

NOTICE: This is the author's version of a work that was accepted for publication in *Tourism Management*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Tourism Management*, 32, 4, 2010 DOI 10.1016/j.tourman.2010.07.009

## **Spatial and Temporal Modelling of Tourist Movements using Semi-Markov Processes**

**Xia, Jianhong (Cecilia) and Zeephongsekul, Panlop and Packer, David. 2010. Modelling Spatio-temporal Movement using Semi-Markov Chains. *Tourism Management* 32 (4): pp. 844-851.**

**Abstract:** Tourist movement is a complex process. It can be modelled from a number of different perspectives; for example, Tourism, Geography, Economics, Mathematics, Computer Sciences and Psychology. This paper aims to establish a sound methodology, using Semi-Markov processes to model the spatial and temporal movement of tourists. The objective is to understand, predict, control for, and optimise the decisions made by tourists in their choice of attractions. Semi-Markov processes have a Markov chain and a renewal process embedded within their structure, and as such, can be used to provide a wide variety of practical models. One of the outcomes of this approach is the introduction of a measure to assess the attractiveness of particular tourist attractions based on spatial and temporal interactions between the attractions. A case study conducted at Phillip Island Nature Park, Victoria, Australia is used to validate the model. The study's results support the proposed model's efficiency. The results are also practical and useful for assisting park managers and tourist operators with park planning and marketing decisions: for example, knowing which attractions are the most popular, how long tourists will spend at any one site; and what the likely routes are that they will follow.

**Keywords:** Semi-Markov process; tourist movement; spatial and temporal modelling

## **1. Introduction**

Tourist movement research is a complex process of understanding physical movement and its link to the decisions tourists make about where, how, and at what pace and time to move from one attraction to the next. The stream of research into tourists' physical movement at a macro level examines a discrete sequence of movements among tourist locations - locations which can be some distance apart. At a micro level, movement is seen as a continuous process (Bierlaire 1998, Hägerstrand 1970, Haldrup, 2004; Lew and McKercher 2006, Miller 2005, Xia, Zeepongsekul and Arrowsmith, 2009, Yamamoto, Kitamura and Kishizawa 2001). The sequence of movements can be represented accurately as a collection of spatial points (x, y) in a coordinate system and tracked using GPS (Global Positioning Systems) (Xia, Arrowsmith, Jackson and Cartwright 2008). Physical movements are visible and, therefore, can be quantified, measured and predicted.

The second major stream of tourist movement research concentrates on tourist wayfinding decision-making, which is understanding the process of deciding which path to traverse when faced with a series of alternatives (Forer, 2002, Golledge, 1999, Itami 2003, Xia, Arrowsmith, Jackson and Cartwright 2008). Knowing more about wayfinding decision-making can be useful to tourism planners, managers and operators in that tourist decisions affect tourist routes and facilities at the macro level.. At the micro level, the planners and developers of tourist routes within individual tourist attractions will benefit from knowing

more about the wayfinding decisions of tourists as the planners themselves make decisions about signage, rest stations and food stops, paths, and internal transport.

Few studies have attempted to model the spatial movement of tourists (Lew and McKercher 2006, Pearce 1988; Fennell 1996). Lew and McKercher (2006), for example, suggested that certain destination and tourist characteristics influence decision-making and thus, tourist behaviour. They then inferred several path and territorial models related to the set of identified characteristics. This modelling was, perhaps, the best attempt so far to model spatial movement, but even these authors did not integrate time into the model, something that has had even less attention in the literature (Haldrup 2004, Pearce 1988; Truong and Hensher 1985). Furthermore, time spent at an attraction can be a measure of the attractiveness of that place, if integrated into a model that can predict probabilities of movement patterns.

Markov Chain models have been used extensively for modelling sequences of events related to each other by a first-order dependency (Cassandras and Lafortune 2008, Isaacson and Madsen 1976; Kemeny and Snell 1976; Stewart 1994;). In a few instances, Markov Chains have been used in tourism research. Mednick (1975) models the travel behaviour of tourists in the United States using Markov chain analysis by predicting the probabilities of various travel patterns. Markov chain models are adopted by Lusseau (2003) to assess the effects of tourism activities on the behavioural patterns of small cetaceans. Xia, Zeepongsekul and Arrowsmith (2009) utilise Markov Chains to analyse the outcomes and trends of events associated with spatio-temporal movement patterns among tourists.

Without a time dimension, i.e., arrival time or time spent at each state included in the modelling process, Markov Chains may predict probabilities of movement patterns but not allow for more robust prediction of tourists moving from one attraction to another or the route they may choose to follow. A Semi-Markov process has a Markov chain and a renewal process embedded within its structure. In the renewal process, it is assumed that events occur randomly and the times between the successive arrivals are independent and identically distributed (Cinlar 1975b). Therefore, Semi-Markov processes can be used to predict the probability of movement patterns, given that we know the time spent at each attraction. Furthermore, time spent at each attraction can be a measure of the attractiveness of a tourist attraction.

This paper examines tourist discrete physical movements among attractions at a macro level, integrating the time spent at attractions into Semi-Markov movement models to estimate the probability of visiting attractions and to arrive at a measure of attractiveness of attractions. Section 2 defines tourist movement and describes the modelling of tourist movement, integrating time using Semi-Markov processes. The approach is to use Semi-Markov processes to assess attractiveness of particular tourist attractions based on spatial and temporal interactions between the attractions. In Section 3, 4 and 5, a case study conducted at Phillip Island Nature Park, Victoria, Australia is used to test the model. The paper concludes with a discussion of the potential applications of Semi-Markov modelling for tourism developers and manager, with suggestions for further work in this area.

## **2. Method**

### **2.1. Tourist movement definition**

Tourist movement, a random spatial and temporal process, can be formally defined as a family of random variables  $\{X_t, t \in T\}$  in a given probability space and indexed by  $t$ . The set  $T$  is often used to represent the time sequence of the process and is usually discrete, i.e.  $T = \{0, 1, 2, \dots\}$  or continuous, i.e.  $T = [0, \infty)$ . The set of values of  $X_t$  as  $t$  ranges over  $T$  is the state space  $S$  of the random process, which could again be discrete or continuous. In this paper, we define a tourist movement as a discrete random process. The state space  $S$  contains the destinations and transit route locations traversed by the tourists. For example, a tourist trip from Melbourne to Sydney and then, to Perth can be recorded as the discrete process where  $T$  is discrete and  $X_1 = \text{Melbourne}$ ,  $X_2 = \text{Sydney}$  and  $X_3 = \text{Perth}$  respectively.

## 2.2. Semi-Markov Processes

Mathematically, the discrete Semi-Markov random process assumes that the cardinality of the set  $S$  is equal to a finite number  $N$  and that set  $T$  is continuous. Let  $X_0, X_1, X_2, \dots$ , be the successive states in  $S$  visits by a random process starting in state  $X_0$  and let  $0 = T_0 < T_1 < T_2, \dots$  be the times of transitions into each of these states. The random process is a Semi-Markov process if it satisfies the following fundamental property (Cinlar 1975b):

$$\begin{aligned} \Pr(X_{n+1} = j, T_{n+1} - T_n \leq t \mid X_0, X_1, \dots, X_n; T_0, T_1, \dots, T_n) \\ = \Pr(X_{n+1} = j, T_{n+1} - T_n \leq t \mid X_n) = Q_{ij}(t) \end{aligned} \quad (1)$$

for  $n = (0, 1, 2, \dots)$ . In words, given that we know the states (attractions) that the movement process will be visiting and the epochs where the process transits into these states up to and including the  $n$ th epoch, the probability distribution of where the movement visits the next attraction, i.e.  $X_{n+1}$ , and the time interval between transitions from the current attraction to the next depends only on the current state  $X_n$ . Note that for a Markov process, the

distribution of time interval between transitions from the current attraction to the next ( $T_{n+1} - T_n$ ) does not only depend on  $X_{n+1}$  due to the “lack of memory property”; it is also restricted to the class of exponential distributions determined by the infinitesimal rates of the process (Isaacson and Madsen, 1976). In this sense, a Semi-Markov process generalises a Markov process since the distribution of the time between transitions is now quite general. This time interval is no longer an exponential distribution, as is the case with a Markov process but could be other distributions such as a normal distribution.

The following two assumptions underlie the Semi-Markov process, making it an ideal tool for modelling spatio-temporal movements of tourists:

- The probability that a tourist will visit a particular attraction depends only on the most recent attraction that was visited by that tourist.
- The distribution of the time spent in each attraction is dependent on both that attraction and the attraction that is visited next and is not necessarily exponentially distributed.

The probability on the right hand side of (1) is called the Semi-Markov kernel of the process and is denoted by  $Q_{ij}(t)$ , i.e.

$$\Pr(X_{n+1} = j, T_{n+1} - T_n \leq t \mid X_n = i) = Q_{ij}(t) \cdot \quad (2)$$

This is the probability of tourists next visiting attraction  $j$ , given that the current attraction visited is  $i$ , and the time between arrivals to these attractions,  $T_{n+1} - T_n$ , does not exceed  $t$ .

Note that  $T_{n+1} - T_n$  is the sum of the duration spent at attraction  $i$  and the travel time between attractions  $i$  and  $j$ .

If we let  $t \rightarrow \infty$ , then since  $Q_{ij}(t)$  is a cumulative distribution function in  $t$ , we obtain

$$\begin{aligned} Q_{ij}(\infty) &= \Pr(X_{n+1} = j | X_n = i) \\ &= P_{ij} \end{aligned} \quad (3)$$

which is just the one-step transition probabilities associated with the underlying Markov chain.

Another relevant distribution associated with Semi-Markov processes is

$$F_{ij}(t) = \Pr(T_{n+1} - T_n \leq t | X_{n+1} = j, X_n = i) \quad (4)$$

which uncouples the two variables,  $X_{n+1}$  and  $T_{n+1} - T_n$  in  $Q_{ij}(t)$ . This is the conditional distribution of  $T_{n+1} - T_n$  given one has information that  $X_n = i$  and  $X_{n+1} = j$ ; i.e. if we know that the current attraction visited is  $i$  and the next attraction is  $j$ , then  $F_{ij}(t)$  is the probability distribution of the time between arrivals at these attractions. As will be demonstrated,  $F_{ij}(t)$  can be used to estimate the mean and variation of the time spent at attraction  $i$ . The three probability distributions defined by equations (2), (3) and (4) are linked. Using the product rule of probability, it follows that

$$\begin{aligned} \Pr(X_{n+1} = j, T_{n+1} - T_n \leq t | X_n = i) \\ = \Pr(T_{n+1} - T_n \leq t | X_{n+1} = j, X_n = i) \Pr(X_{n+1} = j | X_n = i) \end{aligned} \quad (5)$$

that is,

$$F_{ij}(t) = \frac{Q_{ij}(t)}{P_{ij}} \quad (6)$$

if  $P_{ij} > 0$ ; otherwise,

$$F_{ij}(t) = H(t) \quad (7)$$

where

$$H(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0. \end{cases}$$



If  $P_{ij} = 0$ , then it is not possible to move from site  $i$  to site  $j$ . Recall that  $F_{ij}(t)$  defined by equation (6) is the distribution of the time  $T_{n+1} - T_n$  it takes to move between the two sites  $i$  and  $j$ , but if this transition is not possible then  $T_{n+1} - T_n = 0$ .  $H(t)$  is just the probability distribution that a random variable takes on the single value 0.

Based on the key assumption (1) and applying a series of concatenations using the product rule, the *joint distribution* of times  $T_{i+1} - T_i$ ,  $i = 0, 1, \dots, n$ , given  $X_0 = x_0, X_1 = x_1, \dots, X_n = x_n$  is computed using the following rule:

$$\begin{aligned} & \Pr(T_1 - T_0 \leq t_1, T_2 - T_1 \leq t_2, \dots, T_{n+1} - T_n \leq t_n \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \\ & = F_{x_0 x_1}(t_1) F_{x_1 x_2}(t_2) \dots F_{x_n x_n}(t_n) \end{aligned} \quad (8)$$

Based on equation (8), if  $S$  consists of a single state, then a Semi-Markov process also reduces to a *renewal process*.

The advantage of Semi--Markov processes is that they have a Markov chain and a renewal process embedded within their structure, and can be used to predict the mean or deviation of duration and arrival time at an attraction, transition time between attractions, and the probability of visiting an attraction or a sequence of attractions. Associations among pairs of attractions, such as mean time spent at a specific attraction before moving to the next attractions or the probability of visiting an attraction - given that a specific attraction was visited previously - can also be modelled using this method.. We systematically summarise these applications into a Markov tourist movement application matrix (See Figure 1).

Figure 1. Markov Tourist Movement Application Matrix

In our study, we develop an application to assess the attractiveness of particular tourist attractions using the Semi-Markov modelling approach. This application is useful for tourist site design, tour package design, tour scheduling, and park facility management, such as location of benches and toilets and staffing requirements.

### 2.3. Attractiveness of an attraction

Based on the assumptions of the Semi-Markov process, two scenarios are derived for assessing attractiveness of tourist attractions.

- The more tourists visit an attraction, the more attractive it is.
- The longer tourists stay at an attraction, the more attractive it is.

From the first assumption, we developed a transition probability matrix for estimating the probability that a tourist will visit a particular attraction. The theory and detailed examples in regards to the calculation of a transition probability can be found in Kemeny and Snell (1976) and Xia, Zeephongsekul and Arrowsmith (2009). We also calculated a mean time transition matrix, based on the second assumption, to estimate the time spent at each attraction.

#### 2.3.1. Mean time transition matrix

Let  $T_{ij}$  represent the time a tourist spends in state  $i$ , given that the process starts in  $i$  before moving to state  $j$ . The mean value of the random variable  $T_{ij}$  is computed by

$$E(T_{ij}) = \int_0^{\infty} tF'_{ij}(t)dt - t_{ij} \quad (9)$$

where  $F'_{ij}(t)$  is the Probability Density Function (PDF) of  $T_{ij}$  and  $t_{ij}$  is the expected travel time between the state  $i$  and state  $j$ . Note that the first term on the right hand side of (9) is the expected time it takes for a tourist to move between state  $i$  and  $j$ . This includes the time it will take to travel between the two states; therefore, subtracting  $t_{ij}$  from this term gives the expected time spent in state  $i$ . Note that  $t_{ij}$  can be estimated easily, since the travel times between various destinations are known. These mean times can be conveniently displayed, like one-step transition probabilities, in matrix form as:

$$E(T_{ij}) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ E(T_{A_2(n),A_1(n+1)}) & 0 & E(T_{A_2(n),A_3(n+1)}) & \dots & E(T_{A_2(n),A_k(n+1)}) \\ E(T_{A_3(n),A_1(n+1)}) & E(T_{A_3(n),A_2(n+1)}) & 0 & \dots & E(T_{A_3(n),A_k(n+1)}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ E(T_{A_k(n),A_1(n+1)}) & E(T_{A_k(n),A_2(n+1)}) & E(T_{A_k(n),A_k(n+1)}) & \dots & 0 \end{pmatrix}$$

Let  $S = \{A_1, A_2, \dots, A_k\}$  where  $A_i$ ,  $i = 2, 3, \dots, k$  are the tourist attractions and  $A_1$  represents the state “OUT”, i.e. the region exterior to the space  $S$  where the tourist sites are located. The Distribution Fitting Tool in the Matlab software package (The Mathworks Inc, 2008) was used to fit different distribution functions to the time spent in each destination. Twenty one distribution functions are available in Matlab, and we explored all these distribution functions to fit our data. However, in order to choose the best fit distribution, a more objective criterion had to be adopted: Akaike Information Criterion (AIC) (Akaike 1973) defined by

$$AIC = 2 \times (\text{number of parameters in the model}) - 2 \times \text{Log Likelihood} \quad (10)$$

was chosen to evaluate the goodness of fit of each distribution. The smaller the value of AIC, the better the distribution fit of the data. Note that AIC utilises the log-likelihood function determined by the data set, with a penalty imposed based on the number of parameters estimated. Once the best fit distribution function was selected, equation (9) was used to estimate the mean time spent at various attractions.

### 2.3.2. *A measure of attractiveness*

A measure of attractiveness of state  $i$ , denoted by  $A(i)$ , is defined by

$$A(i) = \sum_{j=1}^N P_{ij} E(T_{ij}) \quad (11)$$

where  $N$  is the total number of states. Since we multiplied  $E(T_{ij})$  by  $P_{ij}$  and then summed over all state  $j$ , equation (11) is the expected time spent in state  $i$ , irrespective of which state the process moves to next. Hence,  $A(i)$  is not an unreasonable measure of the attractiveness of that particular attraction.

## 3. Case study area

Phillip Island, located at the mouth of Westernport Bay, Australia, is 140 kilometres south-east of Melbourne. There are a large number of penguins, koalas, seals, and shearwaters living in the mangroves and wetlands, and on the sandy beaches and rugged rocky cliff faces. The major attractions are the Penguin Parade (G), the Koala Conservation Centre (D), Cowes (F), Churchill Island (C), Rhyll Inlet (E), Woolamai (B) and the Nobbies (H) (Figure 2). On the island, visitors can experience wildlife in its natural environment (Phillip Island Nature Park 2005).

Figure 2 Map of Phillip Island (Phillip Island Nature Park 2005)

## **4. Sample Survey**

Tourists' daily movement data for Phillip Island Nature Park were collected via a self-administered questionnaire. The questionnaire was designed to address three different sections. The first section aimed to acquire socio-demographic data or profiles of the tourists. The second section aimed to collect information regarding travel mode, length of stay and with whom the tourists were travelling. The final section gathered information on tourist movement. In the third section, tourists were asked to write down their approximate arrival time and duration of stay at each attraction visited for the entire day. Tourists were also asked to draw the route of travel to each attraction on a street map of Phillip Island. Eight hundred questionnaires were distributed from the 6-8th of March 2004 and from 17-20th Jan 2005 at the Phillip Island Nature Park Information Centre, Churchill Island, Koala Conservation Centre and Penguin Parade. Penguin Parade was the major sampling site, as it is usually the last attraction for tourists to visit: they come to Penguin Parade just before sunset to watch the fairy penguins. Park managers estimate that more than 90% of tourists entering the park visit these four attractions; it is therefore unlikely that significant movement patterns were missed by the survey. Five hundred questionnaires were returned with 456 entered into the database. The remaining 44 questionnaires were incomplete and discarded.

## **5. Results**

### **5.1. Transition probability matrix**

In our case study, we consider the movement of tourists between the nine attractions on Phillip Island shown in Figure 2 above. A stationary discrete Markov Chain is used to

model the movements of the tourists between each attraction, from the moment they entered the park until they completed their visits. The states of the chain are the nine attractions visited with an additional absorbing state labelled, "OUT", which signalled the completion of their tour. Movement of tourists on Phillip Island, therefore, is broken down into ninety one-step transitions in a transition probability matrix (See Table 1). The table reads from left row to right column as one pair of the transition. For example, from Cowes (F) to Penguin Parade (G), the transition probability is 56.8%, which means 56.8% corresponded tourists on Philip Island have moved from Cowes (F) to Penguin Parade (G). The other highly associated pair attractions are The Nobbies (H)  $\rightarrow$  Penguin Parade (G) (74.6%), Koala Conservation Centre (D)  $\rightarrow$  Cowes (F) (48.1%), Cowes (F)  $\rightarrow$  The Nobbies (H) (33.9%), Cape Woolamai (B)  $\rightarrow$  Cowes (F) (28.3%), and Cape Woolamai (B)  $\rightarrow$  Koala Conservation Centre (D) (21.7%) . These pairs of attractions are considered to be more attractive to tourists than other pairs of attractions. For details concerning the computation of the entries of this matrix, the reader is referred to Xia, Zeepongsekul and Arrowsmith (2009).

Table 1 Transition Probability Matrix

## 5.2. Mean time transition matrix

The mean time transition matrix computed using equation (9) shows the mean time spent at attraction  $i$ , given that the movement starts at  $i$  and then moves to attraction  $j$ . The dimensions of the mean time transition matrix are the same as the transition probability matrix (see Table 1) which includes nine attractions and one absorbing state labelled as "OUT".

A range of possible probability density functions (PDF) were adopted to estimate time spent at attractions. Figure 3 illustrates the range of possible probability density functions (PDF) for time spent at attraction *F* before moving to *G*. As the set of possible distributions is large, Akaike Information Criterion (AIC) (Akaike 1973) was adopted to choose the best fit distribution based on numbers of parameters. The smaller the value of AIC, the better the distribution fits the data (see Table 2). In our case study, the lognormal distribution function provides the best fit distribution for time spent at attraction *F* before moving to *G*. We recognise that even though the more common distributions, such as gamma, exponential and Weibull, failed to adequately fit all the times spent at different sites, this does not mean that the model is not good enough; it just means that the available distributions fail to capture all the data and if we wanted to clarify all points, other distributions would be required. For the same reason, we should not expect equation 9 to be the same for all pairs of transitions since tourists spent different amount of times and exhibited different behavioural patterns at different sites, thereby resulting in different fit distribution functions.

Figure 3 Probability density functions (PDF) for time spent at attraction *F* before moving to *G*

Table 2 AIC for Each Distribution Function

The same procedure was repeated 90 times to identify the best fit distribution for each pair of transitions (see Table 3). The value zero occurring in the table means no tourists in the survey data did this transition; e.g. from *A* to *B*. This is because the majority of tourists visiting Woolamai (*B*) are domestic visitors and since they are familiar with Phillip Island, there is no need to visit the Information Centre. In Table 3, "?" indicates that there are no distributions among the ones listed that fit the data satisfactorily.

Table 3 Distribution Fit

After selecting the best fit distribution, mean time is calculated for each pair of transitions (See Table 4). Table 4 reads similar to Table 1. For example, from Cowes (F) (left) to Penguin Parade (G) (up), the mean time is 2.39 hours, which means tourists spent an average of 2.39 hours at Cowes (F) before moving to Penguin Parade (G). We can see the variation in the mean time tourists spent at an attraction before moving to different attractions and might conclude that time spent depends on the nature of the attraction as well as tourist motivation to explore the attraction thoroughly. Generally, tourists spent longer at Cowes (F) and Churchill Island (C), at over 2 hours, than at any other attraction on the island. Therefore, we conclude that these two attractions hold the highest attraction for tourists to the island. We have used average duration to calculate mean time when the best distribution cannot be identified because of the small sample size.

Table 4 Mean Time Transition Matrix

### 5.3. Attractiveness of an attraction

The final step is to calculate the attractiveness of each attraction using equation (11) (see Table 5). Attractiveness of attraction  $F$  is the expected time spent at attraction  $F$  irrespective of which state the process moves to next (refer to Equation (11)). For example, the attractiveness of attraction  $F$  ( $i = 6$  and  $N = 10$ ) is calculated as:

$$A(6) = \sum_{j=1}^{10} P_{6j} E(T_{6j}) = 0 \times 0 + 0.017 \times 1.75 + 0.01 \times 2.5 + \dots + 0.014 \times 1.767 = 2.238$$



Table 5 reveals the attractiveness of each attraction (A(i)). This measure was calculated based on transition probabilities and the mean time of transitions between different pairs of attractions. Higher values indicate higher attractiveness. Based on this measure, G is the most attractive attraction on Phillip Island. We conclude that there is a higher probability of tourists visiting G, and that they will stay there longer than any other attraction. Table 5 also describes the space and time interaction of pairs of attractions. The higher values associated between two attractions, e.g.  $G \rightarrow \text{OUT}$  (2.4),  $F \rightarrow G$  (1.36) and  $H \rightarrow G$  (0.85) indicate the relative attractiveness of pairs of attractions for tourists. For example, relatively more tourists preferred to have dinner or lunch at Cowes (F), and they spent relatively longer there before moving to Penguin Parade (G) to see the little fairy penguins.

Table 5 Attractiveness of Attractions

The advantage of the Semi-Markov model is that it helps to elucidate the interaction between attractions along the movement chain, i.e., spatial and temporal interaction between two or more attractions. For example, tourists visiting attraction A have a higher chance of visiting attraction B, and tourists are likely to spend more time at attraction A before moving to B. The matrix described by Figure 4 illustrates the attractiveness of pairs of attractions by the size of the dots, based on the figures in Table 5. The scale of the dots is 1:20. The row shows the original attractions and the columns are the next attractions visited. Using the size of the dots as indicators, G (left)-OUT (up) is the most attractive tourist attraction pair on Phillip Island. The other attractive attraction pairs are F-H, F-G, H-G, and C-F. Figure 4 also shows the interaction between more than two attractions. The paths linking attractions show the movement patterns. The red path illustrates the most

attractive four attraction movements (CDFG). The purple and green paths show two three-attraction movements (DFH & FHG) that are also attractive.

Figure 4 Attractiveness of Attractions

## 6. Discussion

The aim of this study was to utilise Semi-Markov processes to model the decisions made by tourists in their choice of attractions by assessing both the spatial and the temporal interactions between the tourists and the tourist attractions. A case study conducted at Phillip Island Nature Park, Victoria, Australia was used to validate the model. The findings of this study have implications for tourism management, especially for destination management, on-site movement planning, and attraction marketing.

**6.1 Destination administration and design:** A Semi-Markov Chain model can estimate the probability of visits at attractions; this data can provide policy-makers with a powerful tool to understand the way in which tourists travel in and around the parks and to allocate resources to meet the greatest needs of tourists (Connell and Page 2008). **At Phillip Island,** for example, park managers can use the outcomes of the transition probability matrix to estimate the number of tourists who transit from one attraction to another, i.e.; The Nobbies to Penguin Parade. Then, based on this information, they can decide on the placement of roads and pathways, which is a costly outlay for any tourist attraction. Furthermore, the estimation can help refreshment and food preparation, as this research indicates that most tourists who transit from The Nobbies to Penguin Parade on Phillip Island stop for a meal at the café nearby. At other facilities, knowing where and when

tourists are looking for food would assist in the location decisions and prevent less than optimal use of resources.

Therefore, by integrating time into the model, the Semi-Markov chain can be used to estimate time spent at attractions and travel time between attractions, as well as the probability of routes to travel, given a certain time spent at attractions. Based on this information, destination management teams can create a more sustainable environment for tourists to travel in by providing suitable and economical transport modes, convenient transport schedules, and good signage to enhance the tourist experience.

**6.2 On-site movement planning:** Individual tourist sites can utilise the modelling of tourist movements in time and space to their advantage. For example, decisions on the location of toilets, benches and cafes, timing of the maintenance of roads and toilets, staff allocations, and tours can be made based on tourist spatial and temporal movement information. This research indicates that tourists usually spend an average of four hours on Churchill Island before moving to The Nobbies. Matching attractiveness analyses with demographic data, we find that many tourists at Phillip Island are elderly and have a European background. They tend to have lunch and coffee and spend about two hours walking around the Nobbies. Knowing the length of time they will spend at the Nobbies, and that they will be continuing on a physical movement in this attraction about 2 hours tells planners that these elderly tourists will require more benches so they can sit down and rest as they need it.

**6.3 Attraction marketing:** Knowledge of tourist movements in time and space is particularly useful for marketing, organising, and scheduling individual and group tours. In

combination with knowledge of tourist market segmentation, marketers can facilitate tour package designs, and promote experiences more accurately, meaning that there would be more clarity in overall destination image, allowing for the creation of place identities aimed at specific market niches (Lew and McKercher 2006; Xia et al. 2010). For example, the attractiveness of attractions tool (Figure 4 and Table 5) can assist park managers in making decisions about which attractions, pairs of attractions, or combination of attractions are more attractive. Once these attractive movement patterns are identified, further analysis can be carried out to overlay the characteristics of tourists, such as gender and travel modes, with attraction patterns. Then, based on this information, tour packages can be designed and marketed to appeal to specific tourist groups. For details of this research, the readers are referred to (Xia et al 2010).

## **7. Conclusions**

The majority of tourist movement modelling studies have focused on predicting spatial movement patterns of tourists (Lew and McKercher 2006; Xia, Zeephongsekul and Arrowsmith 2009). Our research developed the modelling further by integrating time into Semi-Markov modelling processes. The modelling included both spatial and temporal movement information. Semi-Markov processes in this model could be broken down into one-step transitions, enabling us to identify how pairs of attractions associate with each other spatially and temporally. Pairs of attractions, then, were joined as a sequence of movements to estimate probability of tourist movement patterns, based on time spent at attractions. Using data on tourist movements on Phillip Island, Victoria, Australia we applied Semi-Markov processes to measure the attractiveness of tourist attractions and the associations of pairs of attractions.

In summary, this method simplifies the measurement of attractiveness and given the relatively straightforward assumptions, the results indicate that a method utilising Semi-Markov processes is effective in deriving probabilities of both tourist movement and attractiveness of specific attractions. In the future, further exploration of the applications of other stochastic processes can be undertaken, such as scheduling procedures to predict arrival times, duration at different attractions, and scheduling of movement between sites in order to improve the information planners, designers, marketers and operators of tourist destinations have on hand to make decisions to enhance the tourist experience.

## References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *2nd Inter. Symposium on Information Theory*. Budapest.
- Bierlaire, M. (1998). Discrete choice models. In Labb, M., Laporte, G., K.Tanczos & Toint, P. (Eds.) *Operations Research in Traffic and Transportation Management*. (pp. 203-227). Springer Verlag.
- Cassandras, C. G. and Lafortune S. (2008). *Introduction to discrete event systems* (2nd ed.) New York :, Springer.
- Cinlar, E. (1975b). *Introduction to Stochastic Processes*. Englewood Cliffs, NJ: Prentice-Hall Cover..
- Connell, J. & Page, S. J. (2008) Exploring the spatial patterns of car-based tourist travel in Loch Lomond and Trossachs National Park, Scotland, *Tourism Management*, 29(3), 561-580.
- Fennell, D. A. (1996). A tourist space-time budget in the Shetland Islands. *Annals of Tourism Research*, 23(4), 811-829.
- Forer, P. (2002) Serial Experiences: Monitoring, modelling and visualising the free independent traveller in New Zealand at multiple scales with GIS. . In Arnberger, A., Brandenburg, C. & Muhar, A. (Eds.) *Monitoring and Management of Visitor Flows in Recreational and Protected Areas*. Bodenkultur University Vienna, Austria.

- Golledge, R. G. (1999) Human wayfinding and cognitive maps. *Wayfinding Behaviour: Cognitive Mapping and Other Spatial Processes*. (pp. 5-45). Baltimore: The Johns Hopkins University Press.
- Haldrup, M. (2004) Laid-Back Mobilities: Second-home holidays in time and space. *Tourism Geographies: An International Journal of Tourism Space, Place and Environment*, 6(4), 434 - 454.
- Hägerstrand, T. (1970) What about people in regional science? *Papers of Regional Science Association*, 247-21.
- Itami, R. M. (2003). *RBSim3: Agent-based simulations of human behaviour in GIS environments using hierarchical spatial reasoning*. ModSim 2003 International Congress on Modelling and Simulation, Townsville, Australia.
- Isaacson, D. L. and Madsen R. W.(1976). *Markov chains, theory and applications*. New York: Wiley.
- Kemeny, J. G. and. Snell J. L (1976). *Finite Markov chains*. New York Springer-Verlag.
- Lew, A. and B. McKercher (2006). Modeling tourist movements: A local destination Analysis. *Annals of Tourism Research* 33(2): 403-423.
- Lusseau, D. (2003) Effects of tour boats on the behavior of bottlenose dolphins: Using markov chains to model anthropogenic impacts, *Conservation Biology*, 17(6), 1785-1793.
- Mednick, H. (1975). A Markov Chain model of travel patterns of U.S. visitors to Ontario. *Journal of Leisure Research* 7(4), 246-255.
- Miller, H. J. (2005) A measurement theory for time geography *Geographical Analysis*, 37(1), 17-45.
- Pearce, D. G. (1988). Tourist time-budget. *Annals of Tourism Research*, 15(1), 106-121.
- Phillip Island Nature Park (2005) How to get to Phillip Island. Melbourne, [www.Penguins.org.au](http://www.Penguins.org.au).
- Stewart, W. J. (1994). *Introduction to the numerical solution of Markov chains*. New Jersey, Princeton University Press.
- The Mathworks Inc. (2008). MATLAB. R2008a ed.
- Truong, T. P. & D. A. Hensher (1985) Measurement of travel time values and opportunity cost from a discrete-choice model, *The Economic Journal*, 95(378), 438-451.
- Xia, J., Arrowsmith, C., Jackson, M. & Cartwright, W.(2008). The wayfinding process relationships between decision-making and landmark utility. *Tourism Management* 29(3): 445-457.

Xia, J., Evans, F. H., Spilsbury, K., Ciesielski, V., Arrowsmith, C. & Wright, G. (2010) Market segments based on the dominant movement patterns of tourists. *Tourism Management*, 31(4), 464-469.

Xia, J., Zeepongsekul, P. & Arrowsmith, C. (2009.) Modelling spatio-temporal movement of tourists using finite Markov chains. *Mathematics and Computers in Simulation*, 79(5), 1544-1553.

Yamamoto, T., Kitamura, R. & Kishizawa, K. (2001). Sampling alternatives from colossal choice set: Application of Markov Chain Monte Carlo Algorithm. *The 80th Annual Meeting of the Transportation Research Board*. Washington D.C., Transportation Research Board.

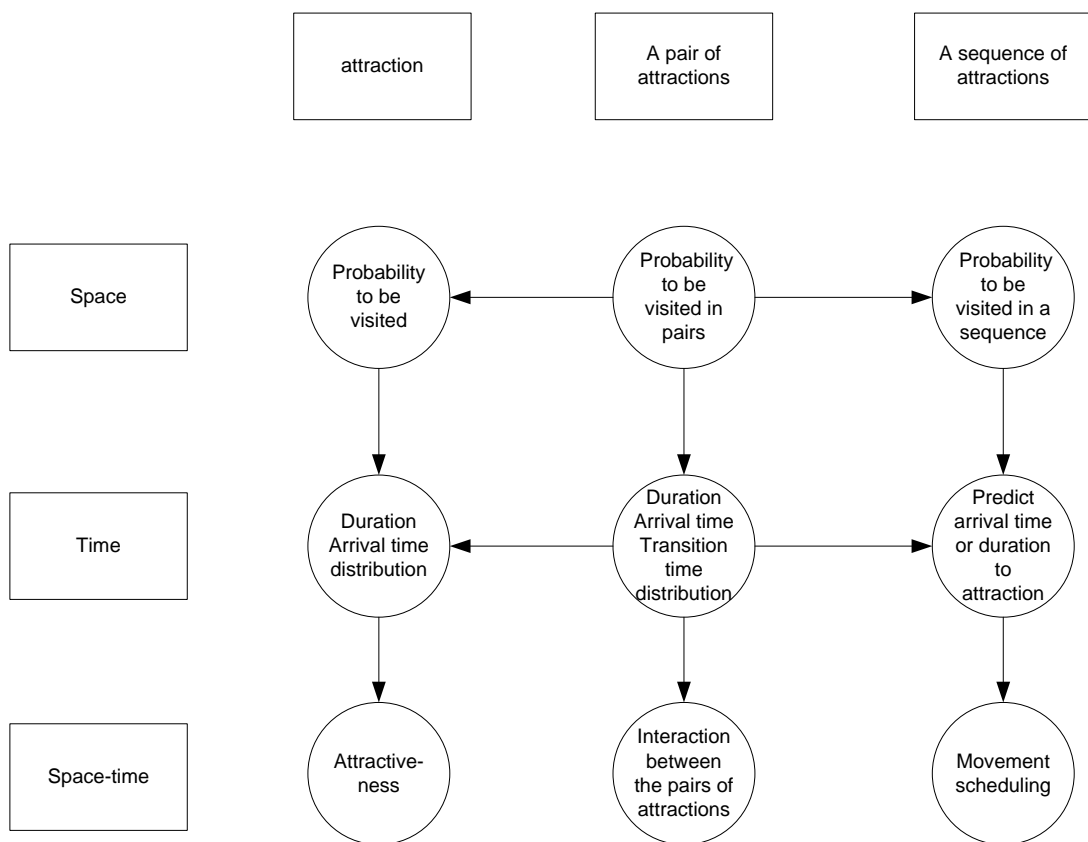


Figure 1 Markov Tourist Movement Application Matrix

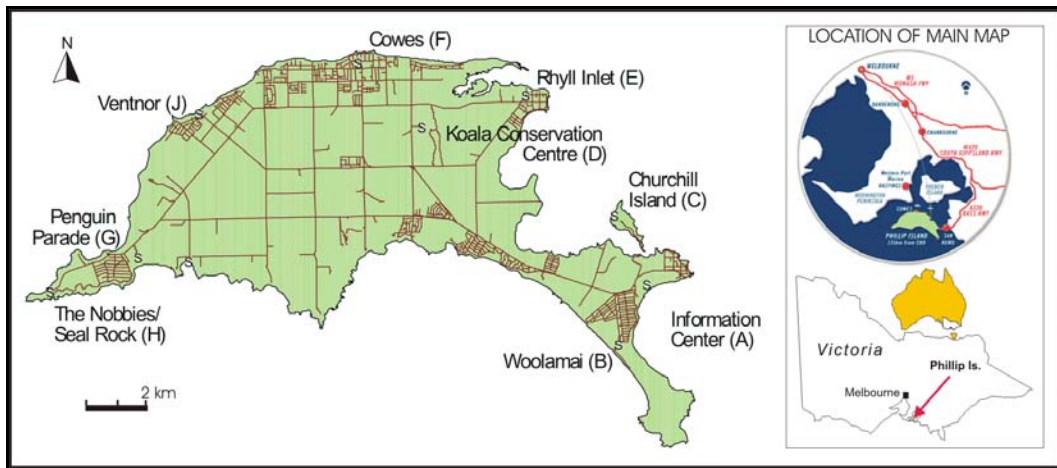


Figure 2 Map of Phillip Island (Phillip Island Nature Park 2005)

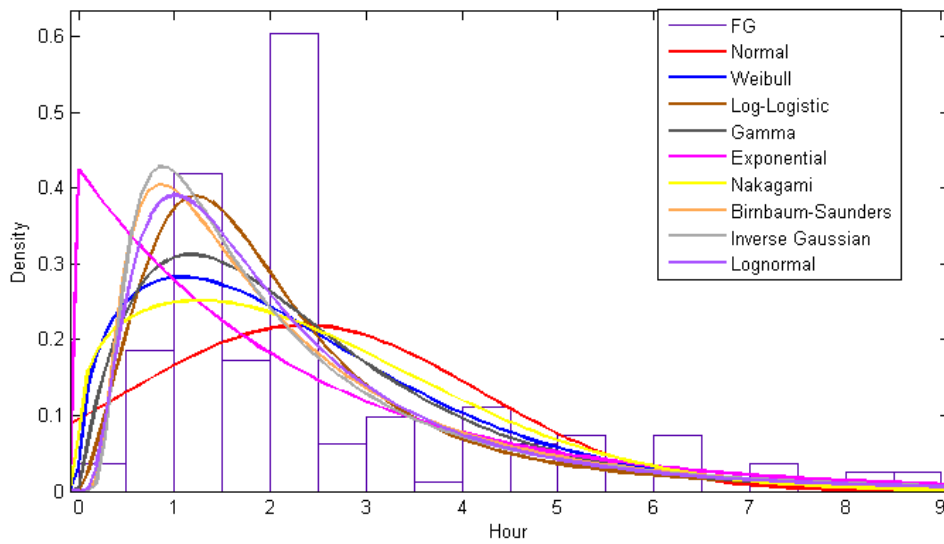


Figure 3 Probability Density Functions (PDF) for Time Spent at Attraction F before Moving to G



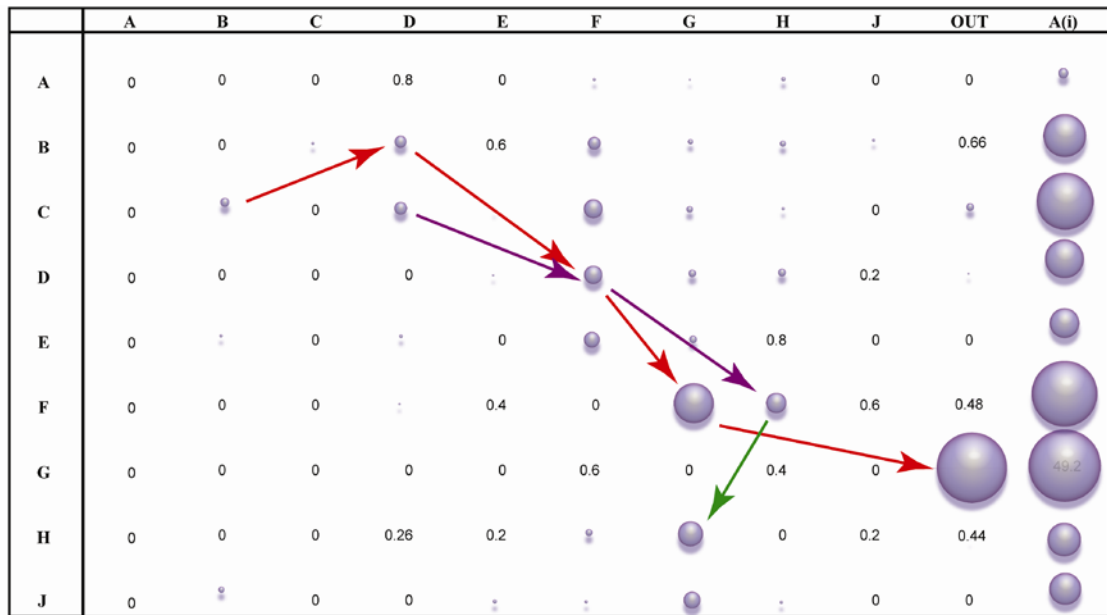


Figure 4 Attractiveness of Attractions

Table 1 Transition Probability Matrix

	Information Centre(A)	Cape Woolamai (B)	Churchill Island (C)	Koala Conservation Centre(D)	Rhyll Inlet(E)	Cowes(F)	Penguin Parade (G)	The Nobbies (H)	Ventnor (J)	Outside the Park (OUT)
Information Centre(A)	0.000	0.000	0.000	0.200	0.000	0.200	0.200	0.400	0.000	0.000
Cape Woolamai (B)	0.000	0.000	0.043	0.217	0.043	0.283	0.087	0.239	0.043	0.043
Churchill Island (C)	0.000	0.103	0.000	0.333	0.026	0.308	0.103	0.026	0.000	0.103
Koala Conservation Centre(D)	0.000	0.022	0.030	0.000	0.030	0.481	0.193	0.185	0.007	0.052
Rhyll Inlet(E)	0.000	0.077	0.000	0.115	0.000	0.577	0.192	0.038	0.000	0.000
Cowes(F)	0.000	0.017	0.010	0.031	0.010	0.000	0.568	0.339	0.010	0.014
Penguin Parade (G)	0.000	0.002	0.002	0.000	0.002	0.022	0.000	0.026	0.000	0.946
The Nobbies (H)	0.000	0.005	0.005	0.019	0.010	0.191	0.746	0.000	0.010	0.014
Ventnor (J)	0.000	0.133	0.000	0.000	0.133	0.200	0.400	0.133	0.000	0.000

Table 2 AIC for Each Distribution Function

FG	Lognormal	Log-logistic	Birnbaum-Saunders	Inverse Gaussian	Gamma	Weibull	Nakagami	Exponential	Normal	Lognormal
Log likelihood	<b>-279.5</b>	-280.331	-280.4	-281.8	-282	-286.1	-291.8	-300.5	-327	-279.5
Number of Parameters	<b>2</b>	2	2	2	2	2	2	1	2	2
AIC	<b>563</b>	564.7	564.8	567.6	568	576.1	587.7	603	657	563

Table 3 Distribution Fit

	A	B	C	D	E	F	G	H	J	OUT
A	0	0	0	?	0	?	?	?	0	0
B	0	0	Lognormal	Birnbaum-Saunders	Gamma	Inverse Gaussian	Gamma	Weibull	Gamma	Gamma Inverse Gaussian
C	0	?	0	Nakagami	?	Logistic	?	?	0	Gaussian
D	0	?	Inverse Gaussian	0	Inverse Gaussian	Lognormal	Lognormal	Weibull	?	Log-Logistic
E	0	?	0	?	0	Inverse Gaussian	Birnbaum-Saunders	Weibull	0	0
F	0	?	?	Weibull	?	0	Lognormal	Logistic	?	Weibull
G	0	?	?	0	?	Nakagami	0	Nakagami	0	Gamma
H	0	?	?	Weibull	?	Nakagami	Log-Logistic	0	?	?
J	0	Birnbaum-Saunders	0	0	?	?	Weibull	?	0	0

Table 4 Mean Time Transition Matrix

	Information Centre(A)	Cape Woolamai (B)	Churchill Island (C)	Koala Conservation Centre(D)	Rhyll Inlet(E)	Cowes(F)	Penguin Parade (G)	The Nobbies (H)	Ventnor (J)	Outside the Park (OUT)
Information Centre(A)	0.00	0.00	0.00	0.20	0.00	0.50	0.30	0.35	0.00	0.00
Cape Woolamai (B)	0.00	0.00	2.30	1.85	0.75	1.54	2.00	0.84	2.00	0.75
Churchill Island (C)	0.00	2.50	0.00	1.29	2.00	2.07	2.00	4.00	0.00	2.50
Koala Conservation Centre(D)	0.00	1.00	1.30	0.00	2.13	1.29	1.29	1.41	1.00	1.10
Rhyll Inlet(E)	0.00	1.00	0.00	1.00	0.00	0.92	1.29	1.00	0.00	0.00
Cowes(F)	0.00	1.80	2.50	2.00	1.75	0.00	2.39	2.04	2.67	1.77
Penguin Parade (G)	0.00	4.00	1.50	0.00	0.00	1.36	0.00	0.89	0.00	2.53
The Nobbies (H)	0.00	1.50	1.00	0.70	1.50	1.19	1.14	0.00	0.75	1.50
Ventnor (J)	0.00	1.30	0.00	0.00	1.00	0.60	1.42	0.75	0.00	0.00

Table 5 Attractiveness of Attractions

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>J</b>	<b>OUT</b>	<b>A(i)</b>
<b>A</b>	0.00	0.00	0.00	0.04	0.00	0.10	0.06	0.14	0.00	0.00	0.34
<b>B</b>	0.00	0.00	0.10	0.40	0.03	0.43	0.17	0.20	0.09	0.03	1.46
<b>C</b>	0.00	0.30	0.00	0.43	0.05	0.64	0.21	0.10	0.00	0.26	1.94
<b>D</b>	0.00	0.00	0.00	0.00	0.06	0.62	0.25	0.26	0.01	0.06	1.32
<b>E</b>	0.00	0.10	0.00	0.12	0.00	0.53	0.25	0.04	0.00	0.00	1.01
<b>F</b>	0.00	0.00	0.00	0.06	0.02	0.00	1.36	0.69	0.03	0.02	2.24
<b>G</b>	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.02	0.00	2.40	2.46
<b>H</b>	0.00	0.00	0.00	0.01	0.01	0.23	0.85	0.00	0.01	0.02	1.14
<b>J</b>	0.00	0.20	0.00	0.00	0.13	0.12	0.57	0.10	0.00	0.00	1.09