Trip chaining patterns of tourists: A real-world case study

Cong Qi [a], Jonas De Vos [b], Tao Tao [c], Linxuan Shi[d], Xiucheng Guo [a, *]

[a] School of Transportation, Southeast University, No.2 Dongnandaxue Road, Nanjing, 211189, China; 230228859@seu.edu.cn; seuguo@163.com;
[b] Bartlett School of Planning, University College London, 14 Upper Woburn Place, London WC1H 0NN, UK; jonas.devos@ucl.ac.uk
[c] Department of Civil and Environmental Engineering, Carnegie Mellon University, PA, USA, 55454, taot@andrew.cmu.edu
[d] Department of Civil and Environmental Engineering, The George Washington University, Washington DC, United States 20052; shilx@gwu.edu

Abstract: Insights into tourist travel behaviours are crucial for easing traffic congestions and creating a sustainable tourism industry. However, a significant portion of the literature analysed tourist travel behaviour by predefined tourist trip chains which result in the loss of more representative classification. Using tourist travel survey data from Nanjing, China, this paper presents an innovative methodology that combines the tourist trip chain identification and the trip chain discrete choice model to comprehensively analyse the travel behaviour of tourists. The discretized trip chains of tourists are clustered using the Ordering Points to Identify the Clustering Structure (OPTICS) clustering algorithm to identify typical tourist trip chains, which will then be considered as the dependent variable in the Nested Logit model to estimate the significant explanatory variables. The clustering results show that there are two main categories, namely single and multiple attraction trip chain, and seven subcategories, which were named according to the characteristics of trip chains. The clustering result is analysed and three main trip chain patterns are derived. Departure city, travel cost, travel time, and travel mode show significant influence on the choice between single and multiple attraction trip chains. The urban attraction trip chain is more favoured by tourists with children, and the typical trip chain shows stronger dependence on travel intention. The first-time to Lishui only affects the choice of the multiple suburban attraction trip chain. These findings are valuable for optimising tourist public transport infrastructure, promoting travel by public transport and better tourism management.

Keywords: trip chain; travel behaviour; tourism; OPTICS cluster; Nested Logit model;

1. Introduction

Tourism plays a major role for economic growth and development. In China, tourism has grown rapidly before the COVID-19 pandemic, with 6 billion domestic tourists in 2019 compared to 3.6 billion in 2014. After the COVID-19 pandemic, the number of Chinese tourists has bounced back quickly. According to Nanjing Culture and Tourism Bureau, the city attracted 44.37 million tourists in the first quarter of 2023, with an increase of 68.5% compared to the same quarter in 2022. The increasing tourist demand has led to traffic congestion, air pollution, and poor travel experiences(Vu et al., 2015). The analysis of travel behaviour can help identify the temporal and spatial characteristics of the routes and destinations, allowing for preventing capacity overload (Lew and McKercher, 2006), as well as generate a tourist flow corridor for optimising tourist public transport infrastructure and promoting tourists travel by public transport.

Many studies have analysed tourist travel behaviour, where the trip chain has been used as the basic unit of analysis. However, despite a growing interest, most studies tend to predefine the tourist trip chain type like commuting trip chain and then categorize the travel data into the corresponding types (Hermawati et al., 2019; Wu et al., 2012). In contrast to commuter activities that are often centred around work or school, the spatial distribution of stops made by tourists may not be clearly tied to the proximity from a specific location. Due to the difference in urban scale and the spatial distribution of attractions, the tourist trip chain may vary in different cities or districts. Predefining the trip chain may lead to the loss of more representative classification, resulting in poor fit of the parameter regression. In this study, tourist trip chain is defined as a trip involving single or multiple attractions, starting from the same place of accommodation and returning in one day. The tourists refer not only to visitors staying overnight at the destination, but also to local residents travelling to and from the attractions on the day of departure from home.

To fill the research gaps, we apply the Ordering Points to Identify Clustering Structure (OPTICS) clustering method and the Nested Logit(NL) method to the analyse the data collected in Lishui District of Nanjing, China. Specifically, we aim to address two research questions: 1) Compared to the predefined trip chain, whether the other typical type of trip chain can be recognized? 2) What are the significant factors that influence the choice of trip chain and how do they influence?

The present study contributes twofold to the literature. Methodologically, the OPTICS clustering method is adopted to capture the typical tourist trip chains which can efficiently identify and describe the travel patterns of tourists depending on the study area. Empirically, it enriches the understanding of tourists' travel behaviour by estimating the significance of the demographic, trip attribute and travel pattern variables on the typical trip chain choice. By revealing the refined relationships, we offer implications for public transport service providers to provide more flexible circulator and paratransit-type services.

The remainder of this paper is organised as follows. Section 2 reviews the related work and Section 3 introduces the tourism travel survey data. The modelling methodology is conceptualized in Section 4, followed by the discussion of the model estimation results in Section 5. At the end, Section 6 outlines the main conclusions of the study and gives policy implications.

2. Literature review

This section provides an overview of the relevant literature that has taken the trip chain as the basic unit of analysis. The methods for identifying the trip chain and analysing tourist travel behaviour are also reviewed.

To explore travel characteristics from people's daily travel diaries, a multi-activity trip chain was first defined as a trip involving multiple purposes to multiple destinations with some incidental stops (Shiftan, 1998). Mandatory activity, maintenance activity, and optional activity were categorised according to the frequency, duration, and location options of the activities (Krizek, 2003). Hedau and Sanghai, (2014) found that simple activity patterns with one activity purpose, such as home-work-home (HWH), and home-other-home (HOH), make up the majority of weekdays. A person tends to have similar activities every day, so the activity patterns are stable in the trip purpose and destination space. The trip chain was then used as the basic unit of analysis to estimate travel behaviour, focusing on specific scenarios. Trip chains of different occupations and genders, including non-workers, out-of-home travellers, and women were modelled using statistical modelling tools. Scheiner and Holz-Rau (2017) concluded that women have higher levels of entropy, and children have a positive effect on the entropy of their activity patterns. Bautista-Hernández (2022) and Daisy et al. (2018) focused on non-worker trip chaining in Halifax and Mexico City. They all addressed that dense and diverse urban environments were associated with more trip chaining and trip complexity. For out-of-home travellers, driving licenses and car ownership were negatively associated with trip chaining (Daisy et al., 2020). Examining the usage patterns of car-sharing, ride-hailing, and commercial vehicles, Khan and Machemehl (2017), and Tanjeeb Ahmed and Hyland (2022) conduct that ride-hailing trip chains significantly ended in healthcare and social/recreational activities, and shared car usage had peak hours in trip chains (Xiaoyan et al., 2020). The characteristics of multi-activity trip chains (Li et al., 2020) could also be used to quantify the mode choice (Milos Balac et al., 2022; Schneider et al., 2021), location choice (Mariante et al., 2018), holiday travel behaviour (L. Yang et al., 2016), car and bus usage during COVID-19 (S. Kim et al., 2021), and network equilibrium (Gao et al., 2019; Halat et al., 2016). As to the tourist movement patterns, day-to-day travel itineraries were selected as the basic unit of analysis, including all the attractions visited or intended to be visited, and total length of trip (Oppermann, 1995). Five basic spatial patterns, including single destination pattern, en route pattern, base camp pattern, regional tour pattern and trip-chaining pattern were identified by Lue, Crompton, and Fesenmaier (LCF) model for understanding the movement of tourists within a destination (Lue et al., 1993). Lew & McKercher (2006) also identified three types of linear path models, including point-to-point pattern, circular pattern and complex combinations of the point-to-point and circular patterns.

Although previous studies shed light on the recognition and identification of the typical trip chain, there is limited discussion on the tourist trip chain. Different dimensions of travel patterns including the number of trips, the choice of trip purpose, type of travel mode, next destination location, trip start time and stop duration were modelled using clustering algorithms, including K-means (Ma et al., 2013), DBSCAN (Le Minh Kieu et al., 2015), fuzzy K-means (Chen et al., 2019), and PAM clustering (Ma et al., 2016). The main data sources used for this model were smart card data, GPS data, travel survey data, and cellular signal data. Four clusters of freight trip chains were determined according to truck GPS data sets from several trucking companies travelling in Washington State (Ma et al., 2016), while the same trip chains were also classified using data from the Austin Commercial Vehicle Survey (Khan & Machemehl, 2017). Using smart card data in Beijing, (Ma et al., 2013) recognised five clusters of transit rider regularity, including very high (VH), high (H), medium(M), low (L), and very low (VL). (Le Minh Kieu et al., 2015) adopted the DBSCAN algorithm and priori market segmentation approach and segmented transit riders into four identifiable types, such as transit commuters, habitual time riders, regular OD riders, and irregular riders. (Duan et al., 2017) developed an entropy method to predict the daily activity points and trip-chain pattern stability based on anonymous mobile phone data from 5 to 25 September 2011.

Research on tourist travel behaviour has been discussed for a long time. Several studies had adopted the structural model and hierarchical linear model to estimate tourist intention. Travel motivation and its effect on tourist travel participation and behaviours were significant with economic conditions (Wong et al., 2018), destination image (Afshardoost and Eshaghi, 2020), social media influencers (Pop et al., 2021; Xiang et al., 2015). While the intention to travel abroad was directly influenced by the number of family members, the level of the residence, and macroeconomic factors (H.-R. Kim et al., 2019; Wong et al., 2016). Apart from tourist intention, the choice behaviour of tourist travel mode (Hermawati et al., 2019; Qi et al., 2020) and destination (Karl, 2018; Vu et al., 2015) was further modified using tourist travel survey data. Classical discrete choice modelling frameworks (i.e., MNL model) were selected to estimate the destination and mode choices. The results showed that destination choices were associated with tourism motivations (Wu et al., 2012), demographic characteristics (Tang et al., 2020), and spatial configuration of destinations (Y. Yang et al., 2013). García et al. (2015) investigated that the importance of previous visits to Majorca increased the probability of revisiting. People preferred to use the same type of travel modes when they travel (Wu et al., 2012). On the other hand, the frequency of using public transport decreased when the destination was located outside of the main tourist area (Tang et al., 2020). Few researchers had combined the tourists' travel behaviour with the trip chain, where the travel intention, travel mode, and destination choice could be considered together with logistic interactions. Hermawati et al. (2019) argued that tourists who performed trip chains 2,3 and 4 mostly chose car renting, while the highest probability of renting a motorcycle was the tourists who performed trip chains 5 or more in Bali. Pop et al. (2021) concluded that the choice of transport mode also differed between tour and non-tour activities. When non-local tourists decided to engage in tour activities during their holidays, public transport was more favourable than private cars. Due to the devastating effects of the COVID-19 pandemic on the tourism sector from 2020 onwards, recent studies have asked whether the COVID-19 risk perception would influence tourists' intention to behave responsibly (Chen et al., 2021) and have examined the travel intentions during and after the pandemic (Talwar et al., 2021). Proximal post-COVID travel behaviours were categorised into five types, including mortality salience during the pandemic, fright-and-flight travel abstinence behaviour, invincible me disruptive travel behaviour, corona light rational travel behaviour, and compensatory binge travel behaviour, while the distal post-COVID travel behaviour consisted of travel restriction salience, de-globalisation and bounded tourism behaviour, posttraumatic growth and travel as a search for meaning (Miao et al., 2021).

In summary, there are two main research gaps in the literature. On the one hand, the tourist trip chains analysed in the existing research are all predefined. However, the spatial distribution of stops made by tourist travellers may not be clearly tied to the proximity from a specific location. Due to the difference in urban scale and the spatial distribution of attractions, the tourist trip chain may vary in different cities or districts. On the other hand, most studies analysed the travel behaviour by travel mode or travel intention that used disaggregated data used trip-based models. In other words, they only considered each trip separately and ignored the interdependence between trips. The tourist trip chain could combine each trip into a complete chain and, thus show the whole trip information in explaining travel patterns.

3. Data

3.1 Data source

The primary data set used in this study was derived from the 2021 Tourism Travel Survey carried out in the Lishui District of Nanjing, China. The study area comprises 5 sub-district totalling 1067 square kilometre (Figure 1). The survey was designed to collect detailed information for analysing tourist travel behaviour and estimating the tourist flow. Demographic information is important for this study and only survey data provide them. Smart card data, GPS data, and cellular signal data have the large sample size but lacks demographic information. In addition, these types of data are not available for this study. Therefore, we chose to use survey data. The stay-at-home order was relaxed during the survey period, and the statistics for tourist arrivals and tourism revenue in 2019 and 2021 are similar. In 2019, there were more than 10 million tourist arrivals and RMB 13.7 billion in tourism revenue for Lishui District. In 2020, there was a significant decrease in both figures. In 2021, tourist arrivals reached 13.6 million and tourism revenue exceeded RMB 15.7 billion. Therefore, this study assumes that the impact of the pandemic on tourist travel behaviour and patterns is trivial.

The 2021 Tourism Travel Survey was conducted in April and May of 2021. The survey period covered both weekdays and weekends, as well as the Qingming Festival holiday, which is a three-day national holiday when people can travel to distant places. Twenty-three attractions in Lishui were selected to distribute the questionnaires (Figure 1). The respondents were mainly selected based on the distribution of age and gender. The form of the interview is face-to-face, and the respondents were asked to finish two questionnaires: demographic form and travel diary. The demographic form collected information, including gender, education, driving license and private vehicle ownership, and the trip-related attributes, including the departure city, trip cost, trip duration, travel intention, number of accompanies, and children. The travel diary collected all trips made by the participants on the day the survey was carried out, including the origin and destination (OD) of the trips, travel time, travel mode, and activity time. A total of 417 tourists completed the survey. They made 1087 trips during their travels.

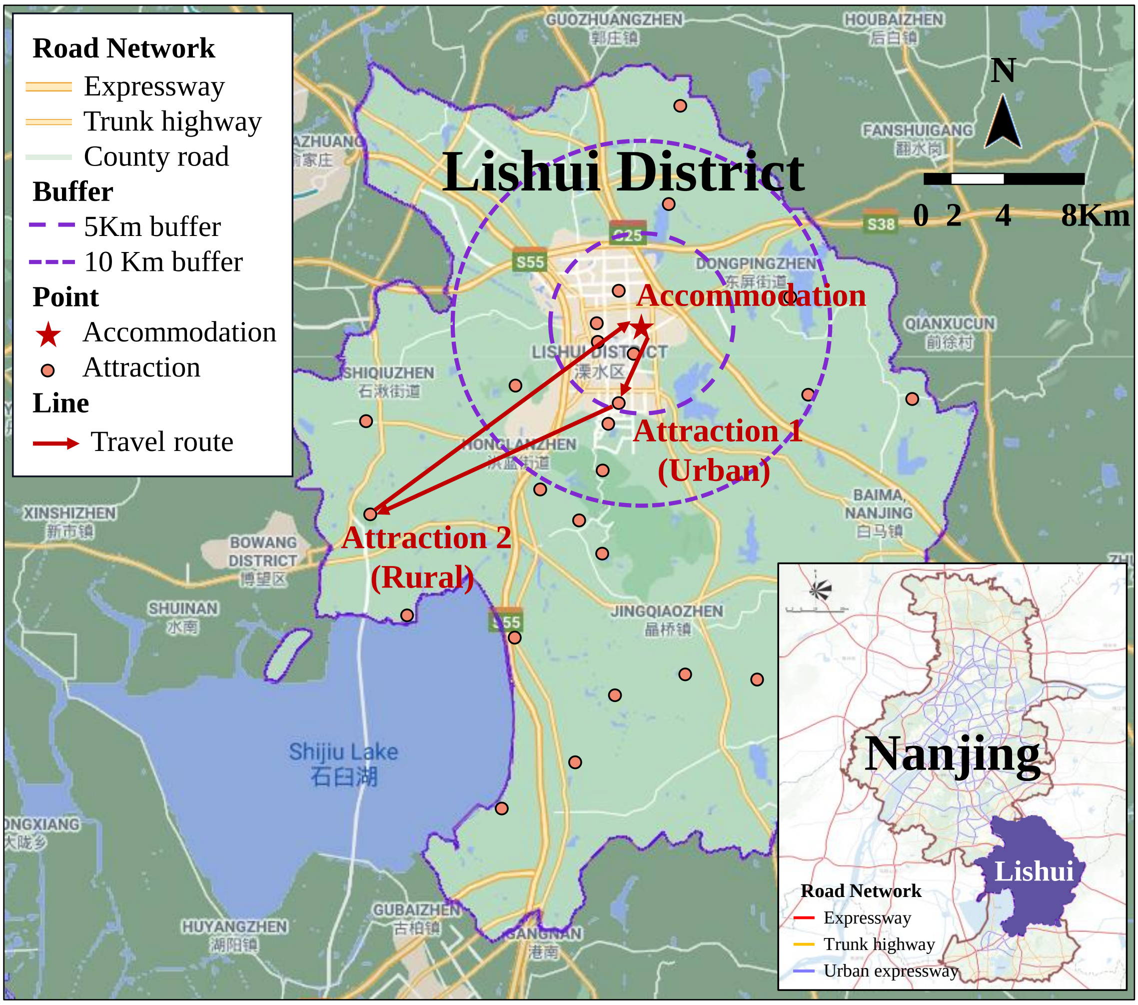


Figure 1. Studied Area

3.2 Trip chain pattern

In a tourist trip chain, a traveller makes one or more stops at attractions other than the accommodation, where a travel can be defined as the movement between origin and destination (Vu et al., 2015) and an activity can be defined as a visit to an attraction. In this study, we applied travel distance discretization to construct the basic trip chain. The discretized distance is categorized conceptually into five categories according to the state of travel and distance from the attraction to the starting point. The detailed definition of discretised travel distance of tourist trip chain has been described in Table 1. By discretizing the spatial distance between the tourist's location and accommodation at each time point, the trip chain can be described by a set of point sequences as shown in Equation (1):

$$Ch_Q = \{(k_{Q1}, t_{Q1}), (k_{Q2}, t_{Q2}), \ldots, (k_{QN}, t_{QN})\} \quad (1)$$

where Ch_Q is the point sequence set of the trip chain Q, and N is the number of sequence points in that trip chain Q. For any point $n \in \{1,2,\ldots,N\}$, k_{Qn} and t_{Qn} represent the discretised distance of the travel/activity following the sequence point n and time stamp of the sequence point n in the trip chain Q, respectively.

Table 1. Definition of Discretised Travel Distance of tourist trip chain

Name	Discretised distance	Explanation
Accommodation	0	Tourists are at home/hotel, and no trip took place.
Travel	1	Tourists are on their way to the next destination.
Urban attraction	2	If the distance between the attractions and the starting points is smaller than 5 km, then the attraction is regarded as urban attractions.
Suburban attraction	3	If the distance between the attractions and the starting points ranges from 5 to 10 km, then the attraction is regarded as suburban attractions.
Rural attraction	4	If the distance between the attractions and the starting points is longer than 10 km, then the attraction is regarded as rural attractions.

For example, the trip chain exhibited in Figure 2 shows that the tourists first set off to the first attraction from accommodation. Since the travel distance is less than 5 kilometres, the discrete distance of the first attraction is 2, and attraction 1 is defined as an urban attraction. After visiting the first attraction, the tourists travel to the attraction 2, which is defined as a rural attraction with its discrete distance value as 4. Finally, the tourists returned accommodation after visiting two attractions. The tourist trip chain can be visualized as in Figure 2 and discretized as follows:

$$Ch_Q = \{(1, t_{Q1}), (2, t_{Q2}), (1, t_{Q3}), (4, t_{Q4})(1, t_{Q5})(0, t_{Q6})\}$$

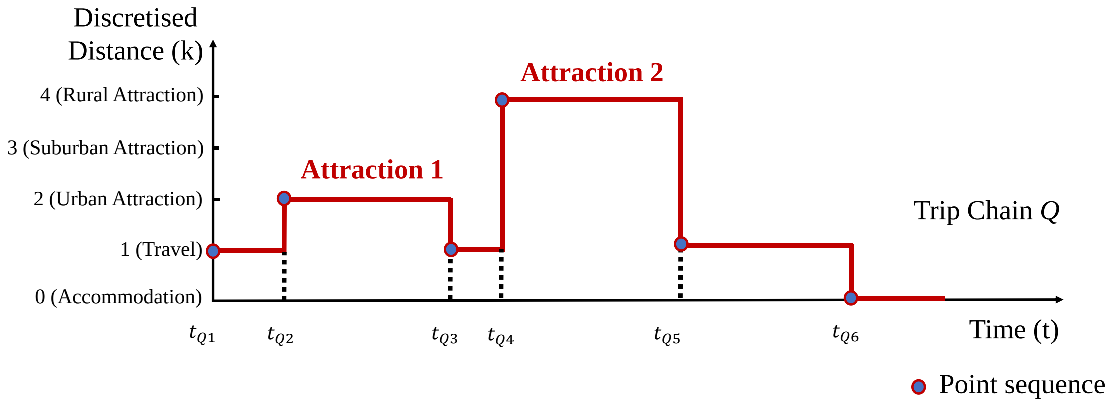


Figure 2. The discretized tourist trip chain

Table 2 presents descriptive statistics by trip chain pattern. The total travel time refers to the overall duration of all travels in a trip chain which includes the transit time and waiting time, while the total activity time refers to the time spent at each attraction during the trip chain. The terms average travel time and average activity time represent values obtained by averaging corresponding total value with respect to the number of travel and attractions visited. The ratio column of the table indicates the percentage of each trip chain. More than 80% of the tourists have an average travel time less than 1 hour, while the total travel time mostly ranges from <1h to 2-3h. 44.25% of the tourists stop at each attraction for 2-3h on average, while the total activity time of 2-4h and 4-6h accounted for the most with 34.77% and 37.18% respectively. Out of these 417 trip chains, 52.28% contain 2 trips, 36.45% contain 3 trips, 9.59% contain 4 trips and 1.68% contain 5 trips. Among all travel modes, the private vehicle is the most preferred by tourists, with a usage rate of 69.78%, suggesting a high baseline utility preference for a private vehicle. Public transport is the second most popular travel mode, followed by active travel, which includes walking, biking and other non-motorised travel, and tour coach.

Table 2. Descriptive statistics of trip chain pattern

Variable	Description	Ratio
Average Travel Time	<0.5h	55.01%
	0.5-1h	29.07%
	1-1.5h	6.44%
	1.5-2h	4.14%
	>2h	5.34%
Total Travel Time	<1h	27.58%
	1-2h	29.98%
	2-3h	21.82%
	3-4h	11.03%
	>4h	9.59%
Average Activity Time	<1h	3.28%
	1-2h	23.62%
	2-3h	44.25%
	3-4h	18.09%
	>4h	10.76%
Total Activity Time	<2h	6.95%
	2-4h	34.77%
	4-6h	37.18%
	6-8h	13.67%
	>8h	7.43%
Length of Trip Chain	2	52.28%
	3	36.45%
	4	9.59%
	5	1.68%
Travel Mode	Active Travel	7.42%
	Public Transport	16.21%
	Private Vehicle	69.78%
	Tour Coach	6.59%

3.3 Explanatory variables

In this study, we included three types of explanatory variables: demographic variables, trip attributes, and travel pattern variables. According to the literature, the choice of trip chain is primarily influenced by gender, education, car ownership, and driving licence ownership (Gross and Grimm, 2018). In addition, more travel companions as well as longer duration of the trip could lead tourists to choose a simple trip chain (Gutiérrez and Miravet, 2016). Table 3 presents the descriptive statistics of the demographic and trip attributes, while the Table 2 has exhibited the travel pattern variables in Section 3.2.

Table 3. Descriptive statistics of demographic and trip attribute variables

	Variable	Description	Percentage
Demographic	Gender	Male	47.36%
		Female	52.64%
	Education	Junior High	10.24%
		Senior High	24.70%
		Undergraduate	54.82%
		Graduate	10.24%
	License Ownership	Yes	82.42%
		No	17.58%
	Car Ownership	Yes	75.76%
		No	24.24%
	Departure City	Lishui District	33.09%
		Other District in Nanjing	45.32%
		Other City	21.58%
	First-time to Lishui	Yes	50.52%
		No	49.48%
	Children	Yes	29.5%
		No	70.5%
Trip Attribute	Accompany Number	0	14.39%
		1	30.94%
		2	22.06%
		>3	32.61%
	Trip Cost	<100 RMB	27.34%
		100-500 RMB	32.85%
		500-1000 RMB	11.99%
		1000-3000 RMB	16.55%
		>3000 RMB	11.27%
	Trip Duration	1 day	69.34%
		2 day	10.53%
		>2 days	20.14%
	Travel Motivation	Scenery Experience	45.51%
		Cultural	24.86%
		Relaxation	29.64%

4. Methods

The method consists of four steps (Figure 3). (1) Compute the distance between discretized tourist trip chain. (2) Use the distance as an indicator for the similarity measurement of the

OPTICS clustering algorithm. (3) Build the discrete choice model and specify the model structure based on the clustering results. (4) Select the explanatory variables and include in the NL model by examining the covariance and significance, then the coefficient can be estimated.

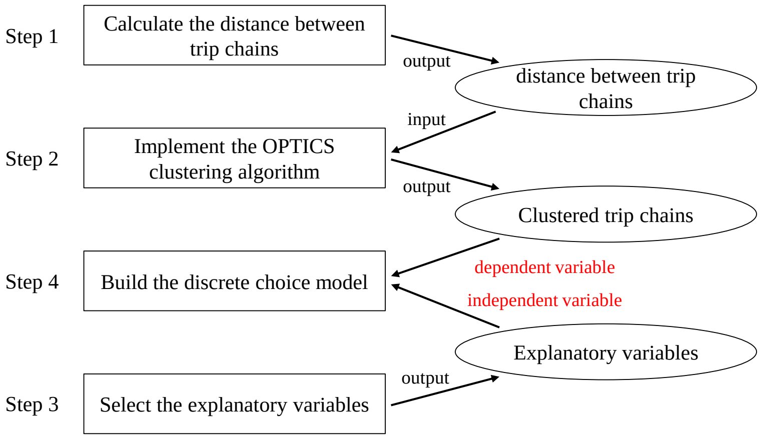


Figure 3. Workflow diagram

4.1 Computation of distance between trip chains

In order to cluster various tourist trip chains into distinct and representative categories using a density-based algorithm, we computed the distance between any two discretised trip chains as the similarity between them. We illustrated the computation of distance between two trip chains, P and Q (Figure 4).

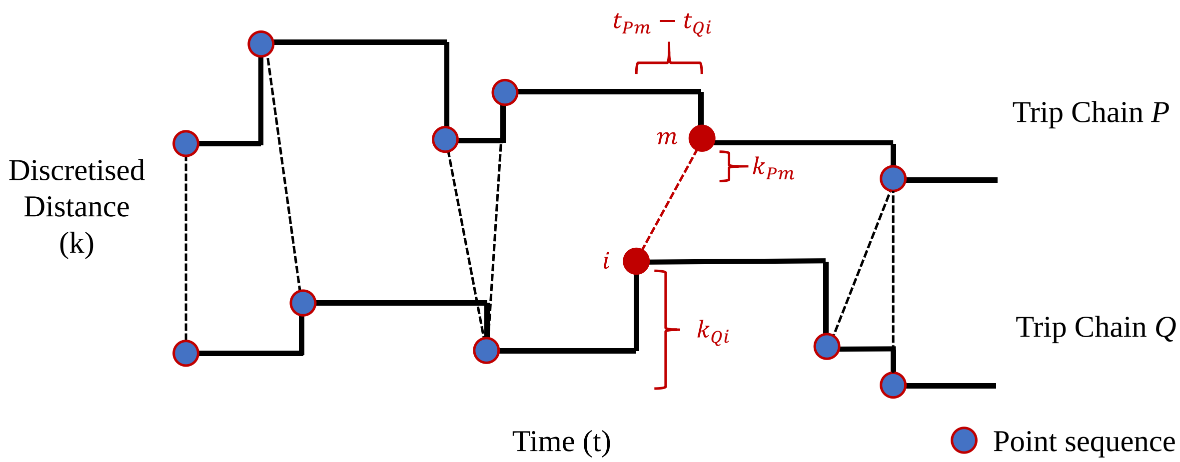


Figure 4. Distance between two trip chains

First, we assumed that there are M sequence points in trip chain P and N sequence points in chain Q. Each specific sequence point can be denoted as $(k_{Pm}, t_{Pm}), m \in \{1,2,\ldots,M\}$ and $(k_{Qn}, t_{Qn}), n \in \{1,2,\ldots,N\}$, respectively. The first step is to loop through all points m in trip chain P. For each point m in trip chain P, its closest point i in chain Q is defined as (k_{Qi}, t_{Qi}), which means $i = argmin_{n'} (t_{Pm} - t_{Qn'}), \forall n' \in N$. The discretised distance between these two points is then defined as Equation (2),

$$d(m,i) = \begin{cases} k_{Pm} - k_{Qi} & if\ t_{Pm} - t_{Qi} < T \\ +\infty & otherwise \end{cases} \quad (2)$$

If the time difference $t_{Pm} - t_{Qi}$ between the sequence point m and its closest corresponding point i is greater than T, which is selected as one hour in this study, the distance between these two points is considered infinite.

Similarly, for each point n in chain Q, its closest point j in chain P can be found as $\left(k_{Pj}, t_{Pj}\right)$, where $j = argmin_{m'} \left(t_{Qn} - t_{Pm'}\right), \forall\, m' \in M$. Their discretised distance is also defined as Equation (3),

$$d(n, j) = \begin{cases} k_{Qn} - k_{Pj} & if\ t_{Qn} - t_{Pj} < T \\ +\infty & otherwise \end{cases} \tag{3}$$

Finally, the distance $D(P, Q)$ between the two discretised trip chains P and Q can be calculated through Equation (4),

$$D(P, Q) = \frac{\sum_{m=1}^{M} d(\mathrm{m}, \mathrm{i}) + \sum_{n=1}^{N} d(\mathrm{n}, \mathrm{j})}{M + N}. \tag{4}$$

4.2 OPTICS clustering algorithm

After the computation of distance between trip chains, we used the Ordering Points to Identify the Clustering Structure (OPTICS) algorithm to cluster tourist trip chains into categories. OPTICS borrows the concept of core density-reachable from the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ankerst et al., 1999). Compared to the DBSCAN clustering algorithm, OPTICS is an extended ordering algorithm from which either flat or hierarchical clustering results can be derived (Hahsler et al., 2019). Nanni & Pedreschi (2006) applied the OPTICS clustering algorithm to the clustering of spatio-temporal trajectories and obtained good clustering results. Since the focus of this study is tourist trip chain, which has a similar data structure to that of spatio-temporal trajectories, we choose the OPTICS clustering algorithm in our study. Clustering starts with a dataset D containing a set of points $p \in D$. The ε and $minPts$ are the basic required parameter for the clustering algorithm, where the ε represents the radius of neighborhood of a point p and the $minPts$ is the minimum number of objects in the ε-neighborhood. OPTICS generates a reachability plot by incorporating two additional concepts, the core-distance and the reachability-distance, which are defined below,

The core-distance of a point p is the smallest value of radius such that the ε-neighbourhood of p has at least $minPts$ objects. If the number of objects is less than the $minPts$, the point will not be recognised as a core point with an undefined core-distance. Given $minPts$ and ε, the core-distance is defined with Equation (5),

$$\text{core-dist}(p;\varepsilon,minPts)=\begin{cases} UNDEFINED & if\ |N_\varepsilon(p)|<minPts \\ MinPts\text{-}dist(p) & otherwise \end{cases} \quad (5)$$

where $N_\varepsilon(p)$ is the number of objects within its ε- neighbourhood and $minPts\text{-}dist(p)$ is the distance from p to its $minPts$-th smallest distance in $N_\varepsilon(p)$.

The reachability-distance from a point $p \in D$ to a point $q \in D$ parameterised is defined with Question (6),

$$\text{reachability-dist}(p,q;\varepsilon,minPts) = \begin{cases} UNDEFINED & if\ |N_\varepsilon(p)|<minPts \\ \max(\text{core-dist(p)}, d(p,q)) & otherwise \end{cases} \quad (6)$$

The reachability-distance of a core point p concerning an object q is the smallest neighbourhood radius that p would be directly density-reachable from q. In this research, ε is set to a large value to limit the number of points considered in the neighbourhood search. Therefore, OPTICS will consider more nearest neighbours in the core-distance calculation. $minPts$ affects the smoothness of the reachability distribution, with larger values leading to a smoother reachability distribution.

4.3 Nested Logit model

In this study, the Nested Logit (NL) model was chosen to explore travel behaviour, where dependent variable was the typical trip chain identified by the OPTICS clustering algorithm (Section 5.1). The hierarchical structure of the NL model differs from the traditional logit models in terms of that NL allows the choice probabilities of any two alternatives in different nests depend on the attributes of each other. In this way, the independence of irrelevant alternatives (IIA) property does not generally hold for alternatives in different nests, and the NL model can determine the decision order more intuitively. The nesting structure was set up to explore the choice behaviour of trip chains, as shown in Figure 5.

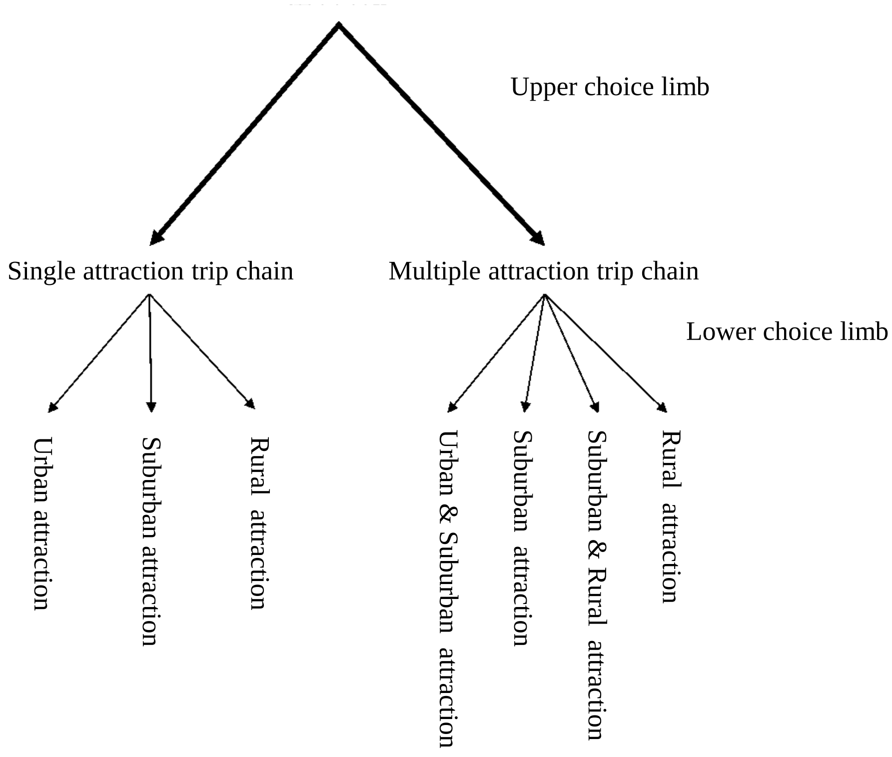


Figure 5. NL model structure

In the NL model, decision maker n, faced with J alternatives, chooses the alternative with the greatest utility among the choice set. The probability of choosing alternative i is $P_{ni} = Prob(U_{ni} - U_{nj} > 0, \forall j \neq i)$. The component of utility U_{ni} can be decomposed into three parts as shown in Equation (7) below: (1) a constant part labelled as W_{nk}, which remains the same for all alternatives within a nest, (2) a part labelled Y that varies depending on the specific alternative within a nest, and (3) a random component μ_{ni}.

$$U_{ni} = W_{nk} + Y_{ni} + \mu_{ni} \tag{7}$$

It is often reasonable to express the observed part of the utility in terms of linear parameters as Equation (8,9):

$$W_{nk} = \sum_{l=1}^{L} \alpha_l x_{nkl} \tag{8}$$

$$Y_{ni} = \sum_{m=1}^{M} \beta_m z_{nim} \tag{9}$$

where x_{nkl} is a vector of features l relating to alternatives for nest k; z_{nim} is a vector of features m relating to alternative i; α_l and β_m are the coefficients of the variables, which are the parameters to be estimated. Once an alternative is selected within the nest B_k, the probability of selecting alternative i can be expressed as Equation (10):

$$P_{ni} = P_{ni|B_k} P_{nB_k} \tag{10}$$

The marginal and conditional probabilities can be expressed as Equation (11-13), where I_{nk} is the inclusive value linking the upper and lower levels of the nested structure.

$$P_{nB_k} = \frac{e^{W_{nk}+\lambda_k I_{nk}}}{\sum_{l=1}^{K} e^{W_{nl}+\lambda_l I_{nl}}} \tag{11}$$

$$P_{ni|B_k} = \frac{e^{Y_{ni}/\lambda_k}}{\sum_{j\in B_k} e^{Y_{nj}/\lambda_k}} \tag{12}$$

$$I_{nk} = ln \sum_{j\in B_k} e^{Y_{nj}/\lambda_k} \tag{13}$$

The coefficient λ_k reflects the degree of independence between the unobserved fractions of utility for alternatives in nest B_k, with lower values indicating less independence. The parameter λ_k may differ across nests, reflecting different correlations between unobserved factors within each nest. The likelihood ratio index is a statistical measure used in the context of a nested logit model to compare the goodness of fit between two models. In a nested logit model, the ρ^2 value is computed as the log-likelihood of a model in which all the parameters are set to zero, divided by the log-likelihood of the fully specified model. This statistic ranges from 0 to 1, with higher values indicating a better fit of the model to the data.

4.4 Explanatory Variable selection

Based on the cross-correlation table, the $Pearson\ \chi^2$ is used to analyse the correlation of the independent variables with the upper choice limb and the lower choice limb separately. If the significance value of the $Pearson\ \chi^2$ is less than 0.05, the variable is considered to have an effect on the choice of tourist trip chain and is included in the model. The variance inflation factor (VIF) is then used to check the autocorrelation of the independent variables. If the VIF is greater than 5 or the tolerance is less than 0.2, there is covariation and the covariates are removed. All the explanatory variables listed in Section 3.2 and 3.3 are examined. They will be estimated in the NL model if the corresponding significance level is less than 0.005. Finally, departure city, accompany number, trip cost, travel time, and mode are used in the upper choice model, where the single attraction trip chain and the multiple attraction trip chain are selected first. Other district in Nanjing and other city in the departure city variable are merged into other district, and active travel, private vehicle and tour coach in the travel mode variable are merged into non-public transport mode. The independent variables used in the lower choice model to predict the choice of typical trip chains are children, travel intention, and the first-time to Lishui. The other variables are not estimated in the NL model.

5. Results

We applied the OPTICS clustering algorithm on the collected tourism travel survey data to first categorize the typical trip chains, and then used the NL model to explore the significant explanatory variables that influence the choice of trip chains. We discussed the related results in this section.

5.1 Typical tourist trip chain identification

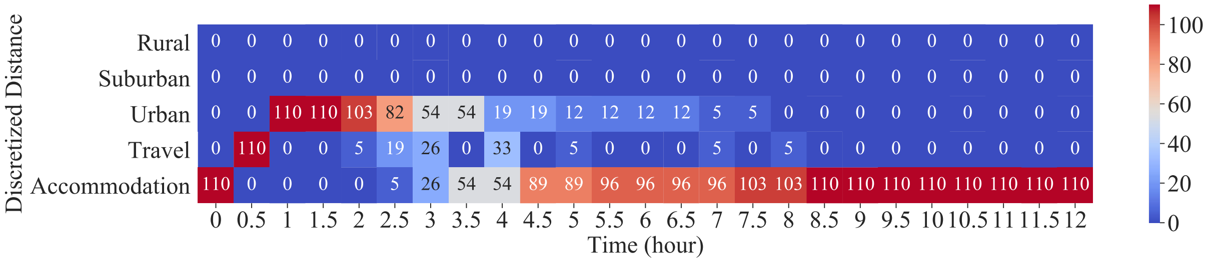
5.1.1 Division and visualization of trip chain

Two main categories with 7 typical tourist trip chains are labelled by OPTICS clustering algorithm. The grid matrix of trip chains in the same category is aggregated to reveal the typical trip chain patterns (Figure 6 and Figure 7). The horizontal coordinates represent the total time of the trip chain, while vertical coordinates represent the discretised distance of the activity chain at each time point. For example, in a tourist trip chain where a tourist is visiting an urban attraction at Hour 1.5 after the departure from accommodation, the cell with the horizontal coordinate of 1.5 and the vertical coordinate of the Urban is noted as 1. The colour of each cell indicates the number of trip chains recorded at that point, with red showing a larger number and blue showing a smaller number. We then analysed the travel time, activity time, the number of trip chains, and activity points based these visualizations. The two categories include single attraction trip chains and multiple attraction trip chains. Single attraction trip chains involve only one visited location, while the multiple attraction trip chains include at least two activity points, indicating that tourists would have more than three trips in one day.

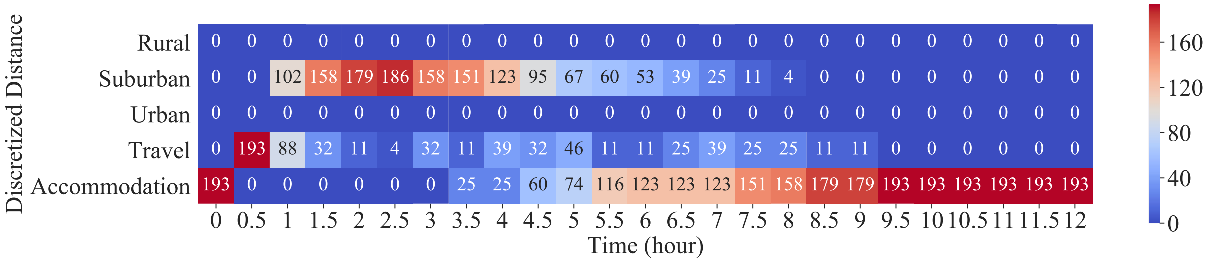
5.1.2 Single attraction trip chain

There are three typical trip chains in the single attraction trip chain, where the urban attraction trip chain, suburban attraction trip chain, and rural attraction trip chain account for 17.98%, 31.46%, and 50.56% respectively. As shown in Figure 6, the travel time decreases from 1.5h to 0.5h, where the travel time of the rural attraction trip chain cost the most, followed by the suburban attraction trip chain and urban attraction trip chain. This is consistent with our expectation that longer travel time is correlated with greater distance between the attraction and accommodation. The activity time of urban attraction trip chain ranges from 1h to 2.5h, while the tourist activity time of suburban and rural attractions fluctuates around 4h. This pattern implies that the outlying attractions may cover a larger area and take tourists a longer time to visit. As the time cost of travel increases, tourists may opt to spend more time in attractions for relaxation and entertainment. The most favourable attractions are Chenghuang

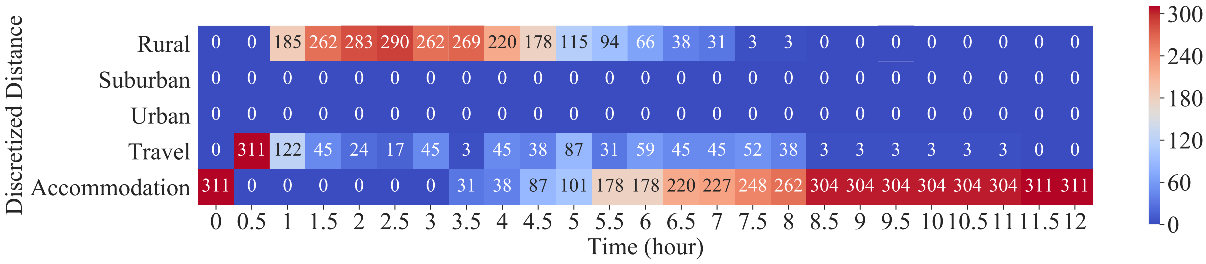
Temple Cultural District, Tian Sheng Qiao Scenic Spot, and Shishu Lake for urban attraction trip chain, suburban attraction trip chain, and rural attraction trip chain separately.



(a) urban attraction trip chain



(b) suburban attraction trip chain

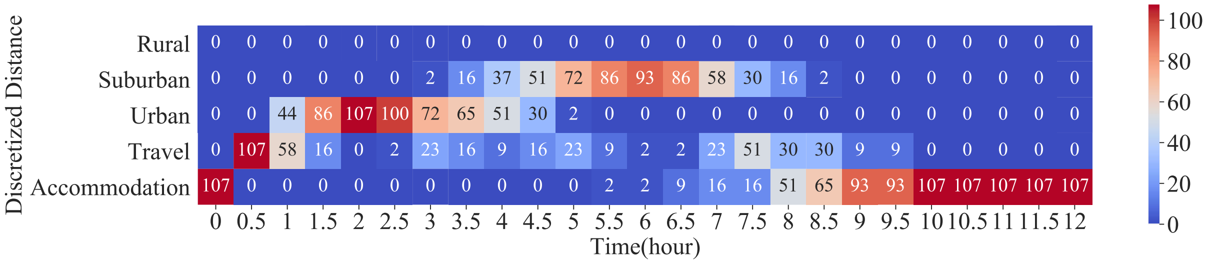


(c) rural attraction trip chain

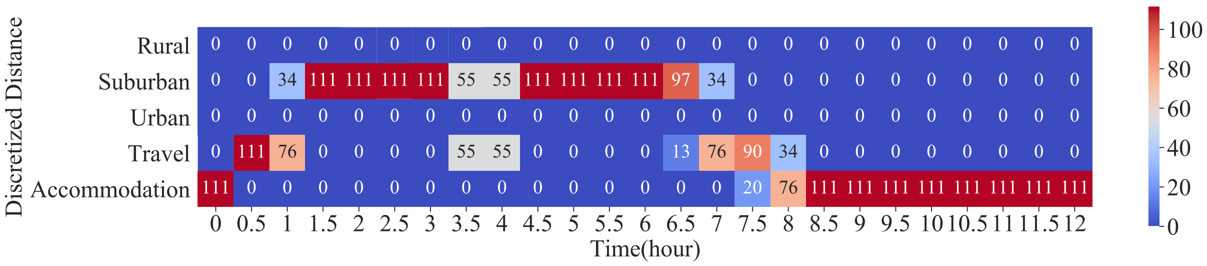
Figure 6. Grid matrix of single attraction trip chains

5.1.3 Multiple attraction trip chain

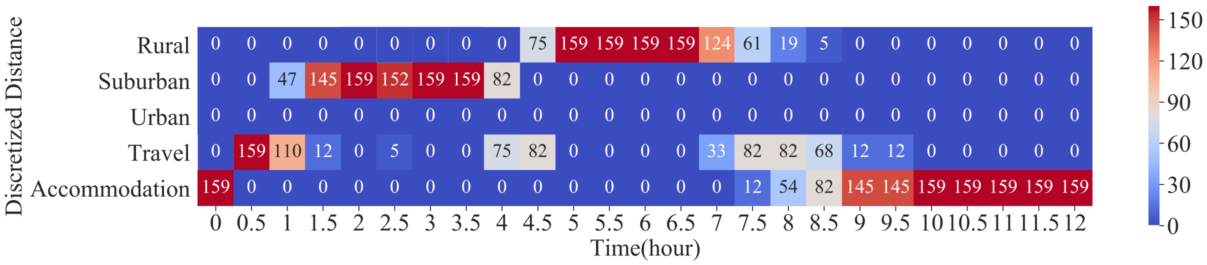
There are four typical trip chains in the multiple attraction trip chain, where the urban & suburban attraction trip chain, suburban attraction trip chain, suburban & rural attraction trip chain, and rural attraction trip chain account for 23.53%, 23.53%, 33.82%, 19.12% respectively. Tourists would not arrange two distant attractions on the same day based on the fact that the trip chain containing both urban and rural attractions is not identified, indicating that tourists prefer destinations that are closer in proximity to minimize the total time cost of the trips. In addition, none of the multiple attraction trip chains connect the two urban attractions, which can be explained by the fact that visiting one urban attraction is sufficient to satisfy the tourist's need for relaxation and there is no need to visit two of the same type of urban attractions. Overall, the total trip time, including travel time and activity time, fluctuates around 9 hours. The average travel time is mostly less than 1 hour, no matter in single or multiple attraction trip chains. One exception is multiple rural attraction trip chains. This is because the formation of two attractions is planned according to the distance, so the final return travel time is almost the sum of the two previous travel times. The total activity time fluctuate between 4h to 6h. Tourists spend more time in the suburban attraction than in the urban attraction in the urban & suburban attraction trip chain. In the suburban, and suburban & rural attraction trip chain, the visiting time in each attraction is roughly equal to 2.5h. Tourists spend more time in the suburban attraction than in the urban attraction in the multiple attraction trip chain, which is consistent with the conclusion in 5.1.2. Furthermore, the activity time in the first rural attraction is twice that of the second one. This could be explained by the fact that the tourists have to complete the trip taking into account the closing time of the attraction or the time needed to return. Tian Sheng Qiao Scenic Spot have the maximum probability of being selected in trip chain including suburban attractions. The most favourable attractions for trip chain including suburban attractions are Shishu Lake and Wuxiang Mountain.



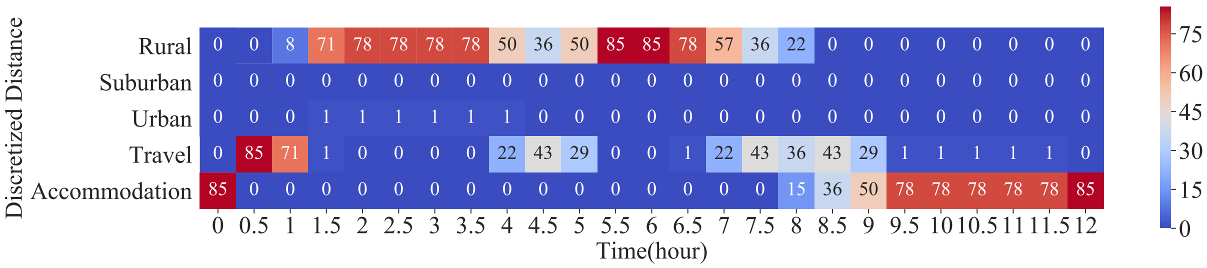
(a) urban & suburban attraction trip chain



(b) suburban attraction trip chain



(c) suburban & rural attraction trip chain



(d) rural attraction trip chain

Figure 7. Grid matrix of multiple attraction trip chains

5.2 Tourist trip chain choice behaviour

Table 4 shows the estimation results obtained with the NL models. The ρ^2 value of the model is 0.4897, indicating a good fitness to the data. Meanwhile, the estimated parameters of the λ_k and most independent variables are statistically significant at 95%, indicating that the model is consistent with utility maximisation for all possible values of the explanatory variables.

5.2.1 Choice between single and multiple attraction trip chain nest

In order to analyse the choice behaviour of trip chains, we have to choose a base alternative. In this study, the multiple attraction trip chain was chosen as the base alternative in the top model. Except for accompanying number, all explanatory variables exhibit statistical significance. The estimates of average travel time and trip cost are both negative. These negative values suggest the fact that longer travel distance and travel time may encourage respondents to include more activities in a one-day trip and lead them to prefer multiple attraction trip chains over the other option. This implies that tourists may be willing to trade off longer travel times and higher trip costs for a more diverse and fulfilling travel experience. On the other hand, the positive values of the parameter for departure city and travel mode indicate that the native tourists intend to choose the single attraction trip chain, and the use of public transport is positively correlated with the single attraction trip chain. This can be explained by the fact that the travel time of public transport is longer than that of private cars, partly because there is no direct public transport between some attractions. As a result, tourists may be constrained by time and opt to visit a single attraction in one day rather than multiple destinations.

5.2.2 Single attraction trip chain choice

In order to estimate the parameter values of the explanatory variables in single attraction trip chains, the urban attraction trip chain is selected as the base alternative. The children factor is significant with negative coefficients of -6.02 and -5.54 for both suburban and rural attractions, respectively. When families travel with children, the urban attraction trip chain is selected more often, including amusement parks, museums and other attractions, which provide entertainment for children. The suburban and rural attraction may lead to physical exhaustion or decline of interest of children's due to the longer time in travelling and visiting. The coefficients of scenery experience motivation for visiting rural attraction and suburban attraction is 5.44 and 4.19 separately, indicating that the tourists with travel purpose for nature sightseeing are more likely to choose the attraction far from the city centre. Tourists with cultural and religious purposes tend to choose the suburban attraction trip chain (6.87), as most of the attractions with cultural value and national 3A+ level tourist attraction are distributed in suburban areas. On the other hand, for the relaxation travel intention, urban attraction trip chain is the most popular one with negative coefficient of -3.26 and -2.17.

5.2.3 Multiple attraction trip chain choice

To estimate the parameter values of the explanatory variables in multiple attraction trip chains, the urban & suburban attraction trip chain is selected as the base alternative. The results show that the first time visit to Lishui positively influences the choice of the suburban attraction trip chain with coefficient of 2.18. This is intuitive, because the national 3A+ level tourist attraction located in suburban areas is the first choice for non-local tourists and they will choose to visit more than one attraction in one day due to the overall travel time and cost budget. The parameters estimated for children (-2.70, -2.22, and -2.57) indicated that the multiple attraction trip chain including urban attractions is more favoured by tourists with children, which is consistent with the result in Section 5.1.3. The motivation of scenery experience is significantly related to the trip chain containing the rural attractions (2.00, 2.00), while the cultural motivation showed significance at 1% with trip chain containing the suburban attractions (3.74, 3.58). The results indicate that tourists with a travel intention focused on natural scenery tend to choose trip chains that include a rural attraction, whereas those with a travel intention for cultural purposes are more likely to choose a trip chain that includes a suburban attraction. Conversely, tourists with a travel intention for entertainment are less likely to choose a trip chain that includes either a rural or suburban attraction. These findings are also consistent with the results presented in Section 5.1.2.

Table 4. Estimation results of tourist trip chain choice

	Single	Multiple	$Sing_{Sub}$	$Sing_{Rur}$	$Sing_{Urb}$	$Multi_{Sub}$	$Multi_{Rur}$	$Multi_{Sub\&Rur}$	$Multi_{Urb\&Sub}$
Constant Term									
X	-	-	-	-	-2.48		-	-	0.65
Y	-	-5.71*	-	-	-		-	-	-
Explanatory Variables for Upper Choice Limb									
Accompany Number	0.30	-	-	-	-	-	-	-	-
Departure City	4.35***	-	-	-	-	-	-	-	-
Trip Cost	-2.01***	-	-	-	-	-	-	-	-
Average Travel Time	-4.19*	-	-	-	-	-	-	-	-
Travel Mode	3.15***	-	-	-	-	-	-	-	-
Explanatory Variables for Lower Choice Limb									
First-time to Lishui	-	-	10.13	9.36	-	2.18*	1.08	0.58	-
Children	-	-	-6.02*	-5.54*	-	-2.70*	-2.22*	-2.57*	-
Scenery experience	-	-	4.19*	5.44*	-	1.01	2.00*	2.00***	-
Cultural	-	-	6.87*	6.16	-	3.74***	0.95	3.58***	-
Relaxation	-	-	-3.26*	-2.17*	-	-0.90	-0.75	-0.20	-
Inclusive Value Parameters									
I_{nk}	0.46***	0.66***	-	-	-	-	-	-	-
λ_k	0.70***	0.78***	-	-	-	-	-	-	-

Note: $N = 1087$, $LL(\beta) = -175.0743$, $LL(\beta_H) = -343.1079$, $\rho^2 =$ 0.4897; ***, * = significance at 1%, 5%.

6. Implication

The results of multiple attraction trip chains showed that travellers prefer to arrange two nearby attractions in one-day trip. This founding suggests that strategies aimed at reducing the inconvenience of travelling may improve the tourist flow and their travel experience. For example, in the suburban & rural attraction trip chain, tourists would most likely to visit Wuxiang Mountain and Tian Sheng Qiao Scenic Spot. For better tourism management, these two attractions could promote a combination ticket, the price of which is much lower than buying two tickets separately. This discounted combination ticket can strengthen the mutual attraction of tourists between the two scenic spots and attract visitors to each other's scenic spots. In terms of transportation, the shuttle bus or customised bus can be provided to connect the attractions, and the cost of transport could be included in the combination tickets. By reducing the time and cost of travelling between scenic spots, tourists will be encouraged to visit more attractions by transit. The use of private cars will be reduced, as well as the traffic congestion and carbon emissions. At the same time, visitors will spend less time travelling and more time enjoying the attractions, thus boosting the economy of the local tourism industry.

In addition, as this study shows, the choice of the trip chain including the urban attractions is sensitive to the number of children along with the visitors. Therefore, tourism and traffic management strategies could be specifically designed for improving children’s experience. Taking the Chenghuang Temple Cultural District, the most favourable attraction in the urban attractions, as an example. Toys, snacks and other tourist products, such as ice cream, puzzles and toy bricks with Chenghuang Temple logo or shape, can be promoted to attract children to buy. Transport facilities should be children friendly. For instance, the phase of traffic signals needs to take into account the walking speed of children. As there is a risk of children suddenly running onto the road, the speed limit of vehicles passing the roads around the scenic area should also be emphasised to avoid traffic accidents.

7. Conclusion

This study applied an OPTICS clustering algorithm to identify the typical tourist trip chain based on the distance between discretized trip chains. Two main categories with 7 typical tourist trip chains were identified, including urban attraction trip chain, suburban attraction trip chain, rural attraction trip chain, urban & suburban attraction trip chain, suburban attraction trip chain, suburban & rural attraction trip chain, and rural attraction trip chain. In addition, we found that travellers prefer to arrange two nearby attractions in multiple attraction trip chains. Then, we used the NL model to explore the factors that influence the tourist trip chain choice. The results showed that tourists from Lishui District and public transport mode are more related to the choice of the single attraction trip chains, and the demand for multiple trip chains increases with higher travel cost and time. Moreover, the estimation of the lower lamb in the NL model revealed that the tourists with children have an obvious preference for the tourist trip chain including urban attractions. The travel motivations are all significant with the tourist trip chain including relevant attractions. The analysis can provide valuable insights into tourist decision-making and behaviour.

The empirical results of this study have important implications for future research on tourist flow corridors and the planning of tourist public transport routes and schedules to provide more convenient and efficient services to tourists. It will also help in analysing road safety and congestion problems around tourist attractions and making appropriate improvements.

There are two limitations of the present study. First, the primary data set used in this study is survey data, which is a relatively single type of data. If multiple types of data, such as smart card data or cellular signal data, can be collected for research, the accuracy of the results can be further improved and more interesting results can be obtained. The second limitation concerns the fact that this study did not consider the built environment of the attraction due to the related data was not available. Future studies could improve the model estimation when the built environment datasets are available, and build new models to analyse the tourist trip chain patterns and tourist travel behaviour based on the big data.

Author contributions

CQ: Conceptualization, Investigation, Formal analysis, Methodology, Writing - original draft, Resources, Software. **JDV**: Project administration, Supervision, Writing - review & editing. **TT**: Validation, Writing - review & editing. **XG**: Formal analysis, Funding acquisition. **LS**: Data curation, Visualization.

Declarations

Conflict of interest on behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Afshardoost, M., & Eshaghi, M. S.: Destination image and tourist behavioural intentions: a meta-analysis. Tourism Management. **81**, 104154 (2020). https://doi.org/10.1016/j.tourman.

Ankerst, M., Breunig, M. M., Kriegel, H.-P., et al.: OPTICS: ordering points to identify the clustering structure. ACM SIGMOD Conference. **28**(2), 49–60 (1999) https://doi.org/10.1145/304182.304187

Balac, M., Horl, S., & Schmid, B.: Discrete choice modeling with anonymized data. Transportation. (2022) https://doi.org/10.1007/s11116-022-10337-1

Bautista-Hernández, D. A.: Individual, household, and urban form determinants of trip chaining of non-work travel in México City. Journal of Transport Geography. **98**, 103227 (2022). https://doi.org/10.1016/j.jtrangeo.2021.103227

Chen, X., Wan, X., Li, Q., et al.: Trip chain based travel mode shares driven framework using cellular signaling data and web based mapping service data. Transportation Research Record. **2673**(3), 51–64 (2019). https://doi.org/10.1177/0361198119834006

Chen, Y., Dai, Y., Anyu Liu, et al.: Can the COVID-19 risk perception affect tourists' responsible behavior intention: an application of the structural equation model. Journal of Sustainable

Tourism. 1–20 (2021). https://doi.org/10.1080/09669582.2021.1977938

Daisy, N. S., Liu, L., & Millward, H.: Trip chaining propensity and tour mode choice of out-of-home workers: Evidence from a mid-sized Canadian city. Transportation. **47**(2), 763–792 (2020). https://doi.org/10.1007/s11116-018-9915-2

Daisy, N. S., Millward, H., & Liu, L.: Trip chaining and tour mode choice of non-workers grouped by daily activity patterns. Journal of Transport Geography. **69**, 150–162 (2018). https://doi.org/10.1016/j.jtrangeo.2018.04.016

Duan, Z., Wang, C., Zhang, H. M., et al.: Using longitudinal mobile phone data to understand the stability of individual travel patterns. Transportation Research Record. **2643**(1), 166–177 (2017). https://doi.org/10.3141/2643-18

Feng, X., Sun, H., Wu, J., et al.: Trip chain based usage patterns analysis of the round trip carsharing system: A case study in Beijing. Transportation Research Part A-Policy and Practice, **140**, 190–203 (2020). https://doi.org/10.1016/j.tra.2020.08.017

Gao, G., Sun, H., & Wu, J.: Activity based trip chaining behavior analysis in the network under the parking fee scheme. Transportation. **46**(3), 647–669 (2019). https://doi.org/10.1007/s11116-017-9809-8

García, J., Juaneda, C., Raya, J. M., et al.: A study of traveller decision making determinants: prioritizing destination or travel mode? Tourism Economics. **21**(6), 1149–1167 (2015). https://doi.org/10.5367/te.2015.0517

Gross, S., & Grimm, B.: Sustainable mode of transport choices at the destination public transport at German destinations. Tourism Review. **73**(3), 401–420 (2018). https://doi.org/10.1108/tr-11-2017-0177

Gutiérrez, A., & Miravet, D.: The determinants of tourist use of public transport at the destination. Sustainability. **8**(9), 908 (2016). https://doi.org/10.3390/su8090908

Hahsler, M., Piekenbrock, M., & Doran, D.: Dbscan: fast density based clustering with R. Journal of Statistical Software. **91**(1), 1–30 (2019). https://doi.org/10.18637/jss.v091.i01

Halat, H., Zockaie, A., Mahmassani, H. S., et al.: Dynamic network equilibrium for daily activity-trip chains of heterogeneous travelers: Application to large-scale networks. Transportation. **43**(6), 1041–1059 (2016). https://doi.org/10.1007/s11116-016-9724-4

Hedau, A. L., & Sanghai, S. S.: Development of trip generation model using activity based approach. International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development. **4**(3), 61-77 (2014).

Hermawati, P., Adisasmita, S. A., Ramli, M.I., et al.: Choices models of trip chain and transportation mode for international tourists in tourism destination island. International Journal of Geomate. **16**(55) (2019). https://doi.org/10.21660/2019.55.88460

Karl, M.: Risk and uncertainty in travel decision making: tourist and destination perspective. Journal of Travel Research. **57**(1), 129–146 (2018). https://doi.org/10.1177/0047287516678337

Khan, M., & Machemehl, R. B.: Analyzing tour chaining patterns of urban commercial vehicles. Transportation Research Part A-Policy and Practice. **102**, 84–97 (2017). https://doi.org/10.1016/j.tra.2016.08.014

Kieu, L. M., Bhaskar, A., & Chung, E.: Passenger segmentation using smart card data. IEEE Transactions on Intelligent Transportation Systems. **16**(3), 1537–1548 (2015). https://doi.org/10.1109/tits.2014.2368998

Kim, H.-R., Yi, C., & Jang, Y.: Relationships among overseas travel, domestic travel, and day trips for latent tourists using longitudinal data. Tourism Management. **72**, 159–169 (2019). https://doi.org/10.1016/j.tourman.2018.11.018

Kim, S., Lee, S., Ko, E., et al.: Changes in car and bus usage amid the COVID-19 pandemic: Relationship with land use and land price. Journal of Transport Geography. **96** (2021). https://doi.org/10.1016/j.jtrangeo.2021.103168

Krizek, K. J.: Neighborhood services, trip purpose, and tour based travel. Transportation. **30**(4), 387–410 (2003). https://doi.org/10.1023/a:1024768007730

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

Li, L., Cao, M., Yin, J., et al.: Observing the characteristics of multi activity trip chain and its influencing mechanism. Ksce Journal of Civil Engineering. **24**(11), 3447–3460 (2020). https://doi.org/10.1007/s12205-020-1927-8

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

7383(93)90056-9

Ma, X., Wang, Y., McCormack, E., et al.: Understanding freight trip chaining behavior using a spatial data mining approach with GPS data. Transportation Research Record. **2596**, 44–54 (2016). https://doi.org/10.3141/2596-06

Ma, X., Wu, Y. J., Wang, Y., et al.: Mining smart card data for transit riders' travel patterns. Transportation Research Part C-Emerging Technologies. **36**, 1–12 (2013). https://doi.org/10.1016/j.trc.2013.07.010

Mariante, G. L., Ma, T.-Y., & Van Acker, V.: Modeling discretionary activity location choice using detour factors and sampling of alternatives for mixed logit models. Journal of Transport Geography. **72**, 151–165 (2018). https://doi.org/10.1016/j.jtrangeo.2018.09.003

Miao, L., Im, J., Fu, X., et al.: Proximal and distal post-COVID travel behavior. Annals of Tourism Research. **88**, 103159 (2021). https://doi.org/10.1016/j.annals.2021.103159

Nanni, M., & Pedreschi, D.: Time-focused clustering of trajectories of moving objects. Journal of Intelligence and Information Systems, **27**(3), 267–289 (2006). https://doi.org/10.1007/s10844-006-9953-7

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

Pop, R.-A., Saplacan, Z., Dabija, D.-C., et al.: The impact of social media influencers on travel decisions: The role of trust in consumer decision journey. Current Issues in Tourism. 1–21 (2021). https://doi.org/10.1080/13683500.2021.1895729

Qi, C., Zhu, Z., Guo, X., et al.: Examining interrelationships between tourist travel mode and trip chain choices using the nested logit model. Sustainability. **12**(18), 7535 (2020). https://doi.org/10.3390/su12187535

Scheiner, J., & Holz-Rau, C.: Women's complex daily lives: a gendered look at trip chaining and activity pattern entropy in Germany. Transportation. **44**(1), 117–138 (2017). https://doi.org/10.1007/s11116-015-9627-9

Schneider, F., Ton, D., Zomer, L.-B., et al.: Trip chain complexity: a comparison among latent classes of daily mobility patterns. Transportation. **48**(2), 953–975 (2021). https://doi.org/10.1007/s11116-020-10084-1

Shiftan, Y.: Practical approach to model trip chaining. Transportation Research Record. **1645**, 17–23 (1998). https://doi.org/10.3141/1645-03

Talwar, S., Srivastava, S., Sakashita, M., et al.: Personality and travel intentions during and after the COVID-19 pandemic: an artificial neural network (ANN) approach. Journal of Business Research (2021). https://doi.org/10.1016/j.jbusres.2021.12.002

Tang, X., Wang, D., Sun, Y., et al.: Choice behavior of tourism destination and travel mode: a case study of local residents in Hangzhou, China. Journal of Transport Geography. **89**, 102895 (2020). https://doi.org/10.1016/j.jtrangeo.2020.102895

Tanjeeb Ahmed, & Hyland, M.: Exploring the role of ride hailing in trip chains. Transportation. **50**, 959–1002 (2022). https://doi.org/10.1007/s11116-022-10269-w

Vu, H. Q., Li, G., Law, R., & Ye, B. H.: Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. Tourism Management. **46**, 222–232 (2015). https://doi.org/10.1016/j.tourman.2014.07.003

Wong, I. A., Fong, L. H. N., & Law, R.: A longitudinal multilevel model of tourist outbound travel behavior and the dual cycle model. Journal of Travel Research. **55**(7), 957–970 (2016). https://doi.org/10.1177/0047287515601239

Wong, I. A., Law, R., & Zhao, X.: Time variant pleasure travel motivations and behaviors. Journal of Travel Research. **57**(4), 437–452 (2018). https://doi.org/10.1177/0047287517705226

Wu, L., Zhang, J., & Fujiwara, A.: Dynamic analysis of Japanese tourists' three stage choices: tourism participation, destination choice, and travel mode choice. Transportation Research Record. **2322**, 91–101 (2012). https://doi.org/10.3141/2322-10

Xiang, Z., Magnini, V. P., & Fesenmaier, D. R.: Information technology and consumer behavior in travel and tourism: insights from travel planning using the internet. Journal of Retailing and Consumer Services. **22**(1), 244–249 (2015). https://doi.org/10.1016/j.jretconser.2014.08.005

Yang, L., Shen, Q., & Li, Z.: Comparing travel mode and trip chain choices between holidays and weekdays. Transportation Research Part A-Policy and Practice. **91**, 273–285 (2016). https://doi.org/10.1016/j.tra.2016.07.001

Yang, Y., Fik, T. J., & Zhang, J.: Modeling sequential tourist flows: where is the next destination? Annals of Tourism Research. **43**, 297–320 (2013). https://doi.org/10.1016/j.annals.2013.07.005

Yoo, C. K., Yoon, D. H., & Park, E.: Tourist motivation: An integral approach to destination choices. Tourism Review, **73**(2), 169–185 (2018). https://doi.org/10.1108/tr-04-2017-0085