Tourism Management 31 (2010) 464–469

Contents lists available at ScienceDirect

Tourism Management

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

Market segments based on the dominant movement patterns of tourists

Jianhong (Cecilia) Xia [a,*], Fiona H. Evans [b], Katrina Spilsbury [c], Vic Ciesielski [d], Colin Arrowsmith [e], Graeme Wright [f]

[a] Department of Spatial Sciences, Curtin University of Technology, GPO Box U1987, Perth 6845, Australia
[b] Centre for Environment and Life Sciences, Department of Agriculture and Food, Private Bag 5, PO Wembley WA 6913, Australia
[c] Centre for Population Health Research, School of Public Health, Faculty of Health Sciences, Curtin University of Technology, GPO Box U1987, Perth WA 6845, Australia
[d] School of Computer Science and Information Technology, Australia
[e] School of Mathematical and Geospatial Science, RMIT University, GPO Box 2476V, Melbourne, Victoria 3001, Australia
[f] Office of Research and Development, Curtin University of Technology, GPO Box U1987, Perth 6845, Australia

ARTICLE INFO

Article history:
Received 29 February 2008
Accepted 30 April 2009

Keyword:
Tourist
Market segmentation
Expectation–Maximisation algorithm
Log-linear models
Movement pattern

ABSTRACT

This paper presents an innovative method for tourist market segmentation-based on dominant movement patterns of tourists; that is, the travel sequences or patterns used by tourists most frequently. There were three steps to achieve this goal. In the first step, general log-linear models were adopted to identify the dominant movement patterns, while the second step was to discover the characteristics of the groups of tourists who travelled with these patterns. The Expectation–Maximisation algorithm was then used to partition tourist segments in terms of socio-demographic and travel behavioural variables. The third step was to select target markets based upon the earlier analysis. These methods were applied to a sample of tourists, over the period of a week, on Phillip Island, Victoria, Australia. A significant outcome of this research is that it will assist tourism organisations to identify tourism market segments and develop better tour packages and more efficient marketing strategies aligned to the characteristics of the tourists.

1. Introduction

The movement of tourists is a complex process which can be modelled at a micro level as a continuous process with high resolution, such as in centimetres or at a macro level as discrete processes with low resolution, such as kilometres from one area to another. Tourist movement patterns are a theme of recurring or repeat movement sequences. This paper focuses on movement patterns at the macro level. Movement patterns represent the sequence of movements by tourists from one-attraction site to another. The dominant movement patterns are the sequences or patterns that are used by tourists most frequently. These tourist movement patterns are vital to park managers or tour operators to understand the location of popular sites and the timing of visits. More importantly, an understanding of movement patterns can indicate how tourists combine attractions together and arrange their schedules.

Traditionally, tourism market segmentation is conducted to identify groups of tourists from an origin perspective, for example, analysis of the origin of tourists from China or Australia. Alternatively, analysis can be performed from a destination perspective, where the tourism market is segmentation-based on *a single* destination such as Melbourne. However, tourists usually visit several attractions during a trip. To understand the spatial combination of attractions and to clarify the characteristics of tourists who travel to these attractions will assist tourist organisations to design more appropriate and profitable tour packages.

The aim of this study was to develop a methodology to identify the characteristics of tourists who travel with dominant movement patterns. The first step involved identification of the dominant movement patterns based on the analysis of categorical data using general log-linear models. In the second step, the Expectation–Maximisation (*EM*) algorithm used in a mixture model framework was adopted to identify the characteristics of tourists who travel with dominant movement patterns. Tourists were divided into different segments based on similar socio-demographic characteristics and travel modes such as type of travel group, modes of transport or visit frequency. For example, the tourists in segment 1 who travel with pattern *A* may mostly be females travelling with their family by car. Finally, appropriate target markets can be

* Corresponding author. Tel.: +61 8 92667563; fax: +61 8 92662703.
E-mail addresses: c.xia@curtin.edu.au (J.(Cecilia) Xia), fevans@agric.wa.gov.au (F.H. Evans), katrina.spilsbury@curtin.edu.au (K. Spilsbury), vc@cs.rmit.edu.au (V. Ciesielski), colin.arrowsmith@rmit.edu.au (C. Arrowsmith), g.l.wright@curtin.edu.au (G. Wright).

0261-5177/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tourman.2009.04.013

determined from the socio-demographic data. Sections 2 and 3 of this paper address the first two steps, followed by a case study of tourists visiting Phillip Island Nature Park in Victoria, Australia.

2. Step one: identification of dominant movement patterns of tourists using general log-linear models

This section discusses the methodology used to identify the dominant movement patterns of tourists. Dominant movement patterns were defined as the combination of attractions used most frequently by tourists. The reasons for associations between attractions are diverse. Tourists could combine several attractions together as one trip because those attractions may be in close proximity, pairs of attractions might have been packaged by operators, the combination of these attractions fit to tourist desire or interests, or attraction site managers developed promotion policies, for example, discounted entrance fees for combined attractions. For economic reasons, tourists might visit several attractions to save money or time by optimising their trips.

Generally, more than two attractions are visited by tourists during a trip, which means more than two categorical movement variables are concerned in tourist movement modelling. These multi-way contingency tables of frequency counts can be analysed using symmetrical general log-linear models under the assumption that the counts follow a Poisson distribution (Agresti, 1990; Goodman, 1984). In order to fit tourist movement data to the general log-linear model, patterns of movement can be coded by a series of destination variables such that each represents a stage in the pattern of movement. For example, if a set of three attractions *A*, *B* and *C* were available for tourists to visit in a day, then for each tourist the variable Destination 1 includes the attraction first visited and variable Destination 2 would include the next attraction visited out of the two remaining attractions, and Destination 3 would include the last attraction visited. A log-linear model calculates the expected number of movement pattern counts for each destination combination as if there was no difference between the movement patterns. Chi-square goodness-of-fit tests are then used to compare these expected values to the observed counts. An association between two (or more) attractions can be considered dominant (significant) if the chi-square test p-value is <0.05. It may be observed, for example, that tourists who visit attraction *A* almost always immediately follow this with a visit to attraction *C* leading to a dominant movement pattern of *AC* in contrast to the less common movement pattern of *AB*.

3. Step two: tourist market segmentation methods

In this section we discuss methods that can be used to generate tourism market segments for each of the significant movement patterns identified in step 1. For instance, the significant movement patterns might be (OPERA-HOUSE, HARBOUR-BRIDGE) and (BONDI-BEACH, OPERA-HOUSE). Segmentation can be performed on tourist data, and the proportions of market segments matching the patterns (OPERA-HOUSE, HARBOUR-BRIDGE) and (BONDI-BEACH, OPERA-HOUSE) can be calculated and used to gain insight into the types of tourists that follow those movement patterns.

3.1. Determination of segmentation variables

This section describes a method to determine the tourist market segmentation variables. Generally tourist segments are divided in terms of geographic, socio-demographic, psychographic and travel behavioural variables (Wedel & Kamakura, 2000). However, the determination of segmentation variables is dependent upon the objectives of tourism market segmentation. For example, Bigné and Andreu (2004) identified tourist emotion-based segments and analysed which segment was most satisfactory for leisure and tourism services based on socio-demographic variables, multiple-item scales of emotions, satisfaction and behavioural intentions. The objectives of market segmentation are summarised by Myers (1996) as follows:

- Identifying and characterising groups of tourists
- Focusing advertising efforts for greater impact
- Identifying likely targets for new tourist products
- Improving existing product/service design
- Looking for new product service opportunities
- Assessing the impact of a competitor's new offering
- Establishing a better tourist attraction image

The objective of the tourism market segmentation in this paper belongs to the first category above – identifying the characteristics of groups of tourists who travelled with the same significant spatial movement pattern. Therefore, the geographic, socio-demographic and trip-related behavioural variables are used.

3.2. Tourist market segmentation methods

Market segmentation is a process of dividing a market into homogeneous subgroups. Tourists in the same group are similar to each other, and different from other groups, in the way they react to the market mix such as promotions or advertising which then influences their spatial behaviours (Weinstein, 2004). Market segmentation has been extensively used as a valuable method to identify tourist market groups, select the target market and position the tourist market (Chandra & Menezes, 2001; Haley, 1968; Kotler & Armstrong, 2003; Lee, Morrison, & O'Leary, 2006; Mykletun, Crotts, & Mykletun, 2001). Market segmentation is equivalent to data clustering and as such, algorithms developed in the area of clustering can be applied here.

Model-based clustering, also known as mixture modelling, assumes that each cluster has an underlying probability density function or model, and that the density for the mixture is a weighted sum of the cluster densities. That is, given observed data y_i for $i = 1,\ldots,N$, where each y_i is a vector of M measured variables, the probability density function p for a mixture model composed of K clusters is given by

$$p(y|\theta) = \sum_{j=1}^{K} \alpha_j p_j(y|\theta_j) \tag{1}$$

where $p_j(y|\theta_j)$ is the probability density function for the j-th cluster and the mixture weights α_i represent the proportions of data belonging to the j-th cluster such that $\sum_{j=1}^{K} \alpha_j = 1$. The parameters θ of the mixture model comprise the parameters θ_i of the cluster densities as well as the mixture weights α_i. An advantage of model-based clustering is that it can be used for both numeric and categorical data. Conditional on cluster membership, numeric variables can be modelled by normal distributions and categorical variables can be modelled using multi-way cross-classification tables, with all variables assumed conditionally mutually independent.

The Expectation–Maximisation (*EM*) algorithm for incomplete data (Dempster & Laird, 1977) can be used to perform maximum likelihood parameter estimation for mixture models. It applies the principle of maximum likelihood to find the model parameters, by iteratively repeating the Expectation (*E*) and Maximisation (*M*) steps after randomly initialising the mixture model parameters. The *E* and *M* steps are iterated until a desired convergence is achieved (Witten & Frank, 2000).