior have been to lifestyle travel or transportation planning (Çolak et al., 2015; Iqbal et al., 2014; Tettamanti et al., 2012; Milne and Watling, 2019), while Ahas et al. (2008) analyzed the distribution of outbound tourists in Estonia. With the commercial application of communication technology, we can expect more accurate and timely passenger whereabouts information to be available (Bonnel et al., 2018; Crawford et al., 2018; Mamei et al., 2019). But it also depends on privacy restrictions and how much information consumers are willing to release.

3. Methodology

3.1. Study area

Hualien is a tourism county surrounded by mountains and seas in Taiwan. Despite being constrained by geological conditions, Hualien still attracts over ten million people annually (*Hualien Tourism Department*, 2017). Because traffic accessibility is facilitated by the improved Suhua Highway, it will enhance a flood of holidaymakers from west of Taiwan in the future (Fig. 2).

This paper constructs a spatial network between supply and demand in tourism. The tourism resources of Hualien are concentrated in four recreation systems: Taroko (A), Hualien city and suburbs (B), the East Rift Valley (C), and the east coast (D). After consulting a tourism map, we take into account three components, which are 74 attractions, two transportation systems and the accommodation representatives (E), as the tourism destinations. The accommodation information is extracted

Table 2
An excerpt from CDR data

ID	Call Date	Call Time	Duration	Latitude	Longitude
AAH03JABiAAJKnPAa5	20,120,707	06:15:20	109	23.8153	90.4181
AAH03JABiAAJKnPAa5	20,120,707	09:03:06	109	23.8139	90.3986
AAH03JABiAAJKnPAa5	20,120,707	10:34:19	16	23.6989	90.4353
AAH03JABiAAJKnPAa5	20,120,707	18:44:53	154	23.6989	90.4353

Data Source: Iqbal et al. (2014).

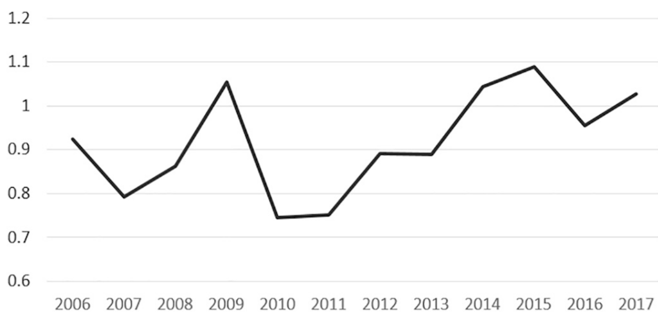


Fig. 2. Visitor arrivals by years (unit:10 million).

from government open data. E is the aggregation of grids such that the number of hotels and bed and breakfasts (B&B) in a grid are more than ten. These nodes are categorized by the resource base, as named attribute, are (Fig. 3), including (1) natural landscape, (2) humanity and culture, (3) recreation, (4) shopping and specialty, (5) transportation, and (6) accommodation. The generating markets (F) are the other counties in Taiwan (Fig. 3, Tables 3 and 4).

3.2. Data collection

The CDR data came from the project of Travel Demand Survey and Transport Planning for Hualien County. We collected CDR data for a week, from October 16 to 22, 2016, consisting of 621,363 cell phone activities. The data generated from resident behaviors in local and visitor behaviors from other counties. A cell phone activity was recorded when a mobile phone signal was detected for over 15 min at a location (origin) and then this phone signal appeared at another location (destination). To protect individual privacy, the telecommunication operator developed an OD (origin – destination) table which simplified the individual journeys that were the aggregation of space-time stamps made from the users. A 250 m × 250 m grid square was superimposed to partition Hualien County (Fig. 4). The users' flow was calculated on an

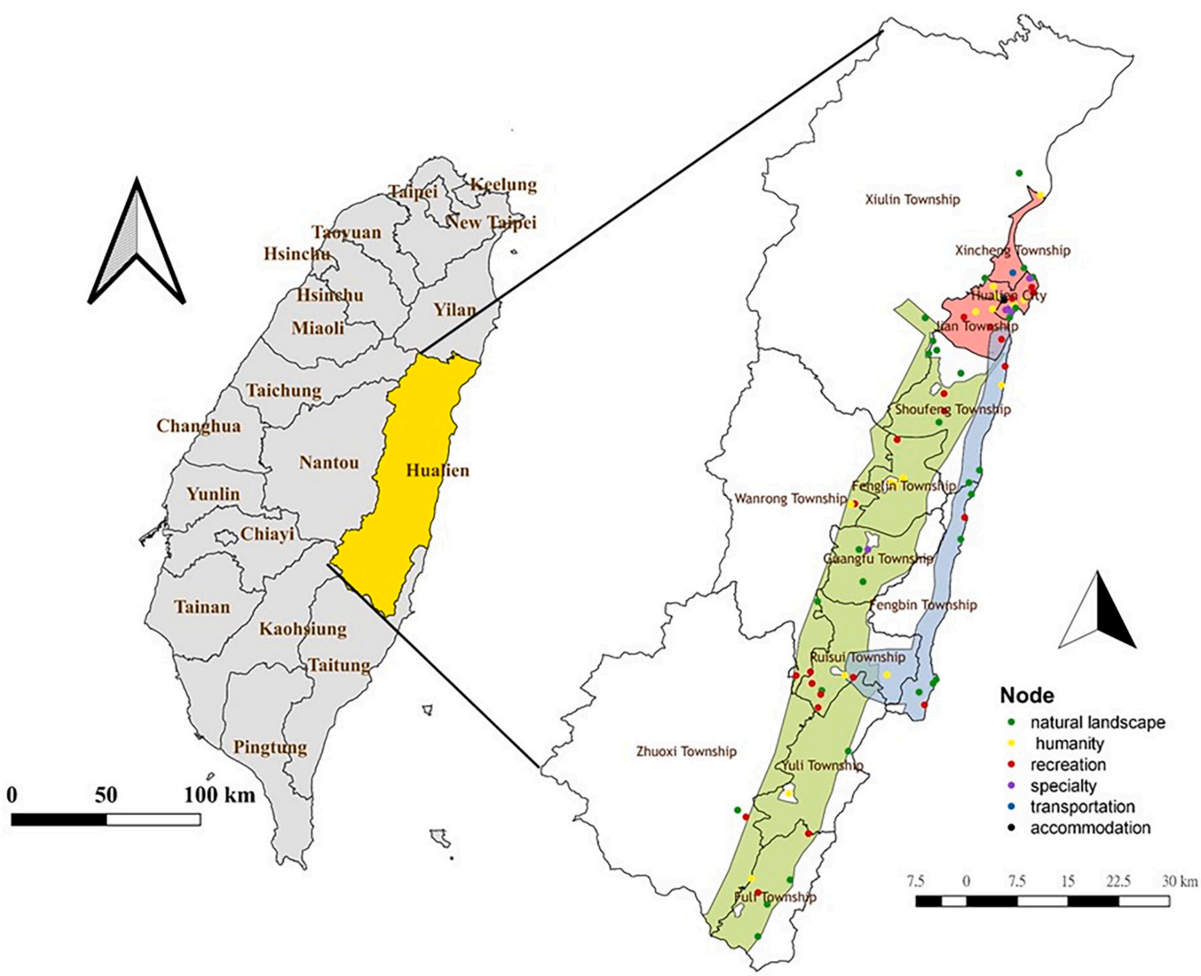


Fig. 3. Map of the study area (left: generating markets; right: Hualien recreation systems).

Table 3
Tourism destinations in Haulien, Taiwan.

Node	ID	Attribute	Node	ID	Attribute
Taroko National Park	A	1	Hualien Airport	B9	5
Lighthouse	B1	1	Danongdafu Forest Park	C1	1
Hualien County Stone Sculptural Museum	B10	2	Yunshanshui Wetlands Ecological Park	C10	1
Hualien Tourism Distillery	B11	4	Ruisui Pasture	C11	3
Shengan Temple and Cihuitang	B12	2	Mingli Hang Gliding Launch Site	C12	3
Gang Tian Temple	B13	2	Liyu Mountain Forest Trail	C14	1
Tzu Chi Cultural Park	B14	2	Luoshan Mud Volcano	C15	1
Tungtamen Shopping District	B15	4	Luoshan Recreation Area	C16	3
Ji'an Keishuin	B16	2	Yuli Xietian Temple	C17	2
Qixintan Scenic Area	B17	1	Antong Hot Springs	C18	3
Beibin Park	B18	1	Chinan National Forest Recreation Area	C19	1
D Park	B19	3	Xiaotianxiang	C2	1
Sakura Trail	B2	1	Yiyuan Resort	C20	3
Shinchen Church	B20	2	National Dong Hwa University	C21	1
Maple Tree Trail	B21	3	Lintian Mountain Forestry Center	C22	2
Pacific Park	B22	1	Hongye Hot Springs	C23	3
Hualien Harbor	B23	3	Ruisui Qinglian Temple	C24	2
Meilun Mountain	B24	3	Wuhe Terrace	C25	3
Tungtamen Night Market	B25	4	Fenglin Principal's Dream Factory	C26	2
General's House	B3	2	Liyu Lake	C27	1
Hualien Railway Cultural Park	B4	2	Chike Mountain	C28	1
Pine Garden	B5	2	Liushishi Mountain	C29	1
Cikasoan Forest Park	B6	3	RuiSui Camp Area	C3	3
Hualien Cultural Creative Industries Park	B7	4	Li Chuan Aquafarm	C30	3
Hualien train station	B8	5	Chi Mei Aboriginal Culture Museum	C31	2

Table 4
Tourism destinations in Haulien, Taiwan (cont.).

Node	ID	Attribute	Node	ID	Attribute
Ruisui Hot Springs	C32	3	Hualien Bay Happy Farm	D8	3
Fei Cui Valley	C33	1	Baqi Gazebo	D9	1
Nan'an Visitor Center	C34	3	Accommodation	E	6
Saoba Ruin	C35	1	Changhua	F	7
Tropic of Cancer Marker_ Rift Valley	C36	3	Pingtung	F	7
Guangfu Sugar Factory	C37	4	Taichung	F	7
Shin Kong Chao Feng Ranch & Resort	C38	3	Tainan	F	7
Tobacco Building	C4, C6, C13	2	Taipei	F	7
Nan'an Waterfall	C5	1	Taitung	F	7
Mataian Wetlands Ecological Park	C7	1	Taoyuan	F	7
Luntianshan	C8	2	Yilan	F	7
Fuyuan National Forest Recreation Area	C9	1	Yunlin	F	7
Ciwidian & Niushan	D1	1	Penghu	F	7
Changhong Bridge	D10	1	Chiayi	F	7
Jiqi Beach Recreation Area	D11	3	Hsinchu	F	7
Tropic of Cancer Marker_Coast	D12	3	Kaohsiung	F	7
Yuedong Recreation Area	D2	1	Keelung	F	7
Shihtiping	D3	1	Miaoli	F	7
Farglory Ocean Park	D4	3	Nantou	F	7
Fanshuliao	D5	1	New Taipei City	F	7
Fengbin Skywalk	D6	1			
Henan Temple	D7	2			

hourly basis (Table 5). The term 'cnt' is used to refer to the number of cell phone activities showing how many users arrive at a destination grid (D-grid) square from an origin grid (O-grid) square. For example, there were 12 users flowing into ID 172632 (destination) from ID 165185 (origin) between 1:00 pm ~ 2:00 pm.

Some steps are required to work the raw OD table into a binary OD matrix for tourism destinations. First, each tourism destination has unique coordinates (longitude and latitude), so Google Earth software can depict the boundaries (Fig. 5). The polygon data overlaps the grids in QGIS system, so we acquire the grid ID of the attractions and the transportations (Fig. 5). The accommodation information was collected from government open data. When the number of hotels and B&B in a grid is more than ten, we define the grid as the accommodation representatives (E). Their density of accommodation facilities is higher than others.

Second, buffer analysis computes the buffer areas with the radii of 250, 500, and 1000 m for each tourism destination (Fig. 6). We seek the appearance of cell phone activities in three spatial dimensions. At the beginning, we view the signal girds within 250 m. If no signal gird is triggered, we will trigger the other dimensions. The nearest grids are the signal girds for the particular tourism destinations.

Third, R software help us data scrubbing. We establish three new tables named TT1, TT2, and TT3. TT1 filters the data from D_gridid consisting of the attractions (A, B, C, D zones), the transportations (B8, B9), Accommodation (E), and the other counties(F). TT2 originating from TT1 keeps the data that O_gridid are the same filtering grids. And then the data of E to E are deleted. Nevertheless, some tourism destinations have business hours. For examples, B4 and B25 share the identical grids (374426, 374427). We discriminate B4 and B25 by the time period (B4: 08:00 am to 4:00 p.m., B25: 5:00 p.m. to 12:00 a.m.) in TT2. The data for B10, B11, B16, C22, C26, C31, C34, D2 and D6 in TT2 are kept from 8 a.m. to 5 p.m. The data for B5, C25 and D4 in TT2 are kept from 8 a.m. to 6 p.m. The data for C30 are kept from 8 a.m. to 8 p.m.

Because the cell phone service provider's market share is 25% (Lin, 2018), the data after preliminary cleanup is multiplied by four to estimate the tourist flow. Table 6 is an example of the final table used to develop the binary OD matrix for tourism destinations.

Fourth, two OD matrices for tourism destinations were created for the weekend (Saturday and Sunday) and two of the weekdays (Wednesday and Thursday). We kept only two of the five weekdays so that the weekend and weekdays would have equal weight in the analysis. Finally, choosing a cutoff value transforms the OD matrix into 1's and 0's. We assume that pairs of tourism destinations with a value lower than 48 on the weekend, or 36 on the weekdays, do not indicate mass tourism demand (Tables 7 and 8).

3.3. Analysis method

Social network analysis is a specialized research method combining sociology with graphing and mathematics. UCINET 6.649 is a software

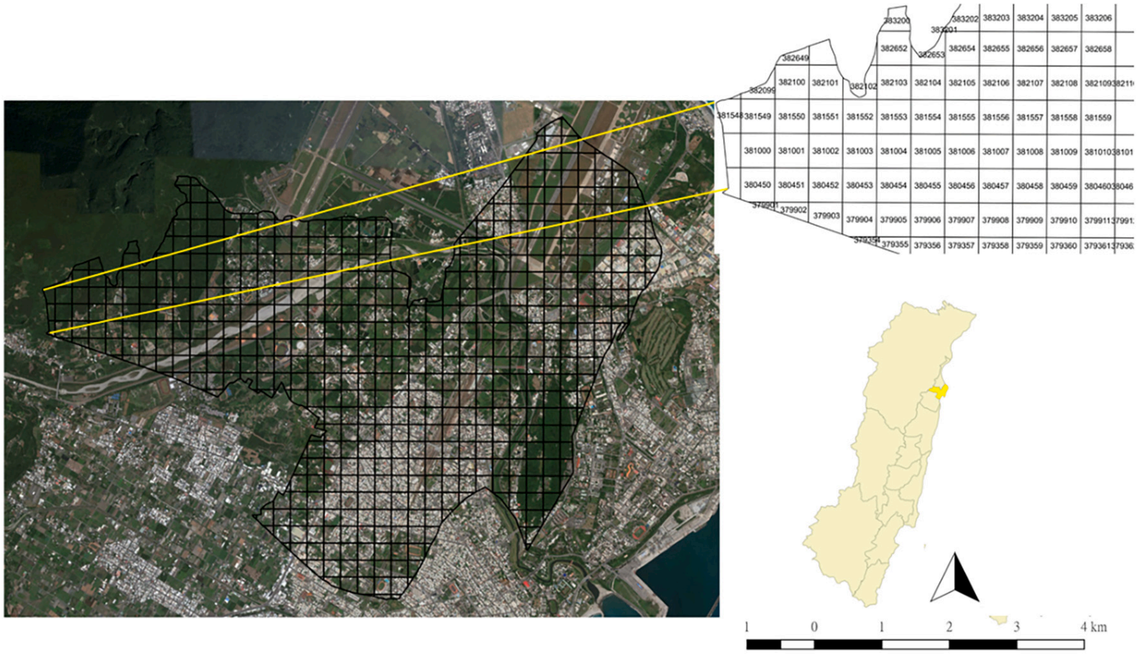


Fig. 4. Grid superimposed on Hualien city.

Table 5
Examples of CDR data; trajectories were recorded on a grid.

Date	Time	O_gridid	D_gridid	cnt
2016/10/21	13	165,185	172,632	12
2016/10/21	14	160,133	200,734	6
2016/10/21	15	245,144	338,477	3

application that executes related indicators for social network analysis. In this study, tourism destinations are regarded as nodes. Tourist mobility (tourist flow) establishes the relationship between the nodes. Trajectories of tourist mobility link pairs of tourism destinations as ties.

3.3.1. Degree centrality (DC)

Degree centrality is the number of direct ties to a given node. Each node has a different degree of connections with others. In a directed network, there are two kinds of degree of connection: in-degree (Eq. (1)) and out-degree (Eq. (3)). Eqs. (2) and (4) calculate the standardized values.

$$C_{D,in} = d(n_i) = \sum_{j}^{n} X_{ij}, \tag{1}$$

$$C'_{D,in}(n_i) = d_{in}(n_i)/N - 1 \tag{2}$$

$$C_{D,out} = d(n_i) = \sum_{j}^{n} X_{ij}, \tag{3}$$

$$C'_{D,out}(n_i) = \frac{d_{out}(n_i)}{N} - 1 \tag{4}$$

X*ij* takes values of 0 or 1, and represents the interaction between *i* and *j*. N is the total number of nodes. For node *i*, there is a tie received from *j*, which means an inward linkage, representing a tourist traveling from *j* to *i*. If node *i* dispatches a tie to *j*, it is an outward linkage, representing a tourist departing *i* to *j*.

A higher in-degree of a tourism destination reflects a great ability to attract tourist flow, because many nodes look for a duct to connect with it. In contrast, a tourism destination with a higher out-degree can deliver its tourists to different other destinations directly, so it is known as an influential actor. This indicator can find the beginning, core and terminal destinations in a spatial network.

3.3.2. Reach closeness centrality (RCC)

Freeman closeness centrality is the original type of closeness cen-

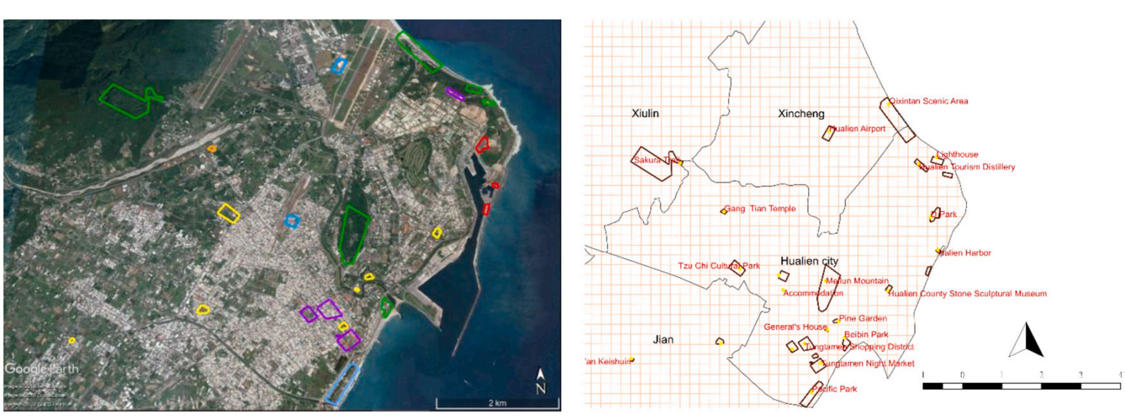


Fig. 5. Geographic ranges of tourism destinations.

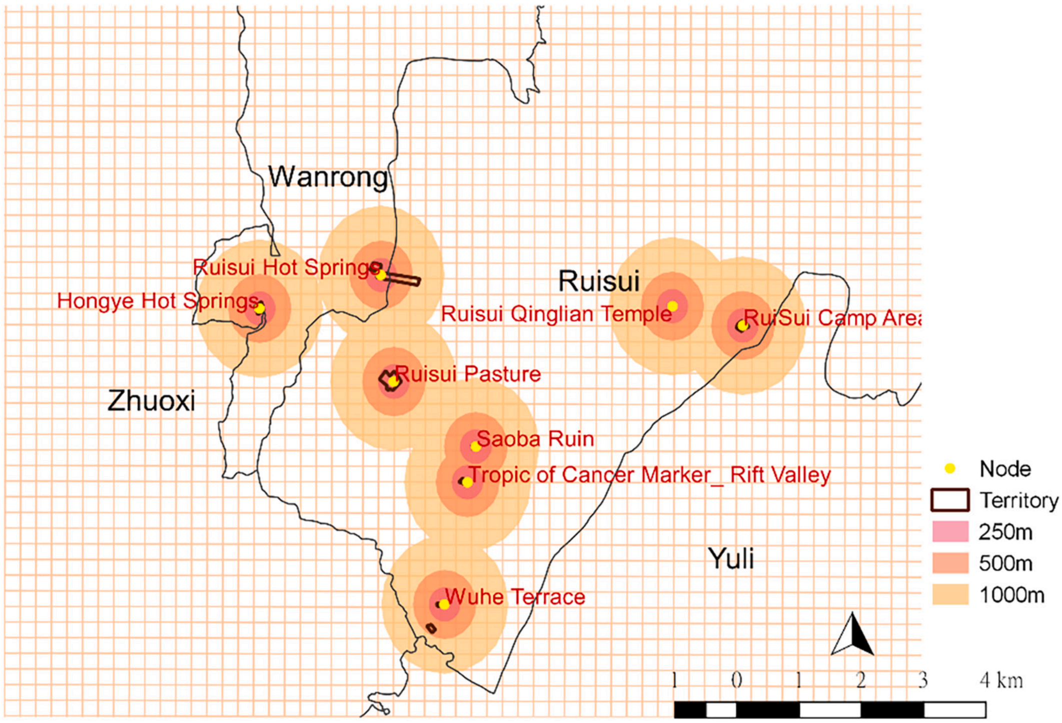


Fig. 6. Buffer polygons of tourism destinations.

Table 6
Examples of CDR data; trajectories were recorded as tourism destinations.

Date	Hour	Origin	Destination	cnt
2016/10/22	11:00	C37	C26	12
2016/10/22	13:00	B17	A	24
2016/10/22	17:00	B15	B25	84

trality, and is defined as the reciprocal of the sum of the geodesic distance from node i to all other nodes j. But Freeman closeness faces a special problem. When the node cannot reach other nodes, there is no geodesic distance generated. The value of total geodesic distance thus becomes smaller (Luo, 2010). RCC is one type of closeness centrality, defined as the sum of the number of nodes that can be reached in k steps divided by k (Eq. (5)). For example, if there are three nodes directly tied, five nodes in two steps, and nine nodes in three steps, then the reach centrality is 1+ 3/1 + 5/2 + 9/3 = 9.5. Standardization is performed by dividing the reach centrality by the largest observed value. This reflects transferability, i.e. the extent to which a tourism destination interacts with all others. Higher in-closeness tourism destinations are easily reached, and thus have more chance to be included in various tourism routes.

$$C_{RC,in} = n_{i0} + \sum_{k=1}^{j} n_{ik} \Big/ k, n_{i0} = 1 \quad (5)$$

3.3.3. Betweenness centrality (BC)

Betweenness centrality evaluates the ability of a node to be an in-

Table 7
Examples of the OD matrix and binary OD matrix of tourism destinations on the weekend.

	A	B10	B11	B12	B13	B14	B15		A	B10	B11	B12	B13	B14	B15
A	0	60	120	132	0	48	540	A	0	1	1	1	0	1	1
B10	48	0	0	48	0	0	324	B10	1	0	0	1	0	0	1
B11	12	24	0	0	0	12	120	B11	0	0	0	0	0	0	1
B12	60	36	12	0	0	24	540	B12	1	0	0	0	0	0	1
B13	0	0	24	0	0	0	0	B13	0	0	0	0	0	0	0
B14	48	0	12	72	0	0	216	B14	1	0	0	1	0	0	1
B15	192	276	84	504	12	264	0	B15	1	1	1	1	0	1	0

Table 8
Examples of the OD matrix and binary OD matrix of tourism destinations on the weekdays.

	A	B10	B11	B12	B13	B14	B15		A	B10	B11	B12	B13	B14	B15
A	0	36	36	24	0	36	312	A	0	1	1	0	0	1	1
B10	0	0	0	60	0	0	300	B10	0	0	0	1	0	0	1
B11	24	12	0	0	0	0	12	B11	0	0	0	0	0	0	0
B12	24	24	0	0	0	156	300	B12	0	0	0	0	0	1	1
B13	0	0	0	0	0	0	0	B13	0	0	0	0	0	0	0
B14	0	48	0	168	0	0	168	B14	0	1	0	1	0	0	1
B15	144	264	12	312	0	180	0	B15	1	1	0	1	0	1	0

termediate (Eq. (6)). When a node does not act as the medium of exchange, the pair of nodes cannot communicate. In other words, a higher betweenness node plays a function as a bridge. If more nodes rely on a particular node to connect with others, the latter node has more power and is a center.

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) \Big/ g_{jk} \tag{6}$$

$$C'_B = C_B / [(N-1)(N-2)] \tag{7}$$

in which *gjk* denotes the number of the shortest paths between node *j* and node *k*. *gjk(ni)* is the frequency of linkages from *j* to k that include *i.* N is the number of nodes in the network. Eq. (7) calculates the standardized values. A tourism destination with higher betweenness is a stopover and can control the distribution of tourist mobility.

3.3.4. Brokerage analysis

The concept of a bridge is very important in a social network. A bridge is also called a broker. A broker carries information from one group to other groups. Gould and Fernandez (1989) invented brokerage analysis. The method focuses on an ego-node and understands each node that acts as a different bridge among different groups. Fig. 7 shows the five kinds of bridges. (1) The coordinator receives information from node A and transmits it to node B in the same group; (2) The consultant acts a bridge between two members in the same group, but the bridge does not belong to this group; (3) The gatekeeper is a channel that keeps the group in contact with outside groups. The gatekeeper controls the inflow of external information; (4) The representative is a spokesman for a group to external groups; (5) The liaison is a bridge between two groups, i.e. between some nodes in one group and some other nodes in another group. (Luo, 2010). This study calculates weighted frequencies to the betweenness center nodes. If nodes B and D both act as bridges between A and C, each gets a score of 1/2(0.5) for the given role (Luo, 2010). Tourism destinations need to be grouped before analysis. We distinguish five groups by geographic zone (A ~ F). Thus, we can further explain which types of bridges the betweenness centers act.

3.3.5. Roles analysis

Roles analysis is based on similarity of the relational pattern. In a social network, position means a collection of individual nodes located in the same roles. This equivalence helps to classify their social roles. There are three approaches: structural equivalence, automorphic equivalence, and regular equivalence. Structural equivalence is the most rigorous. If the overall tendency of the whole network will not change when nodes swap their positions, these nodes are then structurally equivalent and substitutes with each other. They have competitive relations between them (Knoke and Yang, 2008). The CONCOR (convergence of iterated correlations) method based on the correlation coefficient is applied to measure structural positions (Breiger et al., 1975). The formula is as below (Eq. (8)):

$$r_{ij} = \frac{\sum\left(x_{ki}-\overline{x}_{.i}\right)\left(x_{kj}-\overline{x}_{.j}\right)+\sum\left(x_{ik}-\overline{x}_{i.}\right)\left(x_{jk}-\overline{x}_{j.}\right)}{\sqrt{\sum\left(x_{ki}-\overline{x}_{.i}\right)^2+\sum\left(x_{ik}-\overline{x}_{i.}\right)^2} * \sqrt{\sum\left(x_{kj}-\overline{x}_{.j}\right)^2+\sum\left(x_{jk}-\overline{x}_{j.}\right)^2}} \tag{8}$$

r_{ij} is a Pearson correlation. $\overline{x}_i$ is an arithmetic average expressing the number of ties connected directly with node *i* divided by the number of ties nodes *i* could connect with. $\overline{x}_{.i}$ is an inward average. x_{ki} means that the relation from k to *i* is a 0 or 1. x_{ik} and $\overline{x}_{i.}$ just transform into the outward relation. $\overline{x}_{i.}$ is an outward average. x_{ik} indicates that the relation from i to k is a 0 or 1. *j* is the other node. When both *i* and *j* have an inward linkage from k, an outward linkage to k, or do not connect with k, the product is a positive value. The denominator is the square root of variance for i and j multiplied by each other. A new correlation matrix is produced. Then, we repeat the same action again and again. The outcome is an iterated correlation matrix in which each cell is a + 1 or − 1. The CONCOR method utilizes this matrix to distinguish the positions. Nodes can be partitioned into n square groups. The result of grouping for the first time ($n = 1$) separates nodes into two groups. The second time ($n = 2$) separates the nodes of two groups into four groups. The structural positions can be assigned by converging and diverging in varying levels. Next, a blocked matrix assorts the matrix by the positions, and a density matrix is built. Finally, an image matrix is formed by the α criterion. We can thus observe the relationships between the roles (Luo, 2010).

3.3.6. Social group analysis

The goal of the centralities is aimed at the individual node. SNA has developed two perspectives to identify the social groups. From the top-down point of view, the social groups can be found through holes, vulnerabilities, weak spots and solidarity of network such us Component, Bi-component, Lambda sets and Factions. On the contrary, from the bottom-up of view, the rule can observe how the smaller fabrics merge into a bigger network. As a result, overlapping nodes may appear among social groups. Scholars proposed two ways. One is computing

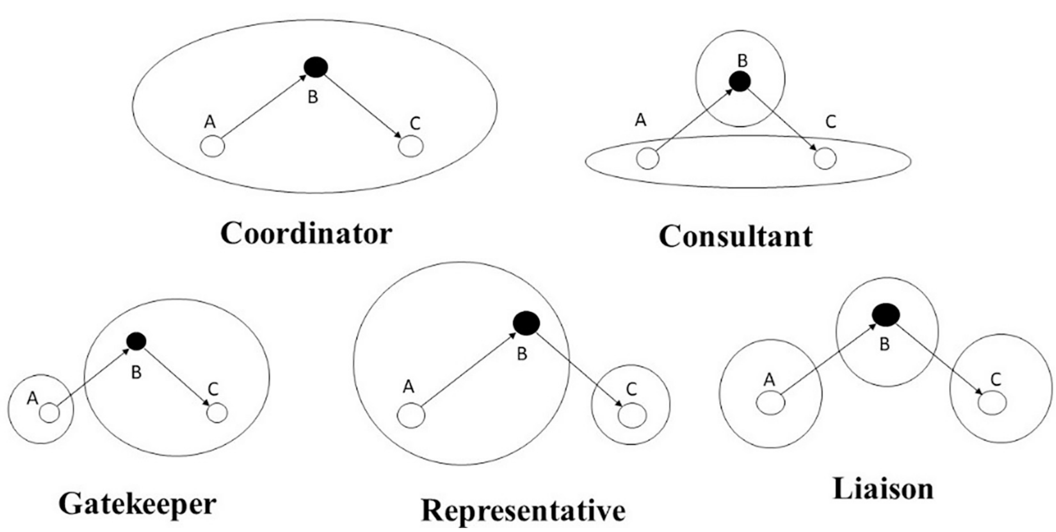


Fig. 7. Five kinds of bridges.

with the degree extent of a node such as k-core and k-plex. The other is applying by the distance (line) such as n-clique and n-clan (Hanneman et al., 2013).

Past researches regarding the spatial network often used n-clique, but n-clique is faced with a physical limitation. All members of an n-clique are connected by the n or less length of geodesic distances (Eq. (9)). *Ns* are the node set. The ties with connecting subgroup members may across external nodes (Wasserman and Faust, 1994). Using N-clique method causes a condition that tourism destination A, B, and C are a subgroup, but B and C may connect through D. N-clique method exists a doubt to present tourism routes in a mobility network. To find effective tourism partnerships, subgroup members should interact by them self. It means partners are several tourism destinations where tourists have been to or will visit. We apply n-clan to find out subgroups as tourism routes and tourism destinations where overlap among subgroups.

$$d(i,j) \leq n, for\ all\ n_i, n_j \in N_s \quad (9)$$

All members of an n-clan are the members of n-clique in which geodesic distance and diameter less than or equal to n within a subgroup. The computing process of n-clan turns a directed network into an undirected network. We think N-clan method is feasible because a one-way tourist flow will establish reciprocity after forming a strategic partnership.

4. Results

4.1. Data description

Fig. 8 shows the trends of cell phone activities on the two selected weekdays and on the weekend by hour. For the raw CDR data, the cell phone activities are 8% less on the weekdays than on the weekend. The cell phone activities start at 6 am and peak at 7 am for the weekdays. The curves diminish during office time and rebound at mealtimes. The peak time on the weekend is after 8 am, and calls continue energetically until 7 pm. The cleaned-up data are the CDR data for which origins and destinations are the tourism destinations. The cell phone activities of tourists have two peaks, which are 8 am to 11 am, and 3 pm to 7 pm. The cell phone activities of tourists on Saturday morning are less than those on Sunday morning, but the cell phone activities Saturday evening are much more than those on Sunday evening. This means those tourists reach Hualien Saturday morning and leave before Sunday evening. The outcome shows that the tourist behavior is different from the general pattern.

4.2. Centrality analysis

4.2.1. Generating markets and tourism gateways of Hualien

Yilan, Taitung, Taipei Metropolitan Area, Nantou, and Taoyuan are main markets generating tourism to Hualien (Table 9). These tourists may be the inhabitants of those areas, or may be tourists to those areas. The four first are the neighbor markets. Yilan, Taitung, and Nantou are the immediate neighbors of Hualien. However, Taipei Metropolitan Area is the most populous area in Taiwan, and connects to Hualien through Yilan. We discovered that the tourists from the neighboring markets go to tourism destinations dispersed throughout Hualien for their first stops. People who come from other counties which are in western Taiwan have fewer direct visits to Hualien, and they tend to go to same tourism destinations for their first stops (Figs. 9 and 10).

The nodes not linking with other counties are excluded. We apply the rest of the nodes to use degree and degree centrality for two periods. The mean degree and the mean flow of each node are standardized into Z scores. Euclidean distance and Ward's method are computed to classify levels of gateways. The nodes are classified as four levels: from low scores (4, both lower degree mean and flow mean) to high scores (1, both higher degree mean and flow mean). Cluster 1 and Cluster 2 are main gateways (Table 10). The number of gateways on the weekend are more than those on the weekdays. A, E, B15, B23, B8, C17 are always gateways. B12, B4, and C37 appear on the weekend. B14 is rejected, because of the lower flow mean. The relationship between generating markets and tourism destinations can be explained by the network graph (Figs. 9 and 10). A is the gateway for northern and central Taiwan. C17 connects with southern Taiwan. E (accommodation) and B8 (train station) are service components. B15 is the city center in Hualien.

4.2.2. Centers of tourism destinations

Generating markets are an influential factor for outcomes; therefore we have to incorporate them into the regional network analysis. Higher-degree tourism destinations are best-known attractions. Normally, these tourism destinations may be the first places strangers visit. They are the driving forces for tourists to travel. Out degree and In degree in Table 11 are the standardized degree centrality. A high-degree centrality destination is defined as one with Out degree and In degree greater than 0.181. Tourism destinations E, B15, B23, A, B4, B22, B24, B12, B18, and B7 have a high degree of centrality at all times. B25, B17, C37 and C30 have high-degree centrality on the weekend, and C38 joins these on the weekdays (Figs. 11 and 12).

There are different largest values of reach closeness centrality between the weekend and the weekdays, so we do not use standardization

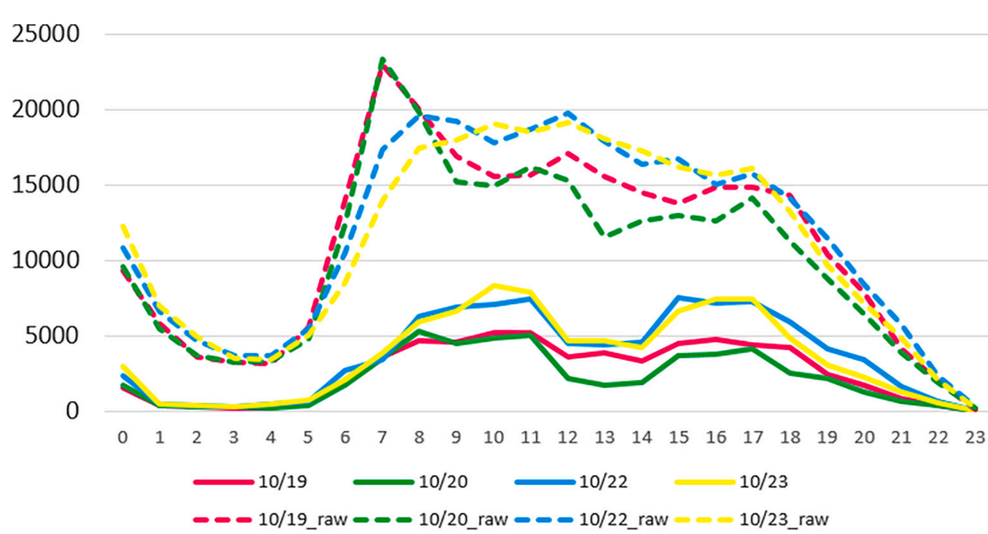


Fig. 8. Comparison of cell phone activities between the raw data and the cleaned-up data.

Table 9
Degree centrality for the generating markets.

	Weekend						Weekdays					
	Out degree	In degree	Outflow	Inflow	flow mean	%	Out degree	In degree	Outflow	Inflow	flow mean	%
Yilan	22	25	8316	8700	8508	34.26	17	19	3120	4740	3930	28.18
Taitung	29	26	7356	6840	7098	28.58	27	22	4680	3300	3990	28.61
Taipei	16	14	4104	3900	4002	16.12	19	7	2640	2508	2574	18.46
New Taipei	12	11	2244	3072	2658	10.70	10	11	1080	2004	1542	11.06
Nantou	2	1	828	456	642	2.59	1	1	504	360	432	3.10
Taoyuan	5	7	648	1164	906	3.65	4	6	348	1176	762	5.46
Keelung	3	2	276	252	264	1.06	0	1	0	180	90	0.65
Kaohsiung	1	3	252	192	222	0.89	1	2	96	216	156	1.12
Taichung	3	2	252	216	234	0.94	2	3	192	252	222	1.59
Pingtung	2	1	156	84	120	0.48	2	2	192	72	132	0.95
Hsinchu	2	1	120	156	138	0.56	2	1	96	60	78	0.56
Changhua	0	0	0	0	0	0	0	1	0	36	18	0.13
Chiayi	0	0	0	0	0	0	0	0	0	0	0	0
Miaoli	0	1	0	84	42	0.17	0	0	0	0	0	0
Tainan	0	0	0	0	0	0	1	0	36	0	18	0.13
Yunlin	0	0	0	0	0	0	0	0	0	0	0	0
Total			24,552	25,116	24,834				12,984	14,904	13,944	

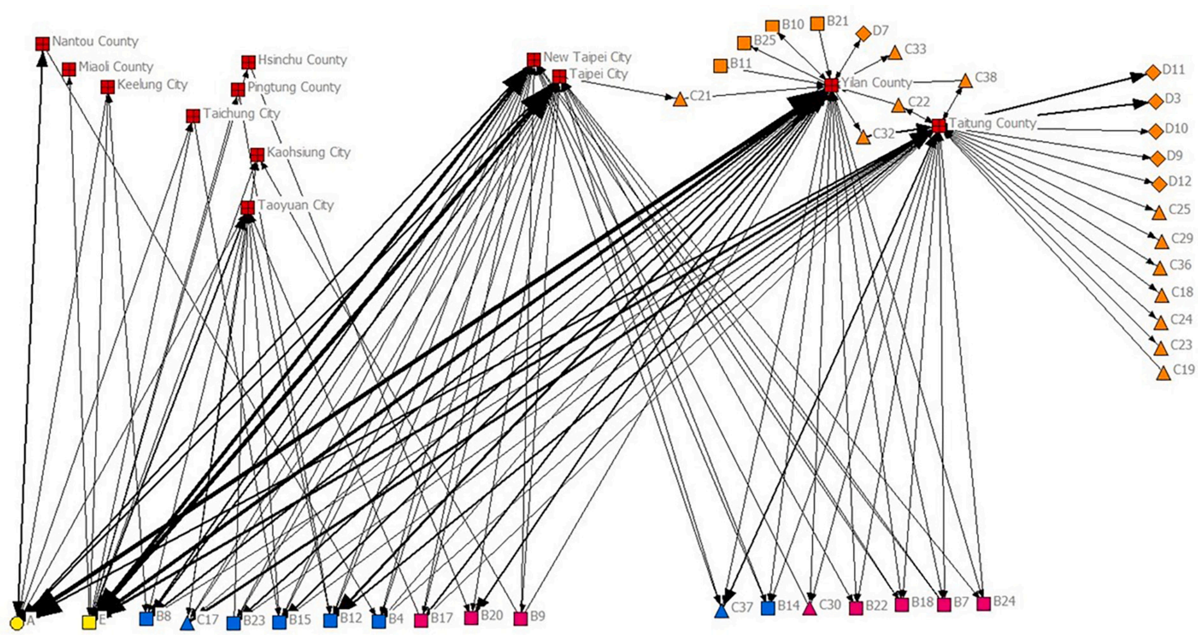


Fig. 9. Network graph of the gateways on the weekend.
Notes: yellow, cluster 1; blue, cluster 2; pink, cluster 3, orange, cluster 4; red, generating markets. Circle, zone A; square, zone B; triangle, zone C; diamond, zone D.

to compare them. The high in-reach destinations (IndwReach >40) are E, B15, B23, B4, and A at all times. B12, C37, and B17 are added on the weekend and B7 is added on the weekdays.

An intermediate tourism destination can guide the tourist distribution to balanced development. That is to say, tourists preparing to visit tourism destination j from i have to pass through the intermediate destination k. When they arrive at k, they might change their mind and go to destination l, not j. The variable betweenness in Table 11 is standardized. The high betweenness centrality tourism destinations (value >1) are E, B15, B23, A, B4, C17, C32 and C38 at all times. In addition, B12 and C22 show this feature on the weekend, while B17, C36 and D12 do so on the weekdays.

4.3. Brokerage advantages of tourism destinations

Table 12 shows the brokerage scores of the betweenness centers. These centers with higher brokerage scores are identified as bridges. They usually get multiple values as different bridges, e.g. one center could be a bridge between nodes A and C, but act as a different bridge between nodes D and E, with each having a different brokerage score value. At the same time, we can know which zones betweenness centers connect. Table 13 shows two examples, E and A.

Destinations E, B15, B23 and A act as steady bridges at all times. E has the highest brokerage score value among all types of bridges. E is the coordinator node, and lets two or more unconnected destinations belonging to the same zone spark the relation (B → E → B). E also acts as a gatekeeper, steering the tourist flow from other zones into the city (C/D/F → E → B), as a representative steering the tourist flow out of the city into other zones (B → E → C/D/F), as a consultant promoting intra-mobility to given zones (C → E → C, D → E → D, F → E → F), and as a liaison as an intersection node among zones (C ↔E ↔D; A/C/D ↔E ↔F). B15 acts as five types of bridges, like E, and has the second highest brokerage score value. A acts as two types of bridges, a consultant (B → A → B; F → A → F) and liaison (B/C/D ↔A ↔F; B ↔A ↔C; D → A → B). But A is not a liaison between F and D on the weekdays. B23 acts as four types of bridges (all but consultant).

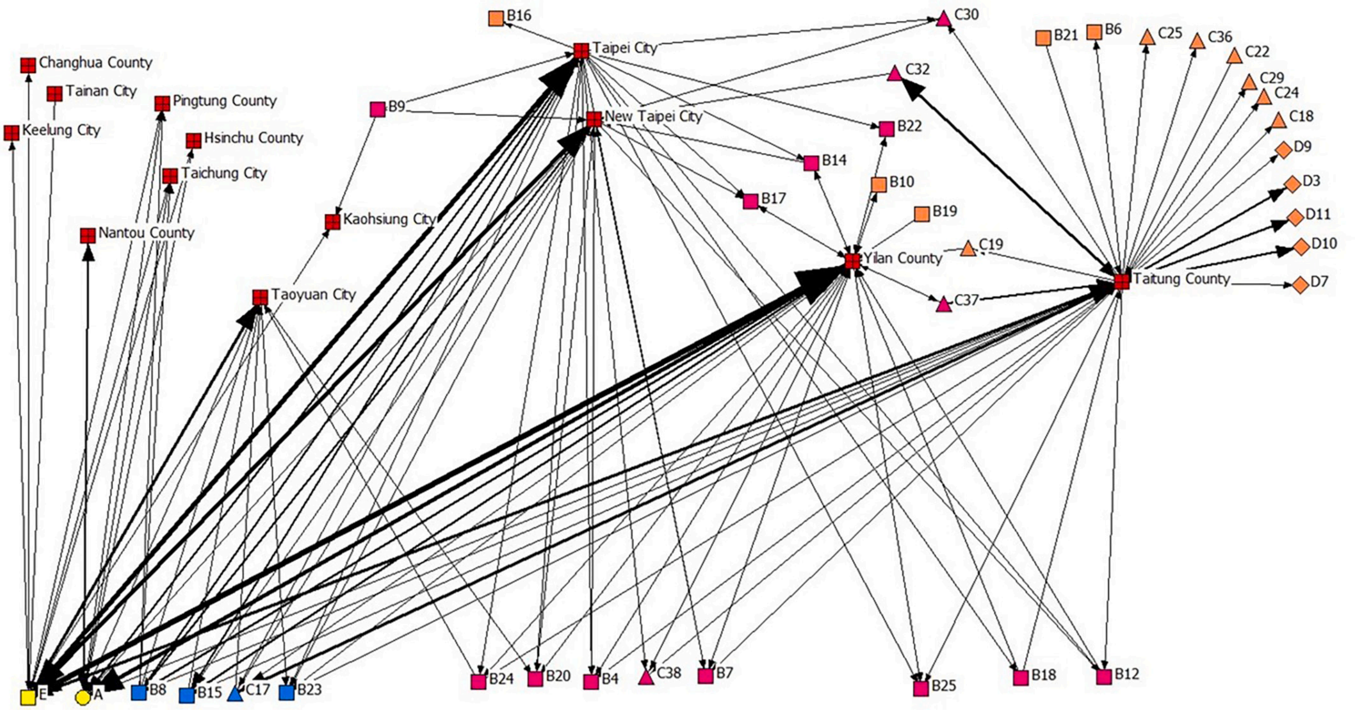


Fig. 10. Network graph of the gateways on the weekdays.

Table 10
Degree centrality for the gateways.

Node	Weekend							Weekdays						
	Out degree	In degree	Z-score	Out flow	In flow	Z-score	Cluster	Out degree	In degree	Z-score	Out flow	In flow	Z-score	Cluster
A	7	8	2.352	3624	4644	2.662	1	6	9	2.645	1944	2712	2.337	1
B12	3	5	0.742	1152	1692	0.607	2	1	3	−0.038	60	168	−0.289	3
B14	3	4	0.512	216	216	−0.307	2	2	2	−0.038	84	144	−0.289	3
B15	4	6	1.202	816	1272	0.321	2	5	4	1.182	612	492	0.231	2
B17	2	4	0.282	540	432	−0.102	3	1	3	−0.038	108	144	−0.275	3
B18	2	3	0.052	216	216	−0.307	3	2	1	−0.281	132	84	−0.296	3
B20	2	3	0.052	288	696	−0.098	3	2	4	0.45	228	276	−0.125	3
B22	2	3	0.052	132	240	−0.329	3	1	2	−0.281	60	156	−0.296	3
B23	5	4	0.972	792	792	0.13	2	3	5	0.938	180	480	−0.033	2
B24	3	2	0.052	288	120	−0.316	3	3	2	0.206	180	156	−0.225	3
B25								0	3	−0.281	0	144	−0.339	3
B4	3	5	0.742	300	720	−0.084	2	2	4	0.45	204	252	−0.154	3
B7	3	2	0.052	204	180	−0.325	3	2	3	0.206	72	144	−0.296	3
B8	6	6	1.662	1728	1104	0.602	2	7	4	1.669	1188	324	0.473	2
B9	4	1	0.052	204	48	−0.375	3	3	0	−0.281	240	0	−0.282	3
C17	6	5	1.432	1560	1260	0.598	2	4	4	0.938	852	1056	0.708	2
C30	3	2	0.052	216	228	−0.302	3	1	2	−0.281	72	72	−0.339	3
C32								2	1	−0.281	456	768	0.302	3
C37	4	4	0.742	972	1092	0.312	2	2	2	−0.038	348	240	−0.075	3
C38								2	3	0.206	108	132	−0.282	3
E	11	10	3.732	8580	6312	5.171	1	11	10	4.108	6624	3120	5.355	1

B17 is a coordinator (B → B17 → B) on the weekends and a representative (B → B17 → C/D) on the weekdays. B4 acts as four types on the weekend including coordinator (B → B4 → B), gatekeeper (C/D/F → B4 → B), representative (B → B4 → C/D/F) and liaison (C ↔B4 ↔D/F; D ↔B4 ↔F), but just two types on the weekdays, coordinator and gatekeeper. C17 acts as two types on the weekend, gatekeeper (C → C17 → B/F) and representative (C → C17 → F), but increases its role as coordinator (C → C17 → C) on the weekdays. C32 acts as three types on the weekend, coordinator (C → C32 → C), gatekeeper (B/F → C32 → C) and representative (C → C32 → B/F), but just as a coordinator on the weekdays. C37 acts as three types on the weekend, coordinator (C → C37 → C), gatekeeper (A/B/D/F → C37 → C) and representative (C → C37 → A/B/D/F), but as two types on the weekdays, coordinator and representative (C → C37 → B/F). C38 is of the representative type (C → C38 → B/F) on the weekend and acts as two types on the weekdays, coordinator (C → C38 → C) and gatekeeper (B/F → C38 → C). B12 has two roles on the weekend, that of gatekeeper (C/F → B12 → B) and representative (B → B12 → C). C22 is a coordinator (C → C22 → C) on the weekend, while C36 serves as coordinator (Cordinat on the weekdays.

However, we discover these above betweenness centers have higher brokerage values in the zones where they are located, too. The destinations in zone D do not act as any center. Understanding the bridges of weaker zones is still necessary. D11 is a coordinator (D → D11 → D) on

Table 11
Results of centralities for tourism destinations.

Node	Weekend						Weekdays					
	Out flow	In flow	nOutdeg	nIndeg	Indw Reach	nBetweenness	Out flow	In flow	nOutdeg	nIndeg	Indw Reach	nBetweenness
E	31,212	31,176	0.625	0.611	52	20.205	23,580	22,272	0.611	0.597	53.333	26.516
B15	17,928	18,420	0.472	0.514	48.5	6.157	10,848	11,196	0.403	0.486	49	6.974
B23	11,124	9660	0.403	0.361	42.667	3.212	8412	8124	0.403	0.375	44.833	6.789
A	7716	7224	0.403	0.319	41.333	4.264	4620	4440	0.306	0.264	40.5	4.685
C37	4464	4764	0.319	0.306	41	4.07	1656	2184	0.167	0.194	38.667	3.396
B4	4800	8052	0.264	0.458	46.5	2.581	3216	3612	0.236	0.264	40.667	1.75
B22	3684	3888	0.25	0.236	38	0.521	2916	3048	0.194	0.208	37.083	0.217
B12	4428	5400	0.222	0.347	42.167	1.387	2424	2340	0.194	0.236	39.5	0.466
B17	4608	4440	0.222	0.292	40.167	0.779	1728	1500	0.153	0.167	35.583	1.516
B7	2352	2604	0.222	0.264	39	0.361	1668	2124	0.194	0.25	40	0.623
B24	3084	3108	0.222	0.208	35.667	0.266	2796	2772	0.222	0.208	36.917	0.445
C30	1656	1428	0.222	0.194	36.833	0.692	456	420	0.111	0.111	35	0.255
B18	2724	2628	0.208	0.181	36	0.149	2424	2220	0.222	0.208	37.75	0.977
B25	6876	5724	0.194	0.208	35.5	0.328	3264	2724	0.167	0.181	37.333	0.316
C17	2856	2784	0.167	0.181	36.333	1.903	2064	2388	0.167	0.181	37.833	3.515
C38	1716	1368	0.153	0.181	36.167	1.524	1524	912	0.181	0.194	38.5	4.241
C32	2808	2364	0.153	0.167	35.667	3.282	1620	1944	0.111	0.111	35.5	2.946
C22	2136	1020	0.153	0.083	32.667	1.376	1092	948	0.042	0.083	29.417	0.499
C36	924	1020	0.069	0.083	29.333	0.4	828	792	0.056	0.069	28.917	1.272
D12	144	240	0.014	0.028	25.583	0	96	144	0.014	0.028	26.417	1.213
Taitung	7356	6840	0.403	0.361	42.833	14.983	4680	3300	0.375	0.306	42.833	17.544
Yilan	8316	8700	0.306	0.347	41.5	1.893	3120	4740	0.236	0.264	39.917	1.254
Taipei	4104	3900	0.222	0.194	35.833	0.739	2640	2508	0.264	0.097	33.917	0.561
New Taipei	2244	3072	0.167	0.153	34.333	0.378	1080	2004	0.139	0.153	36.083	0.643

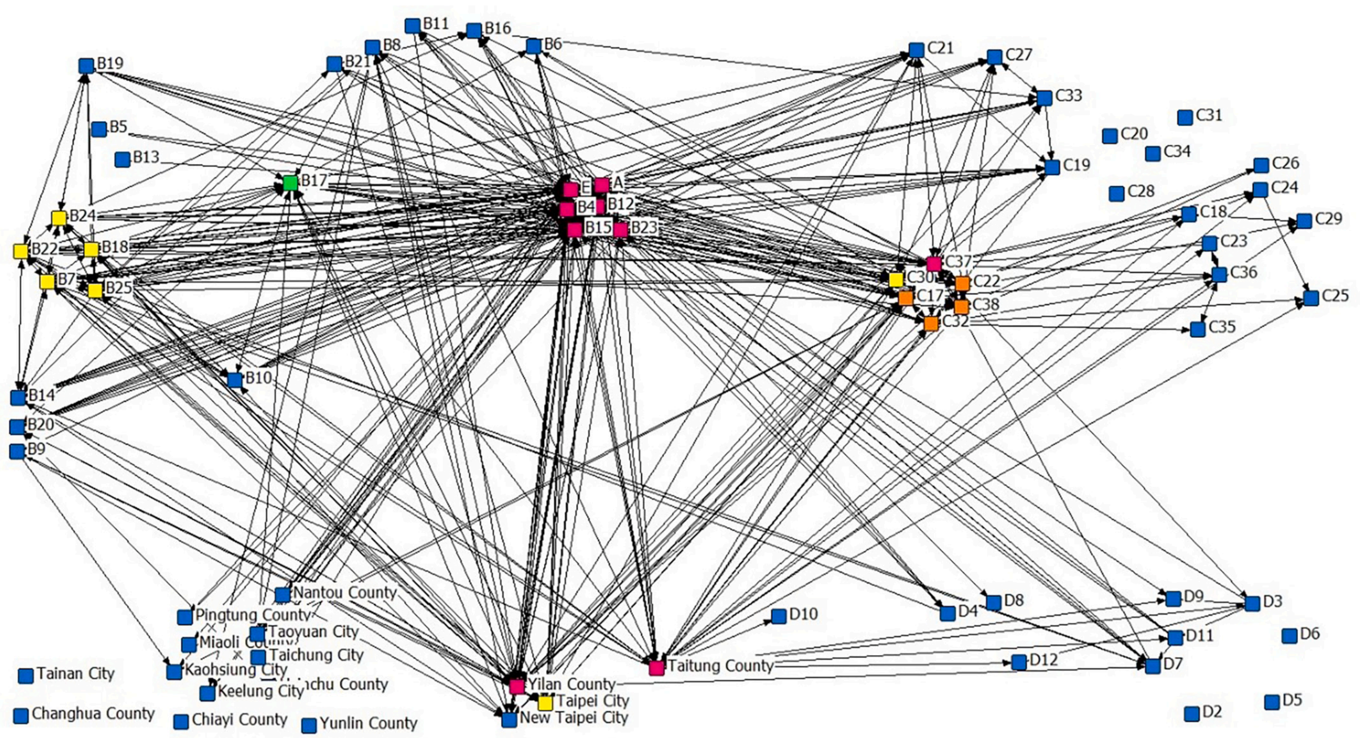


Fig. 11. Network graph of the centralities on the weekend.
Notes: pink, three kinds of centers; green and gray, two kinds; yellow, one kind.

the weekend, but a representative (D → D11 → B) on the weekdays. D12 serves as representative (D → D12 → F) on the weekdays. D3 is a gatekeeper (C → D3 → D) on the weekend. D7 acts as representative (D → D7 → B) on the weekend but as consultant on the weekdays (B → D7 → B).

4.4. Roles of tourism destinations

We split the iterated correlation matrices four times to derive sixteen groups. These groups can be converted into six roles. Table 14 is the density of blocks between roles. The density of a whole network is often used as cut-offs ($\alpha = 0.1$ for two time periods) to transform the density matrices into the image matrices. If we use 0.1, we will lose the information about the weaker blocks. A value which is greater than 0.1 demonstrates a strong relationship. A value between 0.1 and 0.08 has a weaker relationship. So, $\alpha = 0.08$ is the cut-off for our research. A density of more than 0.8 becomes 1; one less than that becomes

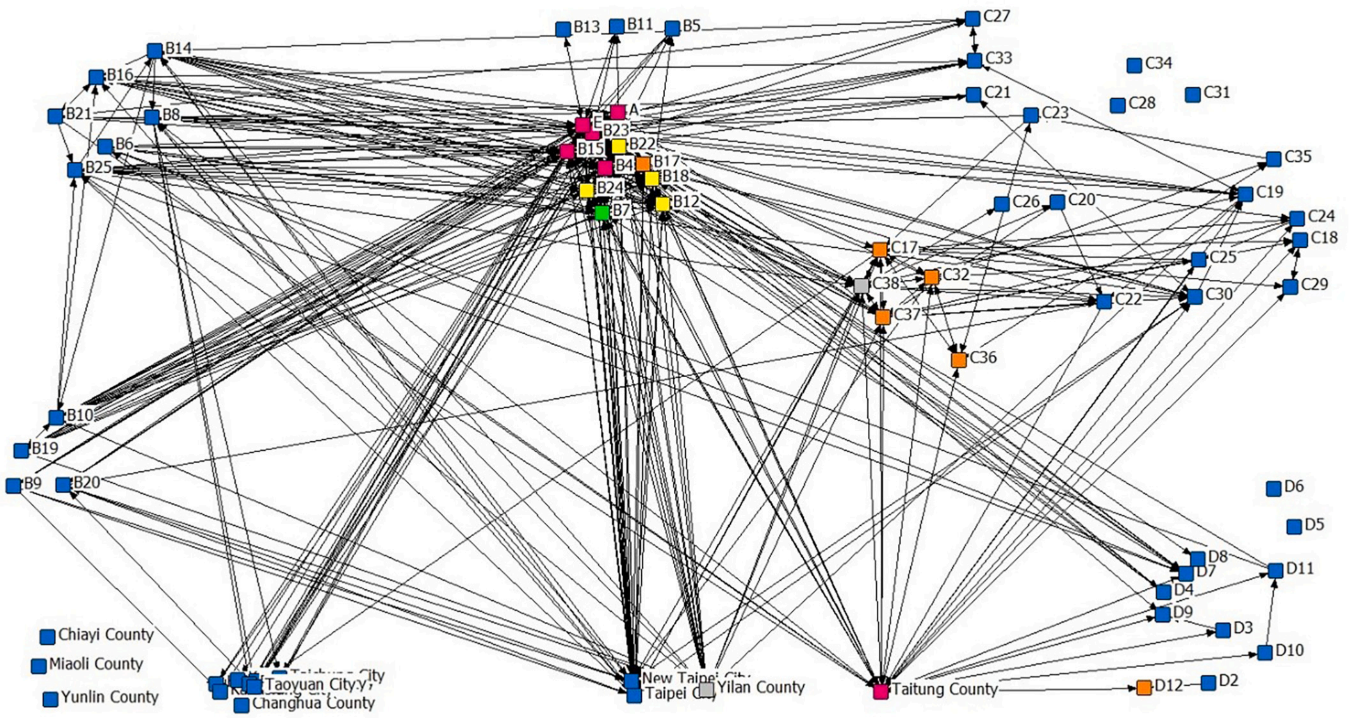


Fig. 12. Network graph of the centralities on the weekdays.

Table 12
Results of brokerage scores for the betweenness centers.

Node	Weekend					Weekdays				
	Coordinator	Gatekeeper	Represent	Consultant	Liaison	Coordinator	Gatekeeper	Represent	Consultant	Liaison
E	82.812	175.804	216.526	85.153	202.6	109.67	226.888	247.533	105.819	189.402
B15	43.245	68.16	54.05	14.698	35.707	39.175	61.512	64.008	13.835	43.419
B23	20.482	16.246	38.827	4.42	18.231	47.389	41.871	66.035	4.574	30.06
A	0	0	0	30.453	95.747	0	0	0	38.261	78.665
B4	15.242	35.604	12.023	3.678	14.152	10.117	16.941	2.634	0.658	2.626
C37	31.26	29.469	33.893	2.848	9.847	17.65	8.251	20.608	0.202	0.494
B12	6.862	8.845	11.513	3.288	4.144	3.853	3.55	3.924	0.269	0.51
B17	11.07	6.975	4.515	0.943	1.252	1.216	1.387	13.374	0.178	5.417
C32	17.283	10.458	8.951	0.077	0	11.033	2.333	5.617	0.125	0
C17	5.667	13.293	17.367	4.767	1.143	22.917	13.093	21.45	1.983	0
C38	4.783	4.719	8.344	0.592	0.817	13.867	30.405	7.976	0.471	0.45
C22	5.31	2.143	1.833	0.392	0.333	1.75	1.95	1	0	0.7
C36	6.167	0.5	0	0	0	3.667	1.5	0	0	0
D3	0	1.333	0	0	0	0	0	0	0	0
D7	0	0	1.6	0.168	0.271	0	0	0	0.541	0.587
D11	0.5	0.25	0.667	0	0	0	0	1	0	0
D12	0	0	0	0	0	0	0	1	0	0
Yilan	0	0	0	23.635	41.748	0	0	0	14.185	20.1
Taitung	0	0	0	76.721	294.036	0	0	0	59.195	270.369

Table 13
Examples of group to group brokering for tourism destinations.

Destination A												Destination E											
Weekend						Weekdays						Weekend						Weekdays					
	A	B	C	D	F		A	B	C	D	F		A	B	C	D	F		A	B	C	D	F
A	0	0	0	0	0	A	0	0	0	0	0	A	0	2	1	1	5	A	0	2	2	2	5
B	0	9	9	2	21	B	0	13	3	0	17	B	2	83	57	28	129	B	4	110	75	25	144
C	0	3	1	1	4	C	0	6	0	0	4	C	1	53	15	13	59	C	3	100	31	12	67
D	0	4	2	0	3	D	0	0	0	0	0	D	1	32	19	5	32	D	1	18	10	2	16
F	0	28	14	3	20	F	0	42	7	0	25	F	2	89	50	20	65	F	2	107	55	16	73

0 (Table 15). Table 16 contains the destination members of six roles.

Fig. 13 is a simple graph to show the relations between six roles. Role 1 and Role 2 have strong intra-connection. Role 5 has relatively low intra-connection, and Role 3 happens on the weekend. Role 4 contains lots of the other counties, and the members do not always directly link between each other, so it is reasonable that Role 4 does not have intra-connection. Role 1 sends tourists to and receives tourists from Roles 2, 3 and 4. Role 2 sends and receives with Roles 1 and 3. Beside Roles 1 and 2, Role 3 has a weak connection with Role 4. Role 5 relies on Role 4 to establish contact with the outside. Role 6 is almost independent. Fig. 13 shows how Roles 1, 2, and 3 concentrate on the north and Roles 5, and 6 concentrate on the south. Role 4 has different ranges for the two periods. We can thus know that several destinations become more energetic on the weekend. The result displays some tourism destination act different roles between the weekend and the weekdays.

4.5. Classification of tourism destinations

We combine the above-mentioned results to identify the functions of tourism destinations. There are six kinds of new functions for tourism destinations in Hualien. Figs. 14 and 15 show the relations between different functions. In the figures, the functions and the roles are differentiated by colors and shapes: (1) The Shape of nodes presents the original attributes based on resource. (2) Color of nodes: red, Role1; yellow, Role2; green, Role3; blue, Role 4; purple, Role5; gray, Role6. (3) Color of border: red, distribution center; yellow, hub destination; green, passageway destination; blue, ordinary destination; purple, secondary destination.

4.5.1. Tourist flow distribution center

Destinations A, B4, B15, B23, and E can be identified as the distribution centers. A distribution center contents the largest tourist flow, has the highest degree, reach closeness, and betweenness centralities. The distribution center destination plays as a multi-functional bridge. There are only few destinations would qualify as distribution centers in a region. These destinations connect tourist generating markets and most destinations located within the region. They have a power to improve the regional tourism.

4.5.2. Hub destination

Destinations B7, B12, B17, B18, B22, B24, B25, and C37 are act as hub destinations. A hub destination contents the large tourist flow, has high degree and reach closeness centralities. A hub is the best-known destination the tourists can arrive easily. The interesting point for C37 (Guangfu Sugar Factory) plays different functional role: on weekend it plays as a hub but on weekdays it is a passageway destination.

4.5.3. Passageway destination

Destinations C17, C22, C32, C38, B8, B9, B20, and D7 act as passageway destinations. C32 (Ruisui Hot Springs), C17 (Yuli Xietian Temple) and C38 (Shin Kong Chao Feng Ranch & Resort) have a high betweenness centrality. C32 acts as a coordinator within East Rift Valley. It makes the tourists visit some secondary touring destinations. C17 brings people from the south of Taiwan and affects the tourist mobility in East Rift Valley. C22 (Lintian Mountain Forestry Cente) acts as a passageway like C32 but only on the weekend. C38 brings people from the north of Hualien into East Rift Valley. B8 (Hualien train station) and B9 (Hualien Airport) are transport facilities. They bring tourists from the generating markets into this region. However, they have slightly influence on the tourist trajectory. Those tourists more rely on distribution centers and hub destinations to decide their travel trajectory in Hualien. D7 (Henan Temple) is a weak passageway destination because it is the only entrance that connects East Coast and the north of Hualien. A passageway destination is a gateway or a bridge in a region.

4.5.4. Touring destination

The touring destinations are classified into three levels: the ordinary, secondary, and the sparse. An ordinary touring destination has a high possibility to be the next destination after tourists visiting a distribution center or a hub. Even the ordinary touring destinations are spread widely, but they could attract certain amount of visitors. A secondary touring destination has the relatively lower centralities and visitors. A sparse destination has the lowest centralities and the lowest number of visitors.

The ordinary touring destinations are identified. They are B6, B10, B11, B14, B16, B19, B21, C19, C21, C27, C30, C33, D3, D4, D9, and D11.

The secondary touring destination are B5, B13, C18, C20, C23, C24, C25, C26, C29, C35, C36, D8, D10, and D12.

The sparse destinations are C28, C34, C31, D2, D5, and D6.

4.6. Tourism social groups

In this section, we present the tourism social groups based on the travel data we collected. The purpose of this section is to demonstrate the analytical results into potential tourism managerial applications. Using the 2-clan method, this paper clusters the destinations into tourism social groups for the weekend and weekdays. A 2-clan is an n-clique in which all visited destinations are connected by paths of length ≤ 2. That is, a social group is a group of destination related closely respect to tourists' travel trajectory.

Excluding the generating market, we find six social groups on the weekend. The destinations of a social group do not have to locate in the same zone. Each social group in a set of destinations that is easy to visit within a few steps. The paths in a social group show the trajectory of tourists that provides business information for tourism industry. Social group 1 implies a regional competition among four zones (A, B, C & D). The destinations of Hualien city (B) are all play as hub destinations. Social group 2 covers the middle and the north of Hualien. The destinations within this district competing each other strongly, such as B19, B10 and B11 because they are closely located on the same road. Social groups 3, 4 and 5 are relatively coordinate, but it still reveals the interior competition within the zone. In social group 3, B6, B11 and B21 are competitors. The similar insights would be investigated with more information (the business type, location, etc.) about these destinations. For example, in social group 5, C18 and C32 are both famous for their hot spring, but they do not have the direct connections. For example, a tourist leaves C29 and wants to visit a hot spring, then he/she has two selections which are C18 with one step or C32 with two steps (through C17).

The results of social group analysis can combine with GIS (Fig. 16). It

Table 14
Density matrices between six roles.

Weekend	1	2	3	4	5	6	Weekdays	1	2	3	4	5	6
1	0.867	0.847	0.548	0.315	0.015	0	1	1	0.829	0.45	0.42	0.059	0
2	0.792	0.697	0.125	0.06	0	0	2	0.657	0.604	0.134	0.036	0.008	0
3	0.631	0.161	0.115	0.06	0.026	0	3	0.5	0.098	0	0.081	0.037	0
4	0.333	0.074	0.083	0.059	0.086	0.005	4	0.4	0.046	0.056	0.042	0.053	0
5	0.015	0	0.013	0.086	0.082	0	5	0.059	0.004	0.015	0.059	0.107	0
6	0	0	0.006	0	0	0	6	0	0	0	0	0.007	0

Table 15
Block images between six roles.

Weekend	1	2	3	4	5	6	Weekdays	1	2	3	4	5	6
1	1	1	1	1	0	0	1	1	1	1	1	0	0
2	1	1	1	0	0	0	2	1	1	1	0	0	0
3	1	1	1	0	0	0	3	1	1	0	1	0	0
4	1	0	1	0	1	0	4	1	0	0	0	0	0
5	0	0	0	1	1	0	5	0	0	0	0	1	0
6	0	0	0	0	0	0	6	0	0	0	0	0	0

Table 16
Roles of tourism destinations.

Two periods											
1	A	B15	B23	E							
2	B10	B12	B14	B17	B18	B19	B22	B24	B25	B7	Yilan
3	B20	C38	D4	D7	D8						
4	B13	B9	C27	C33	Hsinchu	Kaohsiung	Keelung	Pingtung	Taichung	Taitung	Taoyuan
5	C18	C23	C24	C25	C29	C35	C36	D10	D12	D3	D9
6	C28	C31	C34	D2	D5	D6	Chiayi	Yunlin			
Downgrade: weekend → weekdays											
1 → 2	B4										
1 → 3	B8										
3 → 4	B11	C19	C21	C30	C37	Nantou	New Taipei				
4 → 5	C17	C22	C32	D11							
4 → 6	Miaoli										
Upgrade: weekend → weekdays											
2 → 1	Taipei										
3 → 2	B16	B6									
4 → 3	B21	B5									
6 → 4	Changhua	Tainan									
6 → 5	C20	C26									

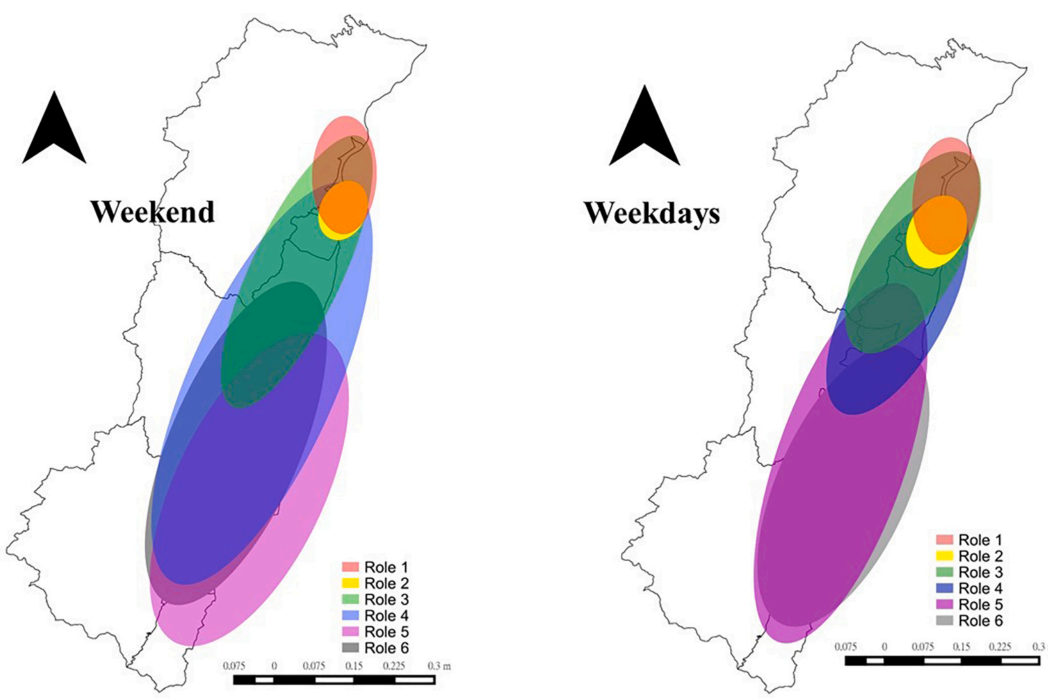


Fig. 13. Maps of destination roles.

generates the tourism strategy for route planning. Social group 1 and 3 are the route-strategy planned to visit the sequential destinations along the main roads. Social group 2 is the landmark-strategy. The tourists use Hualien city as a hub to visit the common touring destinations nearby. Social group 4 and 5 are the tour plan for the tourists prefer specific destinations. Social group 4 is very clear that the tourists go to East Rift Valley and visits the secondary destinations. The tourists in social group 5 go to Taroko and East Rift Valley. Two groups are not interested in city attractions. Social group 6 is a bypassing trip. The tourists are only use Hualien area as an intermedia in their whole trip. They may stay overnight but don't visit any Hualien's attraction destination.

For the weekdays, the data result in eight social groups. Those social

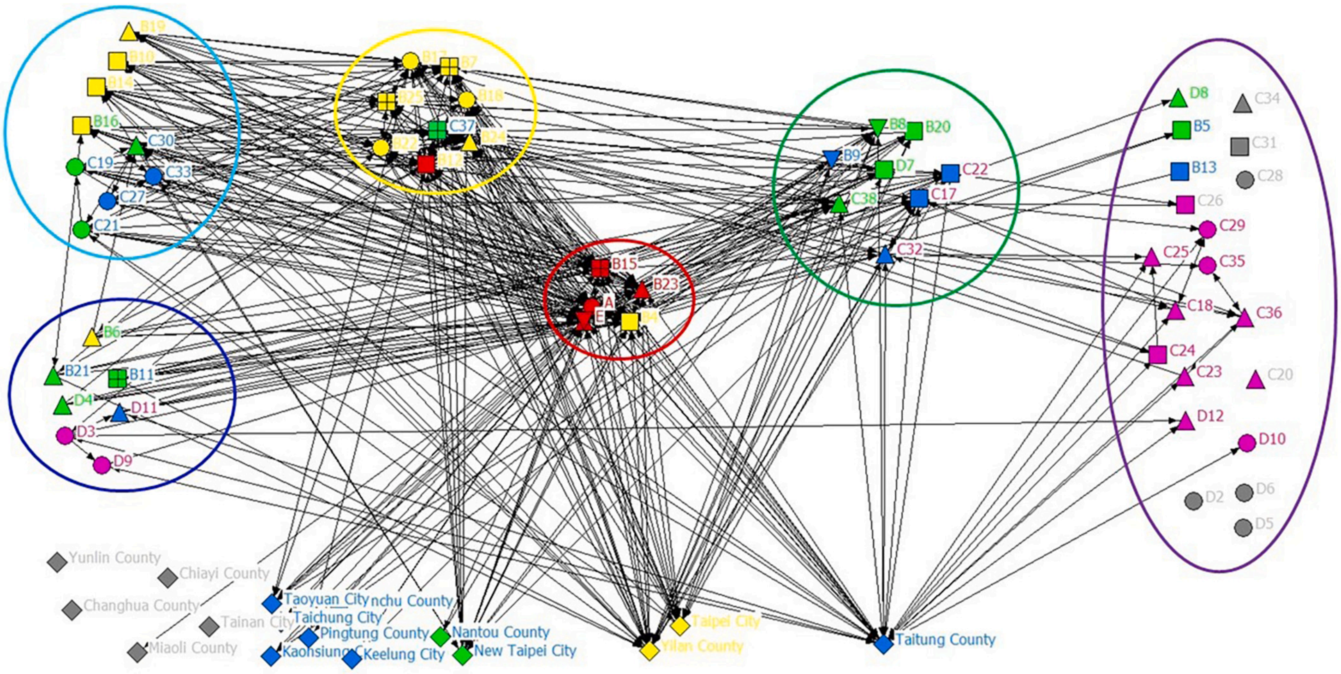


Fig. 14. Network graph of the relationships among six functions of tourism destinations on the weekend.

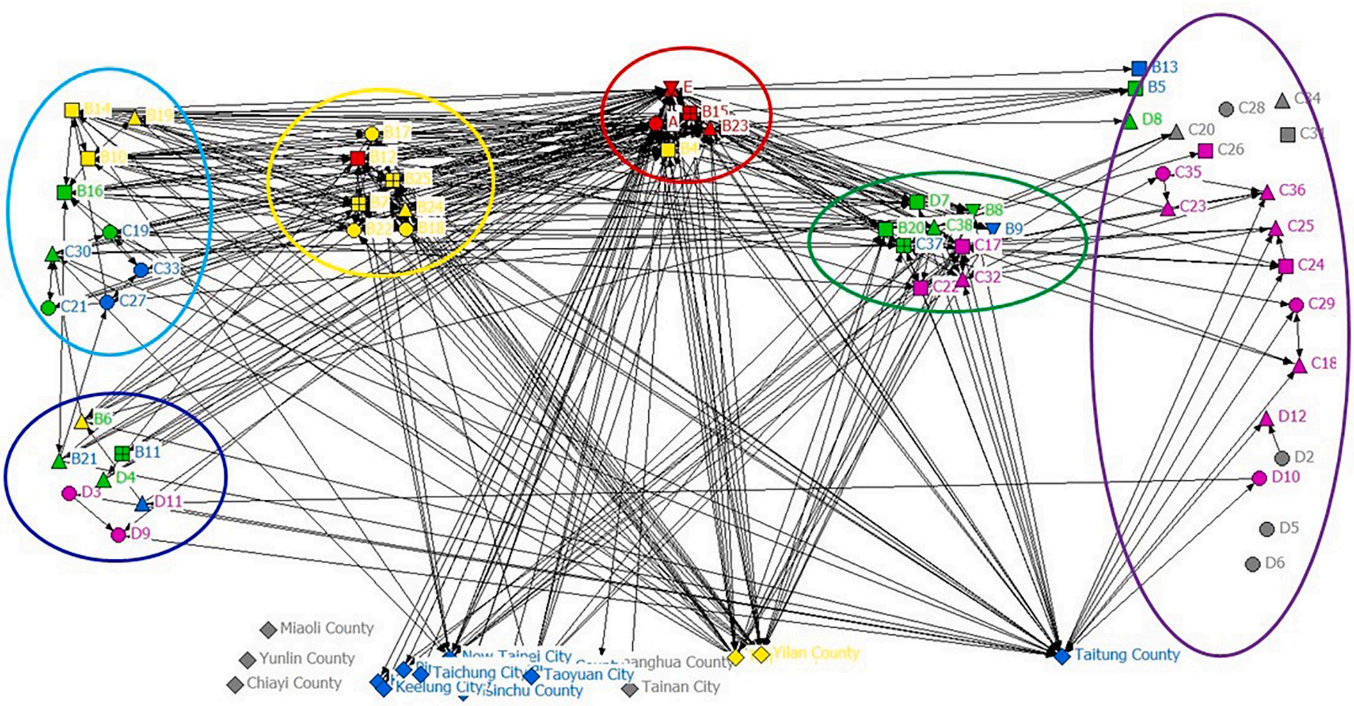


Fig. 15. Network graph of the relationships among six functions of tourism destinations on the weekdays.

groups combine with GIS can convert to the suggest tourist routes in Fig. 17. Social group 1 is the landmark-strategy. Social group 2 and 5 are the route-strategy. Social group 3, 4 and 6 are the orientation-strategy. Social group 7 and 8 are the bypassing trips. We can realize that the landmark-strategy tourists take into account tourist flow distribution centers or hub destinations in advance to choose their trips. Although the route-strategy tourists connect several destinations located along main roads, they do not visit the secondary destinations. The orientation-strategy tourists may visit the region before, so their objective is to specific zones and the secondary destinations.

5. Conclusion

Tourism resource investigation is essential to develop strategic plans for a tourism region. This paper discusses how tourists' trajectories can be analyzed using data collected from high technology. Steps for developing OD matrices for tourism destinations are detailed. The CDR data are a countable resource to automatically data collection for en-

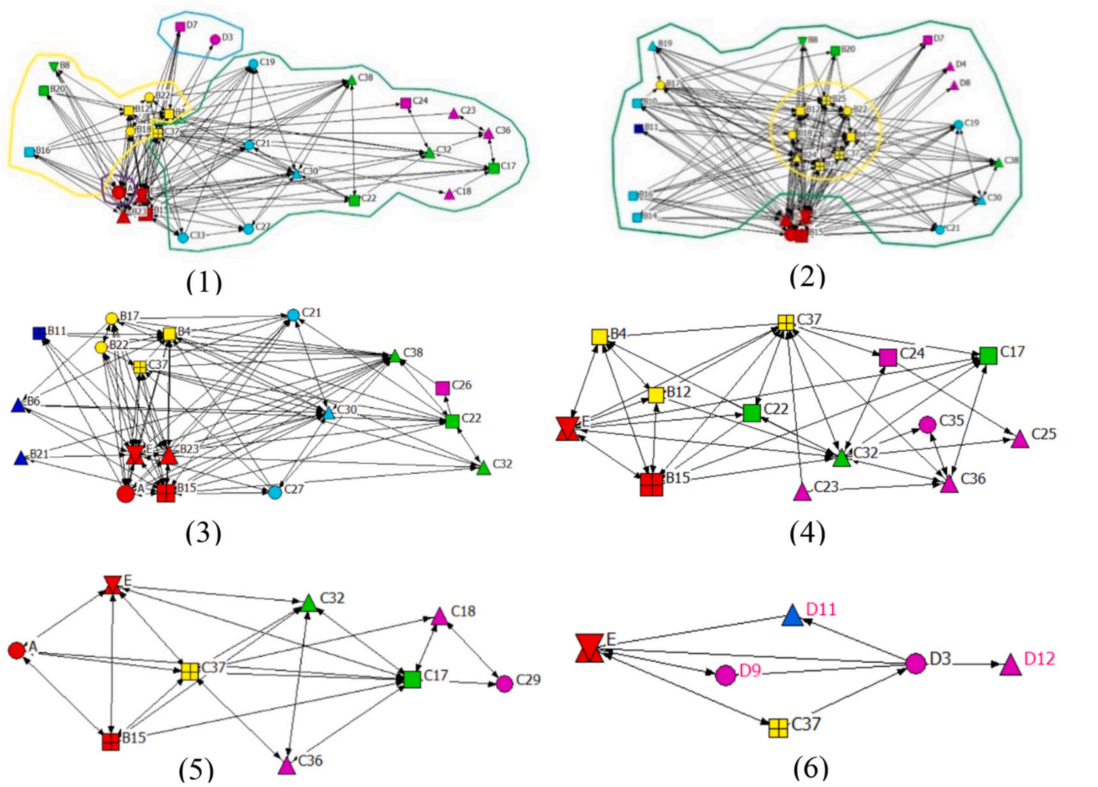


Fig. 16. Network graphs of social groups on the weekend.
Notes: red, distribution center; yellow, hub; green, passageway; blue, common; purple, secondary.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tour trajectory.

Our research method is different from the traditional questionnaire methods and mathematical models, and is able to uncover further information about the roles of tourism destinations and their interrelationships. The locations of tourism destinations are fixed, but the results reveal that their roles and functions are elastic on the weekdays and on the weekend. Uncovering where tourists come from (the generating markets) and which destinations they choose to go to first (gateways) can help tourism planners understand the direction of tourist movements in advance. The key to developing regional tourism management is to understand core destinations.

There are five destinations in Hualien acting as three types of centers at all times in our study. They are the most influential destinations in the spatial network and can control the tourist mobility among all zones. Tourist flow depends on the center destination acting as a core which gathers tourists, from which they spread gradually to other locales. Taroko National Park (A) is famous for its rich nature resources. Accommodation (E) is the beginning and endpoint for each tourism day, so it is necessary for most tourists. Tungtamen Shopping District (B15) and Hualien Railway Cultural Park (B4) are bustling places in Hualien. Local people as well as travelers gather there to buy daily necessities and souvenirs or go for a walk. Although whale watching is a summer activity at Hualien harbor (B23), the harbor still attracts people in other seasons. Tourism destinations with one or two types of centers influence the tourist flow among specific zones or in a zone. Besides popular recreation attractions, the attributes of centers include service components, like commercial districts and accommodation, which supply convenient services.

A tourism destination can act as more than one bridge. As a coordinator or gatekeeper destination, it should provide tourism information and promote the features of attractions in its zone. As a consultant or representative destination, it should focus on promotion of co-operative zones. If it connects with other regions, they can propose cross-regional cooperation together. Finally, liaison destinations, which are most important, can control the tourist mobility between zones. They should provide comprehensive tourism services. If tourists receive the right information in the right places, it can have a significant effect on their choices for their next tourism destinations.

In fact, destinations can be similar, and cannot possibly be absolutely structurally equivalent, as each has its own unique attributes. The role members are greatly influenced by geographic proximity, but do not always abide by this rule. It is remarkable that destinations with a particular role will connect with destinations having the same role. The movements of tourists are based on role relations, and tourist flows spread gradually from destinations with the same role to others with other roles. Thus, the components of a role (i.e. zones sharing the same role) may be located in several zones.

The process of collecting data affects the quality of CDR, especially the definition of capturing raw data. Because each record in the study is based on a person staying put for over 15 min, some data might be lost. Although this is a restriction on the use of existing information, it is also interpretable in behavior. Because generally speaking, it takes more than 15 min to go to a convenience store to buy some snacks or enjoy a cup of coffee for a break.

Based on the conclusions, this study purposes some suggestions as the reference for future researches. First, the future research can join the cross-section data to compare by different time stages. For example, summer vacation is the peak tourist season in Hualien. It might construct the regional network and the functions of destinations that are different from the off season.

Second, the technologies of collecting data are the problems. The positioning accuracy and the grid size are two requirements when the mobile phone positioning data are employed in spatial or tourism planning. Because CDR data are collected by triggering the base stations,

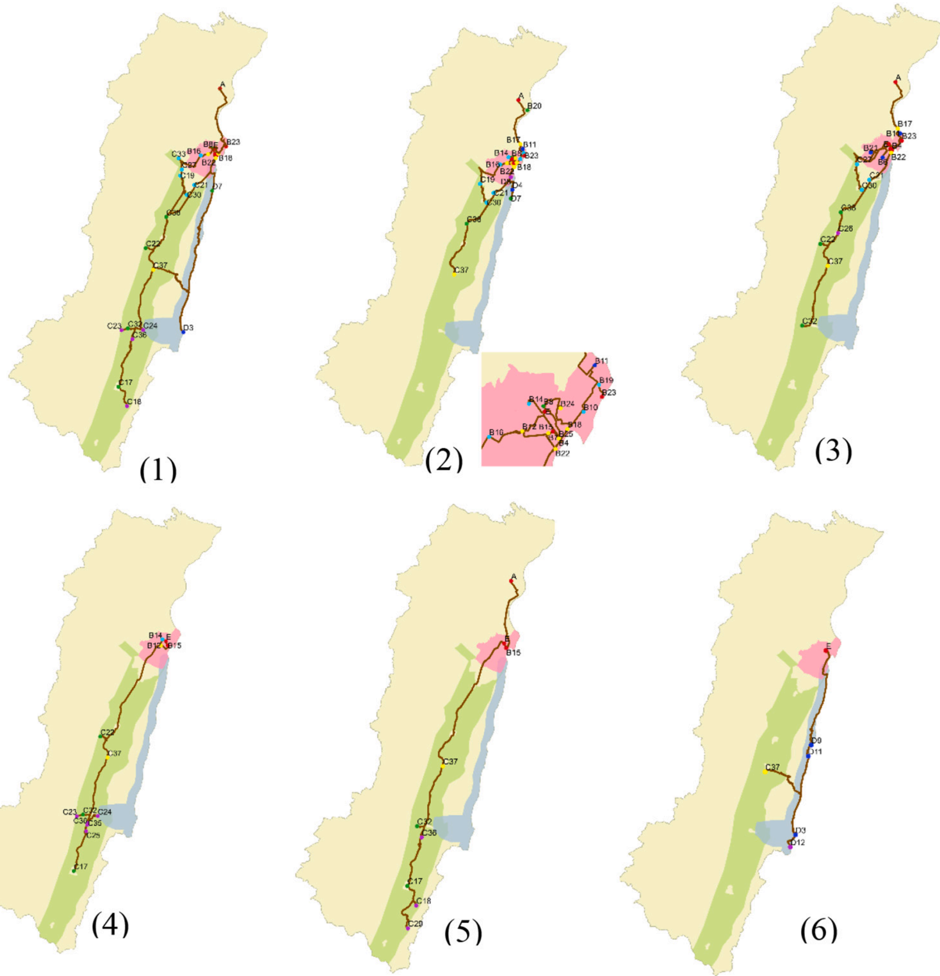


Fig. 17. Tour routes of different social groups on the weekend.

it may cause the loss of data. The network-driven data is better to describe the reality.

Third, there are other expended theories and models in Social network analysis. The future studies can test the adaptability of the indexes for spatial relation.

References

Ahas, R., Aasa, A., Roose, A., Mark, Ü., Silm, S., 2008. Evaluating passive mobile positioning data for tourism surveys: an Estonian case study. Tour. Manag. 29 (3), 469–486.

Ahas, R., Aasa, A., Silm, S., Tiru, M., 2010. Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: case study with mobile positioning data. Transportation Research Part C: Emerging Technologies 18 (1), 45–54.

Barnes, J.A., 1954. Class and committee in a Norwegian Island Parish. Hum Relat 7, 39–58.

Bekkelien, A., Deriaz, M., Marchand-Maillet, S., 2012. Bluetooth Indoor Positioning. Master's thesis. University of Geneva.

Bonnel, P., Fekih, M., Smoreda, Z., 2018. Origin-destination estimation using mobile network probe data. Transportation Research Procedia 32 (2018), 69–81.

Breiger, R.L., Boorman, S.A., Arabie, P., 1975. An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. J. Math. Psychol. 12 (3), 328–383.

Buchanan, T., 1983. Toward an understanding of variability in satisfactions within activities. J. Leis. Res. 15 (1), 39–51.

Burt, R.S., 1992. Structural Holes. Harvard University Press.

Carrasco, J.A., Miller, E.J., 2009. The social dimension in action: A multilevel, personal networks model of social activity frequency between individuals. Transportation Research Part A: Policy and Practice 43 (1), 90–104.

Casanueva, C., Gallego, Á., García-Sánchez, M.R., 2016. Social network analysis in tourism. Curr. Issue Tour. 19 (12), 1190–1209.

Chung, H., Chung, N., Nam, Y., 2017. A social network analysis of tourist movement patterns in blogs: Korean backpackers in Europe. Sustainability 9 (12), 2251.

Çolak, S., Alexander, L.P., Alvim, B.G., Mehndiratta, S.R., González, M.C., 2015. Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. Transportation Research Record: Journal of the Transportation Research Board 2526, 126–135.

Crawford, F., Watling, D.P., Connors, R.D., 2018. Identifying road user classes based on repeated trip behaviour using Bluetooth data. Transp. Res. A Policy Pract. 113, 55–74.

D'Agata, R., Gozzo, S., Tomaselli, V., 2013. Network analysis approach to map tourism mobility. Qual. Quant. 47 (6), 3167–3184.

Dredge, D., 1999. Destination place planning and design. Ann. Tour. Res. 26 (4), 772–791.

Freeman, L.C., 1978. Centrality in social networks conceptual clarification. Soc. Networks 1 (3), 215–239.

Gould, R.V., Fernandez, R.M., 1989. Structures of mediation: a formal approach to brokerage in transaction networks. Sociol. Methodol. 89–126.

Granovetter, M., 1973. The strength of weak ties. Am. J. Sociol. 78, 1360–1380.

Granovetter, M., 1985. On the social embeddedness of economic exchange. Am. J. Sociol. 91, 481–510.

Hualien Tourism Department. Retrieved from. https://td.hl.gov.tw/bin/home.php.

Hanneman, R.A., Riddle, M., Roger S, Chen, 2013. Introduction to social network methods: UCINET applications. Chuliu Publisher, Kaohsiung.

Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C., 2014. Development of origin–destination matrices using mobile phone call data. Transportation Research Part C: Emerging Technologies 40, 63–74.

Jones, P., Lucas, K., 2012. The social consequences of transport decision-making: clarifying concepts, synthesising knowledge and assessing implications. J. Transp. Geogr. 21, 4–16.

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. J. Destin. Mark. Manag. 9, 358–370.

Knoke, D., Yang, S., 2008. Social Network Analysis (No. 154). Sage.

Lew, A., McKercher, B., 2006. Modeling tourist movements: a local destination analysis. Ann. Tour. Res. 33 (2), 403–423.

Lin, Yu-Jun, 2018. Service provider utilized big data made from million users. https://www.ettoday.net/news/20180306/1124651.htm.

Lue, C.C., Crompton, J.L., Fesenmaier, D.R., 1993. Conceptualization of multi-destination pleasure trips. Ann. Tour. Res. 20 (2), 289–301.
Lue, C.C., Crompton, J.L., Stewart, W.P., 1996. Evidence of cumulative attraction in multidestination recreational trip decisions. J. Travel Res. 35 (1), 41–49.
Luo, Jar-der, 2010. Social Network Analysis, 2nd Ed. Social Sciences Academic Press, Beijing.
Mamei, M., Bicocchi, N., Lippi, M., Mariani, S., Zambobelli, F., 2019. Evaluating origin–destination matrices obtained from CDR data. Sensors 19, 4470.
Mayo, G.E., 1933. The Human Problems of an Industrial Society. Macmil-lan, New York.
Milne, D., Watling, D., 2019. Big data and understanding change in the context of planning transport systems. J. Transp. Geogr. 76, 235–244.
Mitchell, J.C., 1969. The concept and use of social networks. In: Mitchell, J.C. (Ed.), Social Networks in Urban Situations. University of Manchester Press, Manchester, England.
Moreno, J.L., 1937. Sociometry in relation to other social sciences. Sociometry 1 (1/2), 206–219.
Pazzaglia, F., De Beni, R., 2001. Strategies of processing spatial information in survey and landmark-centred individuals. Eur. J. Cogn. Psychol. 13 (4), 493–508.
Pike, S., Lubell, M., 2016. Geography and social networks in transportation mode choice. J. Transp. Geogr. 57, 184–193.
Radcliffe-Brown, A.R., 1940. On social structure. J. R. Anthropol. Inst. G. B. Irel. 70 (1), 1–12.
Schiefelbusch, M., Jain, A., Schäfer, T., Müller, D., 2007. Transport and tourism: roadmap to integrated planning developing and assessing integrated travel chains. J. Transp. Geogr. 15 (2), 94–103.
Schiller, J., Voisard, A. (Eds.), 2004. Location-Based Services. Elsevier.
Shih, H.Y., 2006. Network characteristics of drive tourism destinations: an application of network analysis in tourism. Tour. Manag. 27 (5), 1029–1039.
Stewart, S.I., Vogt, C.A., 1997. Multi-destination trip patterns. Ann. Tour. Res. 24 (2), 458–461.
Sugimoto, K., Ota, K., Suzuki, S., 2019. Visitor mobility and spatial structure in a local urban tourism destination: GPS tracking and network analysis. Sustainability 11 (3), 919.
Tettamanti, T., Demeter, H., Varga, I., 2012. Route choice estimation based on cellular signaling data. Acta Polytechnica Hungarica 9 (4), 207–220.
Tourism Bureau, 2019. Republic of China. Survey of Travel by R.O.C. Citizens. Ministry of Transportation and Communications, Republic of China.
Van der Zee, E., Bertocchi, D., 2018. Finding patterns in urban tourist behaviour: a social network analysis approach based on TripAdvisor reviews. Information Technology & Tourism 20 (1–4), 153–180.
Wang, Z., He, S.Y., Leung, Y., 2017. Applying mobile phone data to travel behaviour research: a literature review. Travel Behaviour and Society 11, 141–155.
Warner, W.L., Srole, L., 1945. The Social Systems of American Ethnic Groups.
Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications, Vol. 8. Cambridge University Press.
Wellman, B., 1983. Network analysis: some basic principles. Sociological theory 155–200.
White, H.C., Boorman, S.A., Brieger, R.L., 1976. Social structure from multiple networks. In: Blockmodels of Roles and Positions. American Journal of Sociology, 81, pp. 730–780.
Yue, Y., Lan, T., Yeh, A.G., Li, Q.Q., 2014. Zooming into individuals to understand the collective: a review of trajectory-based travel behaviour studies. Travel Behav. Soc. 1 (2), 69–78.
Zandbergen, P.A., 2009. Accuracy of iPhone locations: a comparison of assisted GPS, WiFi and cellular positioning. Trans. GIS 13 (s1), 5–25.
Zeng, B., 2018. Pattern of Chinese tourist flows in Japan: A social network analysis perspective. Tour. Geogr. 20 (5), 810–832.

Chih-Peng Chu is a professor of Business Administration department and Graduate Institute of logistics Management at National Dong Hwa University and the director of eastern research center for transportation development. He received Ph.D. in Engineering-Economic Systems and Operation Research from Stanford University. His research focuses on business models, network economics, and transport economics. He is currently editor-in-chief of Journal of the Chinese Instituis of Transportation and is as reviewer for multiple journals like Transportation Research Part A, D, and E, European Journal of Operational Research and so on. He has published more than 60 academic journal papers.

Yu-Hsin Chou graduated from Providence University with a bachelor in tourism and obtained a master's degree in business administration at National Dong Hwa University. Her research interests include tourist flow, destination management, network analysis and geographical analysis.