Tourism Management 33 (2012) 672–682

Contents lists available at ScienceDirect

Tourism Management

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

Exploring visitor movement patterns in natural recreational areas

Daniel Orellana*, Arnold K. Bregt [1], Arend Ligtenberg [1], Monica Wachowicz [1]

Wageningen University, Centre for Geo-Information Science and Remote Sensing, P.O. Box 47, 6700 AA Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 14 January 2011
Accepted 21 July 2011

Keywords:
Movement patterns
Tourist movement
Visitor monitoring
Spatial behaviour

ABSTRACT

GPS technology is widely used to produce detailed data on the movement of people. Analysing massive amounts of GPS data, however, can be cumbersome. We present a novel approach to processing such data to aid interpretation and understanding of the aggregated movement of visitors in natural recreational areas. It involves the combined analysis of two kinds of movement patterns: 'Movement Suspension Patterns' (MSPs) and 'Generalized Sequential Patterns' (GSPs). MSPs denote the suspension of movement when walkers stop at a place, and are used to discover places of interest to visitors. GSPs represent the generalized sequence in which the places are visited, regardless of the trajectory followed, and are used to uncover commonalities in the way that people visit the area. Both patterns were analysed in a geographical context to characterise the aggregated flow of people and provide insights into visitors' preferences and their interactions with the environment. We demonstrate the application of the approach in the Dwingelderveld National Park (The Netherlands).

1. Introduction

Monitoring and analysing the flow of visitors in natural recreational areas is key to understanding visitor behaviour, which in turn is needed for effective management that meets both conservation and recreational requirements (Mckercher & Lau, 2008; Muhar, Arnberger, & Brandenburg, 2002 pp. 1–6). To understand these requirements we need detailed information about area usage and the preferences of different target groups (Chiesura, 2004). Analysing the spatial behaviour of visitors by relating different uses and activities to different places and landscape configurations can provide insights into their preferences and purposes (Golicnik & Ward Thompson, 2010). One of the most important aspects of the spatial behaviour of visitors in recreational areas is their movement inside the area (intra-site flow). Monitoring the movement of people during their visits to a recreational area can help to identify which places they visit most or least, how much time they spend in each place and which kind of attractions different target groups prefer. Knowing those preferences, managers can segment the market and offer more diverse and focused options, adapted to the wishes of specific groups of visitors (Holyoak & Carson, 2009). Monitoring and analysing the movement of visitors and area usage can also provide information about potential crowding and conflicts between different groups (Manning & Valliere, 2001; Ostermann, 2009). The movement behaviour of visitors looking for solitude and relaxation may differ from visitors looking for social activities, such as playing and picnicking, and studying this can help us understand how different groups experience crowding. The study of intra-site flow of visitors can also provide information for conservation management. To assess the carrying capacity in sensitive areas, for example, we must know about the spatial and temporal distribution of visitors.

Traditionally, studies on visitors' use of space in recreational areas have been based on data and information collected from interviews, surveys and direct observation. Researchers have used geographic information systems (GIS) to analyse the spatial properties of these data to understand how the spatial behaviour of visitors is related to different places and landscape configurations (Golicnik & Ward Thompson, 2010). GIS has also been used to study how recreational areas are used by different groups to detect and understand processes of appropriation and exclusion (Ostermann, 2009).

To complement these techniques, location-sensing technologies (e.g. GPS, mobile phones, PDA) are providing an inexpensive and unobtrusive way to collect massive datasets on the location in space and time of people in recreational areas (Nielsen & Hovgesen, 2004; van Schaick & van der Spek, 2008; Shoval & Isaacson, 2009; Taczanowska, Muhar, & Brandenburg, 2008; Xia, Arrowsmith, Jackson, & Cartwright, 2008). To make sense of this new source of

* Corresponding author. Permanent address: Charles Darwin Foundation, Pto. Ayora, Galápagos, Ecuador. Tel.: +593 5 2526146x222; fax: +593 5 2527013x103.
E-mail addresses: daniel.orellana@fcdarwin.org.ec, temporalista@gmail.com (D. Orellana), arnold.bregt@wur.nl (A.K. Bregt), arend.ligtenberg@wur.nl (A. Ligtenberg), monica.wachowicz@wur.nl (M. Wachowicz).
[1] Tel.: +31 317482092.

0261-5177/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tourman.2011.07.010

data, researchers are envisaging new methods and techniques for exploring and analysing vast amounts of positioning data to extract patterns that represent the movement of individuals and groups (Laube, 2009). Recent advances in the field suggest that despite the potential diversity of movement behaviour, people usually follow simple and predictable movement patterns (Gonzalez, Hidalgo, & Barabasi, 2008; Song, Qu, Blumm, & Barabasi, 2010). It is accepted that these patterns may provide information that will help to explain the interactions between moving entities and between those entities and the environment (Batty, De Syllas, & Duxbury, 2003; Bierlaire, Antonini, & Weber, 2007; Gudmundsson, Laube, & Wolle, 2007; Hoogendoorn & Bovy, 2005). Taking into account the diversity of movement patterns reported in the literature, some authors have proposed formalisation and classification systems to provide a systematic framework for ongoing research (Dodge, Weibel, & Lautenschütz, 2008; Wood & Galton, 2009).

Spaccapietra et al. (2008) stated that in order to analyse movement data and detect useful patterns, the representation of the movement of an object must go beyond its raw spatiotemporal positions. In their work, the authors proposed a representation called 'semantic trajectories', in which the trajectory of the object is divided into semantic units called 'stops' and 'moves'. Stops are those segments of the trajectories where the object does not move. Among various methods proposed to implement this representation, Alvares et al. (2007) devised a method for detecting stops called IB-SMoT (Intersection-Based Stops and Moves of Trajectories), which is based on an analysis of the intersections of trajectories with user-defined geographical features for a minimal duration. Rinzivillo et al. (2008) proposed a similar approach, in which the stops are those segments of trajectories where a moving entity remains within a distance threshold for a minimum period of time. Palma, Bogorny, Kuijpers, and Alvares (2008) proposed a method called CB-SMoT (Clustering-Based Stops and Moves of Trajectories), which analyses each trajectory and generates stops when the speed value is lower than a given threshold for a minimal amount of time.

More recently, Bogorny, Heuser, and Alvares (2010) suggested a general framework for modelling trajectory patterns during the conceptual design of a database. The authors provided a conceptual description of the framework, an implementation of IB-SMoT and SB-SMoT, and data-mining algorithms to extract three movement patterns (i.e., frequent patterns, sequential patterns and association rules) for semantic trajectories. They also provided examples of how to instantiate the model for different applications by parameterising the spatial and temporal dimensions. Other researchers have proposed methods for analysing aggregated movement data to learn more about the spatial behaviour of visitors. For example, Shoval (2010) proposed using a raster-based representation that divides the area of study into a regular grid of cells, and counting the number of GPS observations in each cell of the grid. Finally, some approaches focus on the aggregation of trajectories to improve the visual exploratory analysis of movement data (Andrienko & Andrienko, 2008; Demšar & Virrantaus, 2010; Scheepens, Willems, van de Wetering, & van Wijk, 2011).

A common feature of these approaches is that the conceptualisation of movement patterns requires a parameterisation of spatial and temporal dimensions, which makes the results highly dependent on the values assigned to those parameters. For example, in order to define a stop, the user must provide values for the minimum time, the minimum speed or the minimum distance to be used to determine whether an individual object has stopped, with the risk of overestimating or underestimating the number of stops. Similarly, to detect sequential patterns, the user must set the intervals for aggregating the temporal data in predefined periods (e.g., morning, afternoon, weekend). In the case of spatially aggregated data in raster-based representations, the size of the cell has a considerable effect on the summary statistics. The parameterisation of these values is not trivial and may be highly sensitive to the inherent GPS inaccuracy and to the spatial and temporal resolution of the observations (Palma et al., 2008). Moreover, the selection of parameters is based on a priori knowledge of the dataset, and therefore may be not suitable for an exploratory approach.

In the present work, we propose a novel approach to explore the properties of the collective movement of visitors in recreational natural areas based on GPS tracking data. We define collective movement to be the aggregated properties of the movement of many people in a defined space and time, not the movement of specific groups of people moving together (i.e., collective movement rather than movement of collectives). Our approach relies on different methods of detecting movement patterns that represent the properties of collective movement. In this contribution, we focus on two kinds of movement patterns – Movement Suspension Patterns (MSPs) and Generalized Sequential Patterns (GSPs) – and demonstrate how they can be used to explore the collective movement of visitors in natural recreational areas.

The next section introduces the proposed approach and details the techniques used for the analysis. Section 3 details how the approach was implemented to analyse the flow of visitors in a national park in the Netherlands. Section 4 presents the results of the analysis and Section 5 discusses the most important findings. In the concluding section we briefly review the proposed approach and identify its current limitations and possible solutions.

2. The proposed approach

We want to represent the flow of visitors in a recreational area, defined as the aggregated movement of people visiting different places in a generalized sequence, regardless of the route followed by each individual (i.e. visitors may follow different routes, but a flow exist if the places are visited in a similar order). To represent this flow, we need to uncover spatial and temporal structures describing the visited places and how they are related in space and time. In other words, this flow is a quantitative and qualitative description of the aggregated spatial behaviour of the visitors. It can be graphically represented on a map by arrows between places (Tobler, 2003) and expanded using a space–time cube representation (Hägerstrand, 1970; Kwan, 2004), which we adapted to represent the sequential order in the Z-axis. This visual representation shows the general structure of the flow at the global level, as well as the local level of movement, the single elements of the flow representing the relations between the places. It aids the analysis of the way in which people use the area and interact with different geographical features.

We propose an exploratory approach to analysing the flow of visitors in natural areas using GPS data. The proposed approach has three aims: a) to determine the main places visited by the people in a recreational area by detecting Movement Suspension Patterns (MSPs); b) to establish the sequence in which each individual visited those places; and c) to detect commonalities in those sequences by extracting Generalized Sequential Patterns (GSPs).

Movement Suspension Patterns (Orellana & Wachowicz, 2011) denote the suspension of movement associated with places where people stop. MSPs are therefore spatial structures and are used to discover the places of interest to visitors. As MSPs are determined by the spatial-statistical properties of the whole dataset, no spatial or temporal thresholds are required. Generalized Sequential Patterns (Agrawal & Srikant, 1995) describe the sequence in which the places are visited, regardless of the trajectory followed. The term 'generalized' implies a relative order and not an absolute order: GSPs are temporal structures used to find commonalities in the order that places are visited. We believe that these two kinds of patterns aid understanding of the spatial behaviour of visitors. The patterns are used to analyse the overall flow of visitors and provide insights into how they interact with relevant places.

The movement of visitors is analysed in consecutive steps, using different techniques (e.g., spatial statistics, data mining and visual exploratory analysis). Each step results in a new dataset representing some specific characteristics of movement. The original data consist of a set of tuples representing the spatiotemporal coordinates of the people, which are recorded using positioning devices (e.g., GPS loggers). First, we compute movement parameters, such as speed and bearing, which generates a dataset of movement vectors representing the properties of movement observed in space and time. Second, we apply a spatial-statistical method to detect MSPs in the movement vectors dataset, which generates a set of spatial clusters representing the locations and times when visitors stopped. These data are used to find the points of interest to the visitors. Finally, we use a data-mining algorithm to extract GSPs that indicate the relative temporal sequence in which the places are visited. The result is a directed graph representing the frequent generalized sequences for those places. These patterns thus represent the aggregated flow of visitors in the study area. Fig. 1 shows the different steps of the process.

2.1. Acquiring positioning data

The movement of visitors can be recorded using Global Navigation Satellite Technology (GNSS), such as GPS receivers and GPS-enabled smartphones. These devices are capable of capturing the location of users in space and time with a certain periodicity. The spatiotemporal positions of a large number of devices tracked in an area during a period of time provide a good basis for analysing collective movement in that area. Preconfigured GPS devices can be handed to the visitors at the entrances of the area. The captured data is stored in one of the standard formats, such as NMEA sentences (National Marine Electronics Association, 2010) or GPX (Topografix, 2010), and can be transmitted in real time to a server via a GPRS or 3G signal (for GPS-enabled smartphones) or imported later (for GPS receivers).

2.2. Computing movement vectors

Movement vectors represent individual observations of movement that can be measured or sensed at a particular place at a particular time. They are defined by the spatiotemporal coordinates of the observation coupled with a magnitude (e.g., speed, acceleration) and an angle (e.g., direction of movement), and can be represented graphically by an arrow. A vector space representation is a set of movement vectors of one or more entities moving in a defined spatiotemporal area. Movement is thus conceived as a spatial property (i.e., how movement is observed in space) rather than as a property of the trajectory of a particular entity (i.e. how does a specific person move).

The speed and bearing values of movement vectors can be derived in real time, depending on the capabilities of the device used to capture the space–time positions. Since not all the devices have those capabilities, a simple computer procedure can be used after the data is collected. The procedure takes two consecutive GPS observations and derives movement parameters (speed, distance, bearing and time step). If two consecutive observations are too separated in space or time (e.g., because the GPS signal is lost or because the GPS is turned off), this separation can be parameterised in the procedure to avoid errors in the computation of movement vectors. The procedure is publicly available online at http://ideasonmovement.wordpress.com. Movement vectors can be used to explore aggregated properties of movement, such as distribution and density and the global and local statistics of speed, direction and other movement parameters.

2.3. Detecting movement suspension patterns

The local statistics of the set of movement vectors can be used to identify spatial clusters of low speed values. In Orellana and Wachowicz (2011) the authors demonstrated the use of a Local Indicator of Spatial Association (LISA) (Anselin, 1995) to find these clusters and detect Movement Suspension Patterns (MSPs). These patterns may indicate the location of geographical features associated with the reduction in speed that characterises the stopping behaviour of pedestrians. Although MSPs are similar to the concept of 'stops' used in other approaches, they are essentially different.

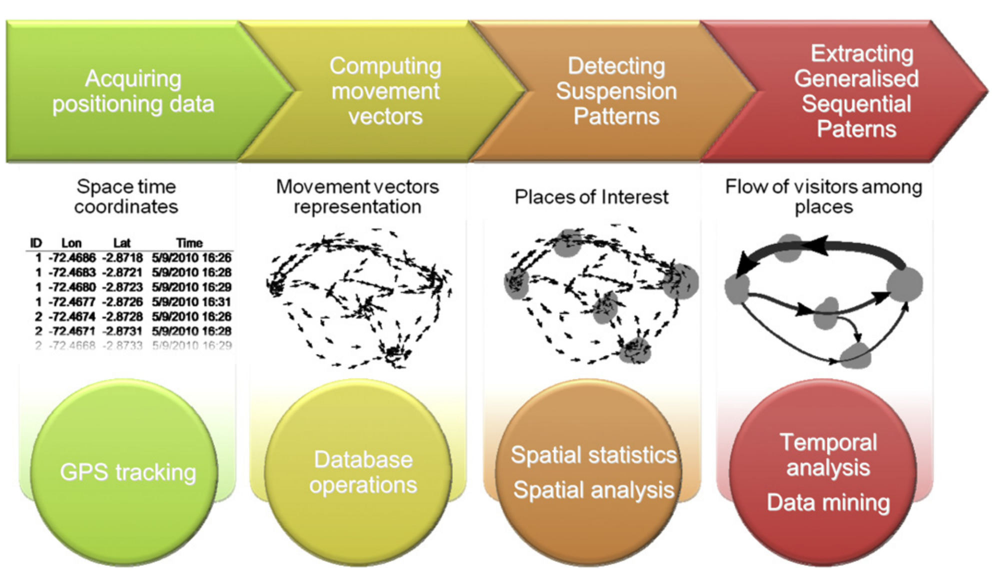

Fig. 1. Schematic description of the proposed approach: First, movement data is captured using GPS devices and imported into a geodatabase. Movement vectors are then computed from observational data using a dedicated database procedure. Next, movement suspension patterns are obtained from spatial statistics and spatial analysis. Finally, generalized sequential patterns are extracted from the temporal sequences of suspension patterns using a data-mining algorithm.

Whereas 'stops' are the parts of an object's trajectory where the object does not move, MSPs are spatial clusters of low speed vectors with a strong spatial association. This difference means that if only one individual stopped for a short time at a place where other individuals continued walking, this may not considered to be a MSP, because the spatial association of the speed values of movement vectors is not strong. This is an advantage for the analysis of collective spatial behaviour. The LISA method was selected because it tends to have the best statistical properties and requires few assumptions about the data. The time and duration of each MSP was obtained from the timestamps of the first and the last movement vectors of each visitor in each spatial cluster.

Fig. 2a shows an example of a vector-based representation of the movement of four visitors and contains four spatial clusters of movement suspension (numbered 1–4). Each cluster consists of the movement vectors of different visitors (numbered from A to D). Fig. 2b depicts the temporal dimension of the MSPs. The vectors classified as suspension are plotted on the timeline of the corresponding spatial cluster, with different markers representing different visitors. The lines above each group of movement vectors represent the duration of each individual MSP.

2.4. Extracting generalized sequential patterns

Generalized Sequential Patterns (GSPs) (Agrawal & Srikant, 1995) are the frequent generalized sequences that can be found in a timely-ordered set of events (in this case, the events are MSPs). They are 'generalized' because the MSPs occurred in a relative order rather than an absolute order. The assumption here is that, given a set of MSPs for a group of people, with their corresponding locations and times, there are structures in the relative order that characterise the collective spatial behaviour of the group.

Using data-mining techniques, GSPs are extracted from the original sequences of MSPs. Each GSP is an ordered list of MSPs (represented by the ID of the corresponding spatial cluster) together with a *support* value. The support value is the ratio between the number of sequences corresponding to the pattern and the total number of sequences in the dataset (Agrawal & Srikant, 1995). The example provided in Fig. 2c shows the sequences of MSPs for each visitor (i.e. the order in which each visitor visited the different places). Fig. 2d enumerates three GSPs with the corresponding *support* values. An interpretation of this example in terms of spatial behaviour is that there are four main places where people stop. In addition, the flow of visitors goes from the signpost (3) to the cafeteria (4) and then to the car park (1), with some visitors deviating via the monument. Moreover, all the people visited the monument (2) before stopping at the car park (1).

Since the potential number of extracted GSPs can be high, the most salient cases are selected during the exploratory analysis, which requires objective and subjective criteria of interestingness. Although *support* is the most frequently used objective measure of interestingness, there is some criticism that it does not provide the flexibility an exploratory approach needs, because it selects the most common cases and dismisses the uncommon, and probably interesting, cases (Laube, 2009). Depending on the application, potentially useful measures are complementary subjective criteria such as *novelty* and *actionability* (Han & Kamber, 2006). These subjective criteria rely not only on the input data, but also on the user examination of the pattern. In general, a pattern can be interesting if it is 'surprising' to users (novelty) or if they can 'do something' with the pattern (actionability) (Silberschatz & Tuzhilin, 1995). For example, a pattern can be considered interesting even when it has a low support value if it reveals a new, unexpected flow of visitors between places that were not initially considered. Similarly, a pattern may be interesting to a park manager if its discovery can be used to improve management practices. In our approach, we combined both objective (i.e. support) and subjective (i.e. novelty) criteria during the exploratory analysis of the flow of visitors.

3. Implementation

The proposed approach was used to explore the spatial behaviour of visitors in the Dwingelderveld National Park in the Netherlands. We analysed the positioning data recorded by GPS devices carried by 372 visitors during their visit to the park. We selected three research questions to illustrate the use of the approach: a) *What are the main visited places in the park?* b) *What are the visitor flows from the entrances to the main places?* and c) *What are the visitor flows between the main places?* The answers to these questions will provide a quantitative and qualitative description of the aggregated flow of visitors in the park.

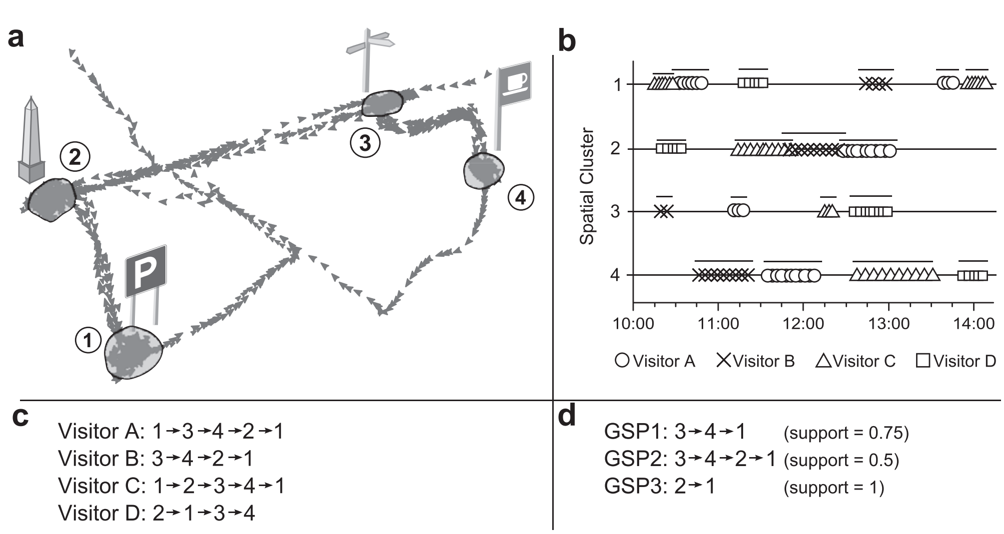

Fig. 2. Examples of different representations of the movement of four visitors. a) Vector-based representation with spatial clusters of movement suspension located at four points of interest. b) Temporal duration of movement suspension patterns for each cluster. c) Sequence of suspension patterns for each visitor. d) Three examples of generalized sequential patterns and their corresponding support values.

Starting with a movement vector representation, the LISA approach was used to detect Movement Suspension Patterns (MSPs). The MSPs were then compared with a geographical dataset, allowing us to answer question a). In the next step, Generalized Sequential Patterns (GSPs) were computed using the BIDE+ data-mining algorithm (Wang, Han, & Li, 2007) and compared for each of the five entrances in the park to answer question b). Finally, the GSPs between the main places were analysed to answer question c).

3.1. Study area

The Dwingelderveld National Park (DNP) is an area of about 3700 ha in the north-east of the Netherlands. It is a typical Dutch recreational area with an extensive network of short strolls (60 km of marked trails, each less than 7 km in length) and long walks, as well as routes for cycling and horse riding. The landscape consists mainly of dry and wet heathlands, pine and deciduous forest, and an important complex of juniper shrubs. Dwingelderveld is a very popular area and receives between 1.5 and 2 million visitors each year. Besides the wetlands, sheep farms and some bird-watching hides, which are the main tourist attractions, the park contains additional amenities for visitors, such as staffed and unstaffed information centres, a tea house and some cultural attractions, including a historic house and a radio telescope (van Marwijk, 2009). Visitors enter and leave the park through one of the five access points (where car parks are located) and follow the paths to one or more points of interest or pursue various leisure activities.

Three different datasets were used. The first was a positioning dataset recorded by GPS receivers given to visitors at the entrances of the park (the beginning of the GPS track). Of the 461 visitors asked to participate, 400 agreed to carry a GPS device during their visit. An evaluation of the quality and completeness of the data led to the inclusion of 372 GPS tracks in the final dataset, which contained about 142,000 time-stamped geographical coordinates. This data was collected over a seven-day period (weekend and weekdays) in the spring and summer of 2006 (details of the data collection can be found in van Marwijk, 2009). The second dataset was a map showing the path network and the locations of the park entrances. The third dataset was a collection of 271 points representing the locations of the attractions and facilities in the park, gathered from specialised web pages (Natuurmonumenten, 2009; Pol-Recreatie, 2003) and a field survey. Fig. 3 is a map of the study area.

3.2. Finding the main visited places

The first task was to identify the main places of interest in the park. We define 'main place' as a site where a movement suspension pattern is detected which can be associated with a relevant geographical feature that can explain the suspension of movement. The assumption here is that visitors are attracted to these places and temporally suspend their movement to perform some activity associated with the place. Some examples are visiting an interesting spot, reading an information board, eating or resting at a picnic bench, etc. These attractions affect the collective movement by shaping the flow of visitors.

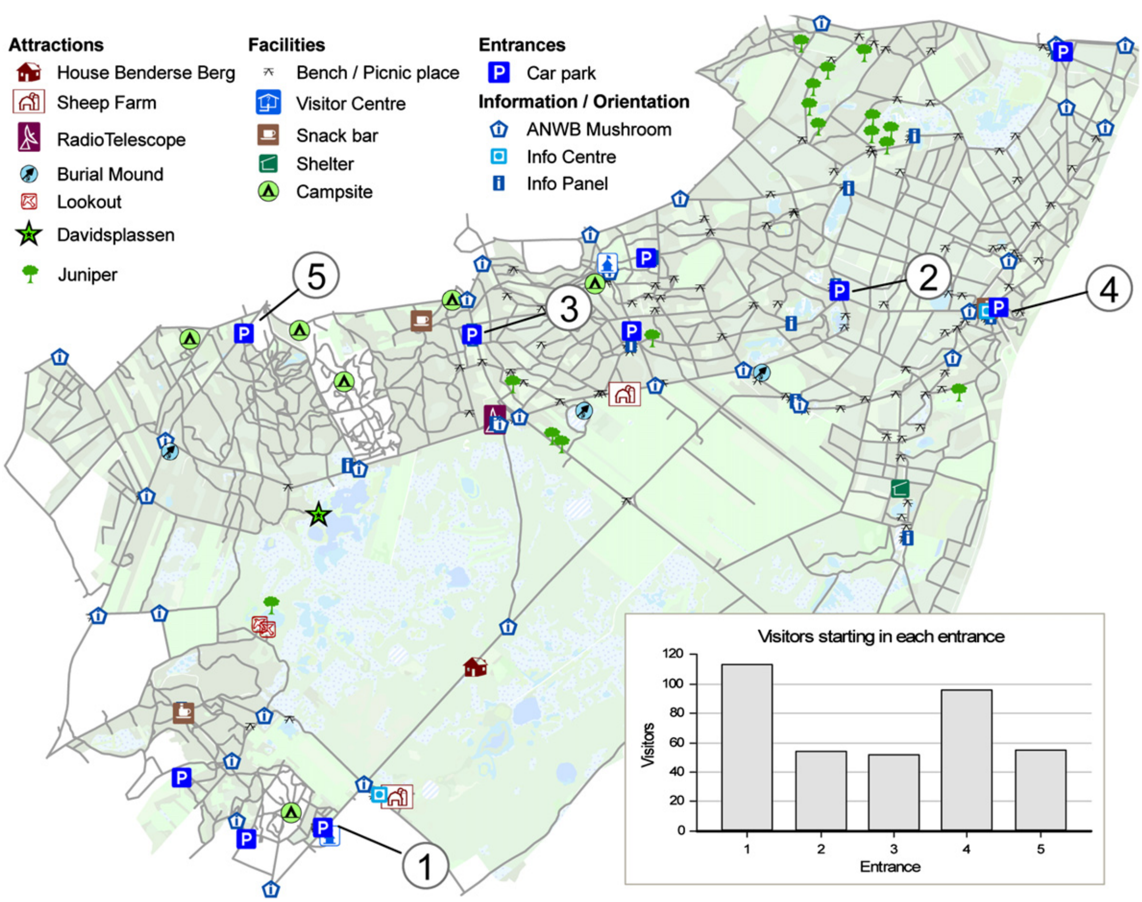

Fig. 3. Map of the Dwingelderveld National Park showing the five entrances and the locations of the main attractions and facilities.

We computed a set of movement vectors representing the movement of visitors using data from the positioning GPS dataset and stored them in a geodatabase. The Local Moran's index was computed for the movement vectors using an adaptive neighbourhood defined by a radius of 50 m and a minimum of 15 observations. These parameters were defined in an exploratory analysis of the spatial distribution of the dataset. The movement vectors with speeds less than the mean and *Z*-scores above 1.96 (5% significance level) were classified as movement suspension and they showed spatial clusters when plotted on a map. To define the boundaries of the spatial clusters, we used a kernel-density function on the set of vectors classified as movement suspension to obtain a continuous surface representing a density estimate. Percent Volume Contours (PVCs) traced on this surface represent the boundaries of the areas that contains *x*% of the volume of the probability density distribution. For example, using a value of 99, the PVCs delineate the areas containing on average 99% of the vectors that were used to generate the kernel-density estimate (Beyer, 2010). The bandwidth for the kernel-density function was determined by the estimated accuracy of the observation (e.g., 10 m for a GPS receiver under ideal conditions). These lines represent the boundaries of compact spatial clusters of suspension of movement. Moreover, the kernel-density function allowed us to find hotspots of movement suspension and differentiate outliers (e.g. clusters with only one movement vector).

GIS overlaying was used to compare the spatial clusters with a point dataset containing the attractions and facilities in the park. An attraction or facility was associated with a spatial cluster if its location lies inside the 99% PVC. However, some attractions were landscape elements outside a cluster but visible from within the cluster (e.g., a water body). In these cases, maps, aerial imagery interpretation and field verification were used to associate the cluster with the corresponding attraction. The relative importance of each place was evaluated using aggregated statistics for the cluster, such as the number of visitors and the number of MSPs. These results were used to answer the first question, *What are the main visited places in the park?*

3.3. Exploring the aggregated flow of visitors

The second and third questions concern the aggregated flow of visitors. This flow is understood to be the aggregated collective movement of visitors between the places and is represented by Generalized Sequential Patterns (GSPs).

We implemented a database procedure in which the MSP dataset obtained in the previous step was analysed to determine the temporal sequences for each spatial cluster and visitor. The database was updated by adding to each MSP an integer indicating the position in the sequence and an identifier for the individual sequence. The result was a dataset of individual sequences of ordered MSPs. This dataset was analysed using the BIDE+ algorithm implemented in the Sequential Pattern Mining Framework, a publicly available JAVA code for analysing sequential patterns (Fournier-Viger, Nkambou, & Nguifo, 2008). We selected this algorithm because it avoids redundancy in the results by extracting only 'closed' sequential patterns. 'Closed patterns' are sequences that are not contained in another sequence having the same support. A closed pattern induces an equivalence class of patterns sharing the same closure, and those patterns are partially ordered, e.g. according to the inclusion relation. The smallest elements in the equivalence class are called minimal generators, and the unique maximal element is called the closed pattern (Fournier-Viger et al., 2008). We modified the original code of the algorithm to produce a formatted output consisting of two tables that can be imported back into the geodatabase. The first table contained the structure of the GSP, in the form of an ordered list of spatial clusters where the MSPs occurred. The second table contained the properties of each GSP, consisting of the ID of the GSP and the values for support, frequency and size.

The results were represented in a dynamic map linked to the geodatabase to indicate the flow of visitors between the main places. This linkage enabled a dynamic visualisation by querying and filtering the datasets.

The second question, *What are the visitor flows from the entrances to the main places?* is answered by exploring the GSPs starting from each entrance in the park. We analysed the GSPs with the largest support values, representing the aggregated flow of visitors from the entrance to the main places. The results were shown on a map using arrows to indicate the direction of the aggregated flow of visitors to each place.

We took a similar approach to answer the last question, *What are the visitor flows between the main places?* The exploratory analysis was performed in a 3D sequential space–time cube, the base of the cube being a two-dimensional map of the park and the *Z*-axis representing the sequential time. The GSPs were rendered in the cube as three-dimensional polylines connecting the different places, ordered in time in the vertical axis, the first MSP at the bottom. The thickness and colour of the lines represent the support values of the GSPs and provided a visual cue for the analysis. To aid visual interpretation, vertical lines connect the places on the map to the MSPs. Moreover, each GSP was linked to a data-space representation, such as multi-dimensional scatter plots. This is a powerful exploratory tool and allowed us to interact with the data by querying, filtering and highlighting the GSP dataset in a multi-dimensional space. The resulting visualisations allowed a quick and effective interpretation of the GSPs and helped to uncover the structures in the flow of the visitors. For example, GSPs with relatively high support values implied that many people visited the places in that order, shaping a visible flow of visitors in the space–time cube.

4. Results

4.1. Most visited places in the park

Using the LISA index to classify the movement vectors, we found that 6.3% ($n = 8988$) corresponded to movement suspension. We identified 184 spatial clusters defined by the 99% PVCs drawn on the kernel-density surface. These clusters contained in total 1581 MSPs. By applying a spatial overlay function, we found that 158 spatial clusters (85.9%) could be associated to at least one relevant geographical feature in the park. These clusters contained 1546 MSPs, or 97.8% of all the MSPs. This result allowed us to discover the places visited by the people participating in the data collection (Fig. 4).

Assuming the degree of interest of a place can be related to the number of times it was visited (represented by the number of MSPs in the corresponding cluster), we analysed a subset containing the 10% ($n = 18$) largest clusters to provide an indication of the most interesting places for the visitors. Details of these clusters are reported in Table 1 and Fig. 4.

The car parks at the five entrances were among the most visited places. In fact, all the visitors started and finished their visits there (e.g., Fig. 4b,f). The remaining main visited places included attractions and amenities in the park. The most visited attraction in the park was the visitor centre close to Entrance 1 (70 visitors, Fig. 4f), where the duration of the visits were typically long (more than 15 min). Other main visited attractions were the radio telescope (41 visitors, Fig. 4a), the sheep farm (32 visitors, Fig. 4e), and some of the wetlands that characterise the landscape of the park and

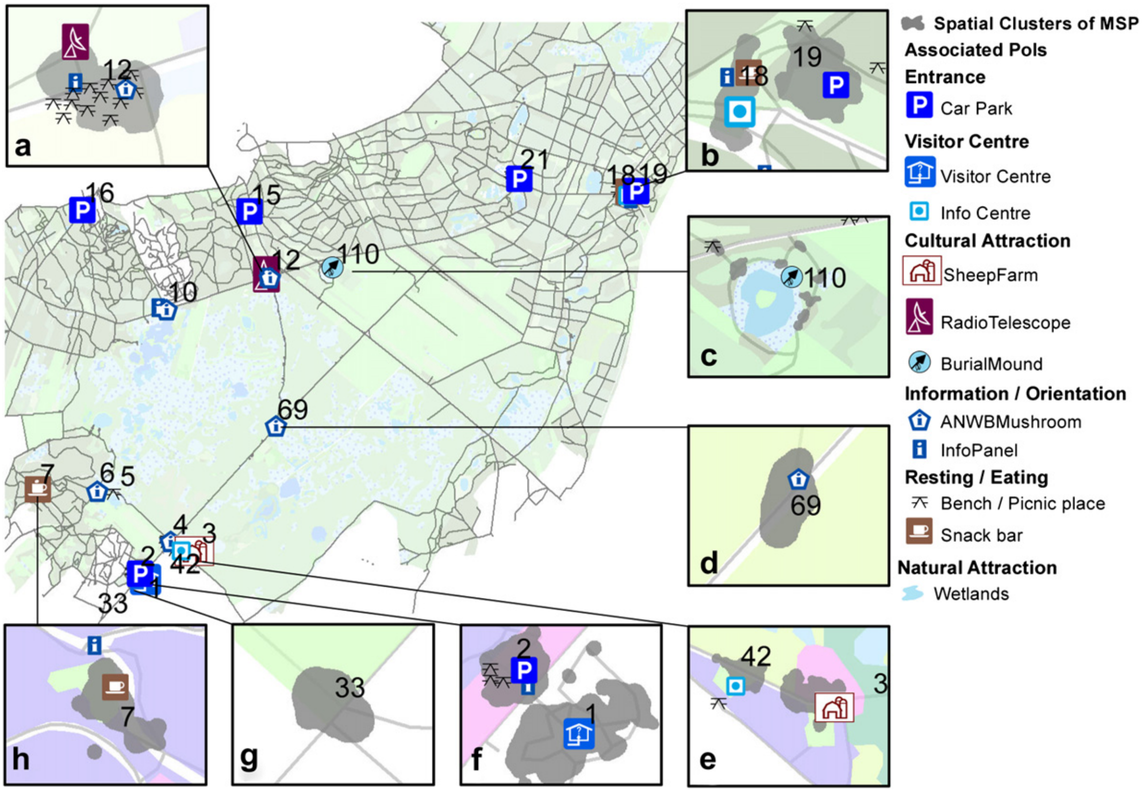

Fig. 4. Map showing the spatial clusters of Movement Suspension Patterns and associated geographical features. The numbers are the IDs of the spatial clusters. Detailed examples are depicted in the insets.

where picnic benches are located. Two of the wetlands that received a large number of visitors are Davidsplassen (31 visitors) and Smitsveen, also the site of an ancient burial mound (17 visitors, Fig. 4c). The average duration of the visits was longer at the radio telescope and the sheep farm (about 7 min) than at the wetlands and other attractions (less than 5 min).

The most visited amenities were those providing information and orientation, including the information point near the entrance at Spier (37 visitors, Fig. 4b) and the one close to the sheep farm (31 visitors, Fig. 4e). The other main visited places providing information were some of the mushroom-shaped ANWB signposts (e.g., Fig. 4d), where the MSPs were of a short duration. Another frequently visited facility was the tea house in the forest (Fig. 4h), with the longest average duration of visits (27 min 50 s).

It is interesting that, besides the attractions and amenities, one spatial cluster of MSPs is located at a path crossing (Fig. 4g), where 27 visitors stopped. This pattern may indicate a specific spatial behaviour of visitors arriving at a crossing and choosing which direction to take. The short average duration of the MSP at this cluster (19 s) supports this interpretation.

Table 1
The 10% largest clusters and their related geographical features. The figures indicate the number of different visitors stopping at the place, the number of Movement Suspension Patterns, the number of vectors classified as suspension, and the average duration of each MSP.

Cluster (Id)	Associated Feature(s)	Visitors (n)	MSP (n)	Vectors (n)	Avg. Duration (min:sec)
2	Car Park	107	207	1212	03:01
1	Visitor Centre	70	82	2011	16:25
19	Car Park	60	92	201	07:19
4	Picnic/ANWB Mushroom	59	74	210	02:01
12	Radio Telescope	41	47	465	07:27
16	Car Park	37	70	257	03:58
18	Snack Bar/Info Spier	37	46	84	01:32
15	Car Park	33	57	198	05:35
3	Sheep Farm	32	32	771	07:34
42	Information at Sheep Farm	31	47	321	04:37
10	Davidsplassen/Picnic	31	31	188	02:13
33	Path Crossing	27	28	60	00:19
21	Car Park	25	45	152	07:38
5	Wetland/Picnic	25	25	132	04:09
7	Tea House	24	24	126	27:50
6	ANWB Mushroom	20	20	35	00:37
69	ANWB Mushroom	19	19	102	01:05
110	Smitsveen/Burial Mound	17	20	96	03:08

4.2. Flow from the entrances to main places

The temporal analysis of the set of MSPs resulted in 282 sequences of MSPs, representing the relative order in which people visited each place. Since a sequence is an ordered set of MSPs, the number of sequences equals the number of people who visited at least two different places (i.e. at least two MSPs detected at different spatial clusters). The BIDE+ algorithm extracted 218 Generalized Sequential Patterns (GSP) with a minimum support of 0.02. This value was low enough to capture uncommon GSPs to allow further filtering and select cases with higher support values. We used a query to extract GSPs in which the first element corresponded to an MSP located at an entrance and found 16 GSPs representing the flow of visitors from the entrances to the main places. This result is reported in Table 2 and Fig. 5.

The shape of the flow of visitors from the entrances to the main attractions reflects the relative importance of Entrance 1 (Fig. 5), where most of the people in the study started their visit to the park. Many of the main attractions are easily reachable from this

Table 2
GSPs from the entrances to the main places.

From (Entrance ID)	To (Cluster ID)	Geographical feature	Support	Frequency	Distance (m)
1	110	Wetland/Burial Mound	0.02	7	4099
1	10	Davidsplassen	0.03	9	3025
1	69	ANWB Mushroom	0.03	10	2283
1	12	Radio Telescope	0.04	15	3683
1	6	ANWB Mushroom	0.04	13	1057
1	7	Tea House	0.04	14	1483
1	33	Path Crossing	0.06	21	188
1	5	Wetland/Picnic	0.06	19	940
1	42	Information at Sheep Farm	0.07	25	576
1	3	Sheep Farm	0.08	27	741
1	4	Picnic/ANWB Mushroom	0.13	46	521
1	1	Visitor Centre	0.16	55	102
3	110	Smitsveen/Burial mound	0.02	6	1184
3	12	Radio Telescope	0.05	16	777
5	10	Davidsplassen	0.04	13	1483
4	18	Snack Bar/Info centre	0.06	20	75

entrance, which may also explain the large proportion of GSPs starting there. This interpretation is supported by the fact that the largest support values corresponded to the GSPs with the shortest lengths (measured as the Euclidean distance between the origin and destination of the flow). This implies that people usually choose an entrance near to their preferred places as the starting point of their visit to the park. For example, the flow to the tea house, the sheep farm and the visitor centre came from Entrance 1. Likewise, the GSP representing the flow of visitors from Entrance 5 to the wetlands in Davidsplassen had larger support than the one from the more distant Entrance 1. We found three exceptions: the radio telescope and the wetlands at Davidsplassen and Smitsveen received visitors from two entrances. Another interesting finding is the flow from Entrance 1 to cluster 4 (Fig. 5, lower inset). This cluster is located at a path crossing, which also has a signpost and a picnic bench. Visitors seemed to stop there before deciding which route to follow to continue the visit.

4.3. Flow between the main places

The first visualisation created to explore the flow in the space–time cube was a display of all GSPs (Fig. 6). This visualisation revealed that GSPs with high support values (*support* ≥ 0.1) usually consisted of only two or three MSPs. They all started at a car park, went to a nearby place and then came back to the starting point. The more MSPs the GSP had, the lower the support value was. This is an expected result because shorter GSPs represent relatively simple flows that many visitors may follow (e.g. Entrance 1 → Visitor Centre → Entrance 1). But GSPs are nested structures, and so shorter GSPs are parts of longer GSPs, which also explains the differences in support values. At this point, it is important to remember that GSPs imply a generalized order, and the example given captures the flow of all the visitors visiting the places in that order, regardless of whether they visited other places in between or not.

The capabilities for dynamic interaction with the data allowed us to explore interesting structures and perform comparisons. For example, we found an interesting example of a GSP with many MSPs and relative high support values (highlighted in dark red in Fig. 6). This example represents the flow of people who visited the places in the following sequence: Entrance 1 → Path crossing → Info Centre → Sheep farm → Info Centre → Entrance 1 ($s = 0.05$).

When we filtered the GSPs for specific places, some properties of the flow at those places were revealed. For example, from the GSPs associated with the radio telescope shown in Fig. 7, it is possible to see that the tea house and Davidsplassen were typically visited after the radio telescope (flow lines go upwards from the radio telescope to those places), whereas the historic house and some of

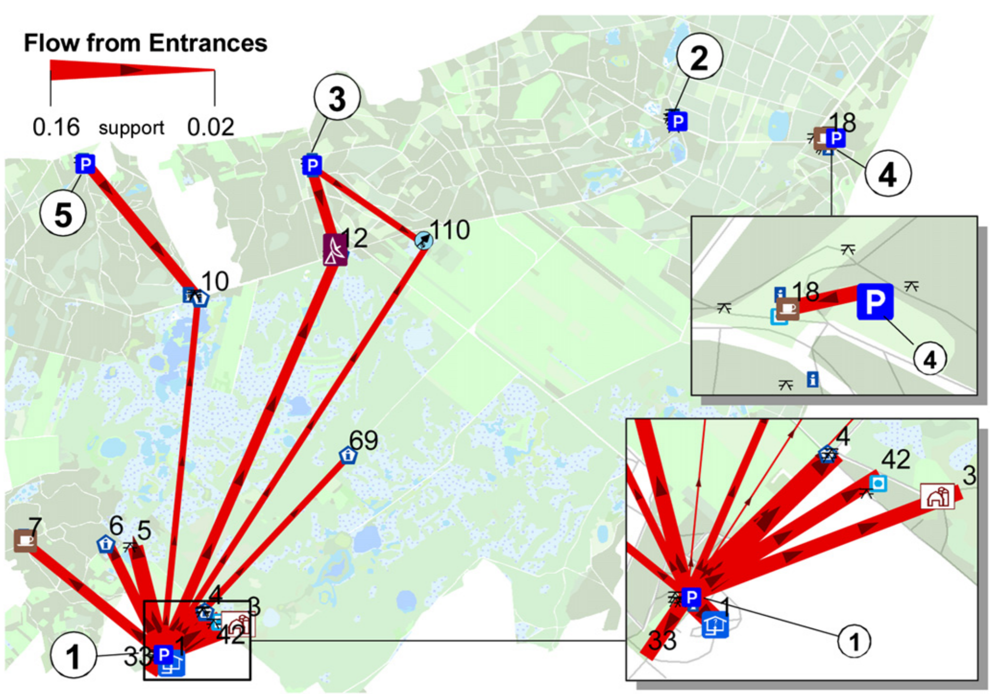

Fig. 5. Flow from each entrance to the main visited places. The numbers in circles are the identifiers of the entrances and the plain numbers are the identifiers of the spatial clusters associated with each place. The thickness of the lines representing the flow is proportional to the corresponding support value.

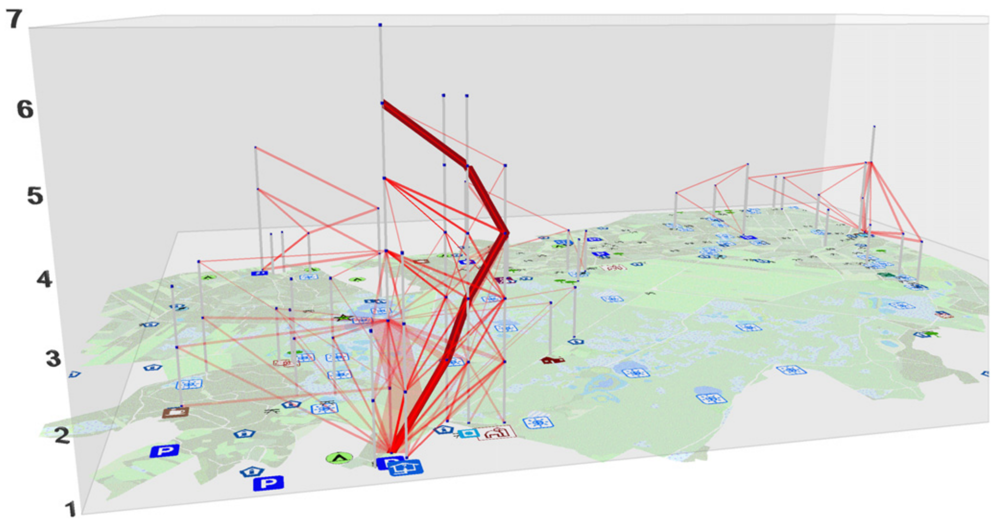

Fig. 6. Sequential space–time cube representing Generalized Sequential Patterns with support ≥ 0.03. The vertical axis indicates the sequence in which the places in each GSP were visited. Dark thick lines indicate high support values; light thin lines indicate small support values. An interesting pattern is highlighted in black.

the wetlands were visited before the radio telescope (flow lines go upwards from those places to the radio telescope). An example of a GSP with more MSPs is Entrance 1 → Burial mound → Radio Telescope → Entrance 1. The support for this GSP was 0.02 (highlighted in dark red in Fig. 7).

Using the filtering and linking functions, we found other interesting properties of the spatial behaviour of visitors. For example, the visitor centre was always visited before the sheep farm, but the nearest information centre was visited both after and before it. Also, of all the visitors going from Entrance 5 to Davidsplassen, 78% had first visited one of the wetlands near the entrance. Finally, of all the visitors that went from Entrance 1 to the tea house, half went first to the visitor centre. There was no flow of visitors from the tea house to the visitor centre.

In general, we found that GSPs with low support values were more common than those with high values. For example, almost 80% of the GSPs had a support value equal to or less than 0.04% (Fig. 8a). This means that there were many different sequences followed by few visitors and only few sequences followed by many visitors (Fig. 8a). The visitor flow is therefore made up of many different patterns. We also found that short GSPs (with three or four MSPs) were more common than GSPs with more MSPs (Fig. 8b), indicating that the flow of visitors consists mainly of common, simple sequences.

5. Discussion

The results obtained to answer the first question, *What are the main visited places in the park?* indicate that despite the large number of points of interest in Dwingelderveld National Park, only a limited number of them attracted significant numbers of visitors and many were hardly visited at all. The commonly visited places can be classified according to their functions as *attractions* and *amenities*. *Attractions* are places mainly associated with natural and cultural leisure activities, where people go to experience and enjoy the landscape and features in the park. Examples of these places are

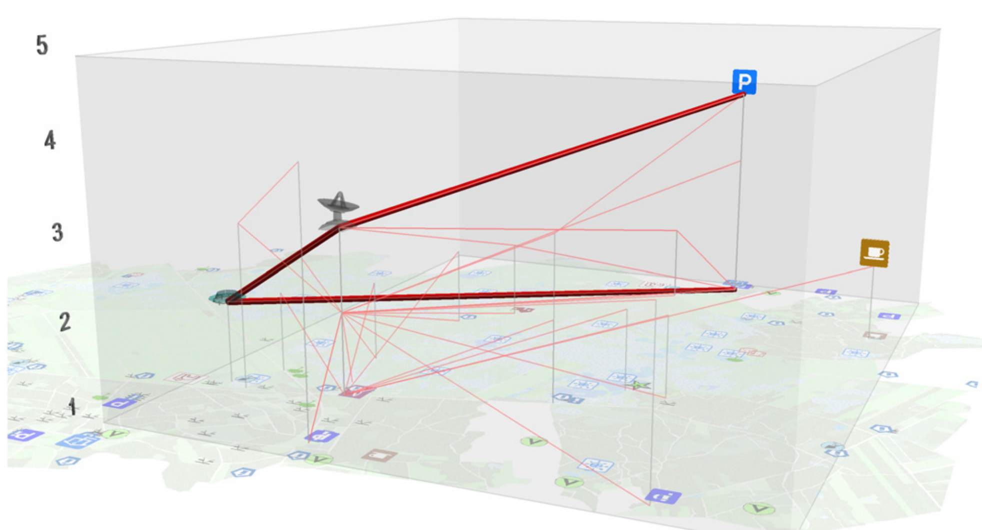

Fig. 7. An example of the sequential space–time cube for the flow of visitors to the radio telescope. The vertical axis represents the temporal sequence, the thin diagonal lines represent the flow of visitors between the places, and the vertical lines link the clusters to the places on the base map. The highlighted Generalized Sequential Pattern represents the flow of visitors starting at Entrance 1, visiting the burial mound and then the radio telescope, before going back to the initial point (support = 0.02).

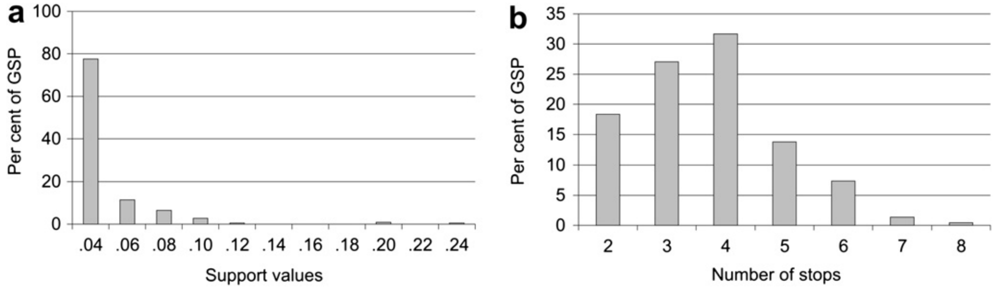

Fig. 8. a) The large percentage of Generalized Sequential Patterns with low support values may indicate a large diversity in the flow of visitors (many sequences followed by few visitors and few sequences followed by many visitors). b) 77% of the GSPs had four Movement Suspension Patterns or less.

the radio telescope, the sheep farm and the wetlands in Davidsplassen and Smitsveen. *Amenities* are places with information facilities and services for visitors. Examples are the tea house, the information centres, the information boards and poles, and the picnic benches. Interestingly, besides these two kinds of places, we found that visitors also stopped at some path crossings. Our interpretation is that these are places where visitors temporary suspended their movement to decide which path to take.

The results obtained to answer the second question, *What are the visitor flows from the entrances to the main places?* show that the flow of visitors to the main places came mainly from the nearest entrance. Few places received a flow of visitors from two entrances. This indicates that the entrance selected as the starting point largely determines the places that are visited, suggesting that the flow of visitors may follow a gravity model.

The results obtained to answer the third question, *What are the visitor flows between the main places?* indicate a great diversity in the flow of visitors in the park. Although few places were visited often, they were hardly ever visited in the same order. This result is interesting since Dwingelderveld National Park, like many natural parks in the Netherlands, has predefined routes that are well marked on maps available to the visitors, as well as on the information boards and signposts. This result seems to contradict the findings of van Marwijk (2009) who reported that 66% of the visitors follow a predefined route in his experiment. Two facts should be borne in mind, though. First, visitors can follow the routes in both directions, changing the order of visited places. Second, the figures reported in van Marwijk's study are based on visitors' answers to a survey after they finished their visit, and not on GPS positioning data.

The answers to the three questions we posed at the beginning illustrate the suitability of our approach to analysing the flow of visitors. The proposed approach has some advantages over previous methods mentioned in the introduction. One advantage is that MSPs, which constitute the building blocks of the analysis, do not need spatial or temporal thresholds. In addition, using the BIDE+ algorithm to detect GSPs helps to avoid redundancy in the results, making it easier to explore and interpret them. Moreover, the combination of different methods provides the flexibility required for an exploratory approach. Our approach is also strongly related to the geographical context in which movement occurs, which helps with interpreting the meaning of the movement patterns. An additional advantage is that all the steps can be performed using publicly available GIS software and open source code — an important advantage for applying the approach in other areas.

The proposed approach can provide useful information for the design, implementation and monitoring of management practices in natural recreational areas. Park managers, for example, can use the proposed approach to assess the popularity of different places in their area and understand the flow of people from the park entrances to those places. This in turn can be used to evaluate the location of signs, design visitor routes and manage the flow of visitors to avoid crowding.

Although this analysis was restricted to the spatial behaviour of the visitors, their socio-economic backgrounds, purposes and motivations, can also be included to differentiate and compare different target groups. Moreover, the flows of different groups and their changes over time can be used as indicators of coping mechanisms in crowded areas (Manning & Valliere, 2001) and for studying possible processes of exclusion and domination (Ostermann, 2009).

The implications of these results are also of potential interest to tourism businesses. Holyoak and Carson (2009) identify several areas that can benefit from the study of the movement of visitors in a broader, regional context. Managers can segment the market and offer more diverse and focused options for specific groups of visitors. Information and marketing strategies can take into account the order in which the places are visited to provide relevant and appropriate information about the destinations. Moreover, researchers in recreational areas can study the flow of visitors to support the design of other data collection methods, such as deciding where and when conduct interviews, surveys and locate direct observers.

Although these results describe the flow of visitors in a relative small area, the approach can be used in larger areas and for longer periods of time. In fact, one of the advantages of Movement Suspension Patterns is that they are derived from the spatial-statistical properties of the dataset, and can therefore work at very different scales. Theoretically, the lower limit corresponds to the spatial accuracy of the data (i.e. it would be not possible to differentiate flows in an area of 10 m radius for data collected with single-frequency GPS receivers without any signal correction) and the upper limit is set by the study design (the area in which the participants will be monitored). In this study, preconfigured GPS receivers were given to the visitors at the entrances of the park. Other portable devices, such as the visitors' own smartphones or tablets could be used to complement data collection. For example, visitors willing to participate in the research could install tracking applications in their devices, in return for receiving location-aware information during their visit to the area.

6. Conclusions

Understanding the spatial behaviour of visitors in recreational natural areas is a key issue for effective management. Using GPS tracking technology, managers and researchers can collect data on the routes followed by individuals to analyse how they interact with the geographical features in the area. When the movement of several people is analysed, some patterns may emerge indicating the existence of common structures in the spatial behaviour. In line with this idea, we suggested that movement patterns representing the flow of visitors could further our understanding of the collective spatial behaviour of the visitors.

In this article, we present a novel approach to exploring the flow of visitors in natural areas through the combined analysis of two kinds of movement patterns extracted from GPS positioning data. Movement Suspension Patterns were useful for uncovering the main attractions, while Generalized Suspension Patterns were helpful in understand common structures in the order that those places were visited. We demonstrated the application of the approach by analysing the movement of visitors in the Dwingelderveld National Park. The results suggest that the proposed approach helps us to understand the aggregated spatial behaviour of the visitors in the park.

In future research, we are envisaging new ways to build the sequential space–time cube to allow a better dynamic exploration of the detected GSPs. We are also developing a better method to establish the duration of the individual MSPs using hierarchical clustering.

Acknowledgements

The authors thank Ramona van Marwijk and Rene Jochem from Wageningen University, for the Dwingelderveld National Park data collection. The authors are grateful with the anonymous reviewers whose suggestions helped to improve the paper.

References

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the eleventh international conference on data engineering* (pp. 3–14).
Alvares, L. O., Bogorny, V., Kuijpers, B., Macedo, J., Moelans, B., & Vaisman, A. A. (2007). A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th annual ACM international symposium on advances in geographic information systems* (pp. 162–169).
Andrienko, G., & Andrienko, N. (2008). Spatiotemporal aggregation for visual analysis of movements. In *Visual analytics science and technology* (pp. 51–58). IEEE.
Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis, 27*(2), 93–115.
Batty, M., De Syllas, J., & Duxbury, E. (2003). The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades. *International Journal of Geographical Information Science, 17*(7), 673–697.
Beyer, H. L. (2010). *Geospatial modelling environment v0.4*. Retrieved 09/04/2011, from. http://www.spatialecology.com.
Bierlaire, M., Antonini, G., & Weber, M. (2007). Behavioral dynamics for pedestrians. In K. Axhausen (Ed.), *Moving through nets: The physical and social dimensions of trave* (1st ed.). (pp. 81–105) Amsterdam: Elsevier.
Bogorny, V., Heuser, C., & Alvares, L. (2010). A conceptual data model for trajectory data mining. In S. I. Fabrikant, T. Reichenbacher, M. V. Kreveld, & C. Schlieder (Eds.), *Geographic information science. 6th international conference, GIScience* (pp. 1–15). Springer.
Chiesura, A. (2004). The role of urban parks for the sustainable city. *Landscape and Urban Planning, 68*(1), 129–138.
Demšar, U., & Virrantaus, K. (2010). Space–time density of trajectories: exploring spatiotemporal patterns in movement data. *International Journal of Geographical Information Science, 24*, 1527–1542.
Dodge, S., Weibel, R., & Lautenschütz, A. K. (2008). Towards a taxonomy of movement patterns. *Information Visualization, 7*, 240–252.
Fournier-Viger, P., Nkambou, R., & Nguifo, E. (2008). A knowledge discovery framework for learning task models from user interactions in intelligent tutoring systems. *Advances in Artificial Intelligence, 5317*, 765–778, MICAI 2008.
Golicnik, B., & Ward Thompson, C. (2010). Emerging relationships between design and use of urban park spaces. *Landscape and Urban Planning, 94*(1), 38–53.
Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779–782.
Gudmundsson, J., Laube, P., & Wolle, T. (2007). Movement patterns in spatiotemporal data. In S. Shekhar, & H. Xiong (Eds.), *Encyclopedia of GIS* (pp. 726–732). Springer-Verlag.
Hägerstrand, T. (1970). What about people in regional science? *Papers in Regional Science, 24*(1), 6–21.
Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques* (2nd ed.). Amsterdam: Morgan Kaufmann.
Holyoak, N., & Carson, D. (2009). Modelling self-drive tourist travel patterns in desert Australia. In *The 32 Australasian transport research forum – ATRF*, Auckland, New Zealand.
Hoogendoorn, S. P., & Bovy, P. H. L. (2005). Pedestrian travel behavior modeling. *Networks and Spatial Economics, 5*(2), 193–216.
Kwan, M. P. (2004). GIS methods in time-geographic research: geocomputation and geovisualization of human activity patterns. *Geografiska Annaler Series B Human Geography, 86*(4), 267.
Laube, P. (2009). Progress in movement pattern analysis. In B. Gottfried, & H. K. Aghajan (Eds.), *Behaviour monitoring and interpretation – BMI – smart environments* (pp. 43–71). Amsterdam: IOS Press BV.
Manning, R. E., & Valliere, W. A. (2001). Coping in outdoor recreation: causes and consequences of crowding and conflict among community residents. *Journal of Leisure Research, 33*(4), 410–426.
van Marwijk, R. (2009). *These routes are made for walking: Understanding the transactions between nature, recreational behaviour and environmental meanings in dwingelderveld national park, the Netherlands*. PhD, Wageningen University.
Mckercher, B., & Lau, G. (2008). Movement patterns of tourists within a destination. *Tourism Geographies: An International Journal of Tourism Space, Place and Environment, 10*(3), 355–374.
Muhar, A., Arnberger, A., & Brandenburg, C. (2002). *Methods for visitor monitoring in recreational and protected areas: An overview. Monitoring and management of visitor flows in recreational and protected areas*. Vienna. pp. 1–6.
National Marine Electronics Association. (2010). *NMEA standards*. Retrieved 09/04/2010, from. http://www.standards.nmea.org.
Natuurmonumenten. (2009). *Natuurmonumenten web site*. Retrieved 09/04/2011, from. http://www.natuurmonumenten.nl.
Nielsen, T., & Hovgesen, H. (2004). GPS in pedestrian and spatial behaviour surveys. In *The fifth international conference on walking in the 21st century, Copenhagen*.
Orellana, D., & Wachowicz, M. (2011). Exploring patterns of movement suspension in pedestrian mobility. *Geographical Analysis, 43*, 241–260.
Ostermann, F. O. (2009). Indicators for socially sustainable park use – results from a case study. In *14th conference on regional planning, Sitges, Spain* (pp. 243–252).
Palma, A. T., Bogorny, V., Kuijpers, B., & Alvares, L. O. (2008). A clustering-based approach for discovering interesting places in trajectories. In *Proceedings of the 2008 ACM symposium on applied computing* (pp. 863–868).
Pol-Recreatie. (2003). *Recreatieverblijf't polletje*. Retrieved 09/04/2011, from. http://www.pol-recreatie.nl/.
Rinzivillo, S., Pedreschi, D., Nanni, M., Giannotti, F., Andrienko, N., & Andrienko, G. (2008). Visually–driven analysis of movement data by progressive clustering. *Information Visualization, 7*(3/4), 225–239.
van Schaick, J., & van der Spek, S. (Eds.). (2008). *Urbanism on track: Application of tracking technologies in urbanism*. Amsterdam: IOS.
Scheepens, R., Willems, N., van de Wetering, H., & van Wijk, J. J. (2011). Interactive visualization of multivariate trajectory data with density maps. In *Paper presented at the Pacific Visualization Symposium (Pacific Vis)* (pp. 147–154). IEEE.
Shoval, N. (2010). Monitoring and managing visitors flows in destinations using aggregative GPS data. In U. Gretzel, R. Law, & M. Fuchs (Eds.), *Information and communication technologies in tourism* (pp. 171–183). Vienna: Springer Vienna.
Shoval, N., & Isaacson, M. (2009). *Tourist mobility and advanced tracking technologies*. Routledge.
Silberschatz, A., & Tuzhilin, A. (1995). On subjective measures of interestingness in knowledge discovery. In *The first international conference on knowledge discovery & data mining, Montreal, Quebec, Canada* (pp. 275–281).
Song, C., Qu, Z., Blumm, N., & Barabasi, A. (2010). Limits of predictability in human mobility. *Science, 327*(5968), 1018–1021.
Spaccapietra, S., Parent, C., Damiani, M. L., Macedo, J., Porto, F., & Vangenot, C. (2008). A conceptual view on trajectories. *Data & Knowledge Engineering, 65*(1), 126–146.
Taczanowska, K., Muhar, A., & Brandenburg, C. (2008). Potential and limitations of GPS tracking for monitoring spatial and temporal aspects of visitor behaviour in recreational areas. In *Fourth international conference on monitoring and management of visitor flows in recreational and protected areas* (pp. 451–455).
Tobler, W. (2003). *Movement mapping*. Center for Spatially Integrated Social Science. Retrieved 09/04/2011, from. http://www.csiss.org/clearinghouse/FlowMapper/MovementMapping.pdf.
Topografix. (2010). *GPX 1.1 schema documentation*. Retrieved 09/04/2011, from. http://www.topografix.com/GPX/1/1/.
Wang, J., Han, J., & Li, C. (2007). Frequent closed sequence mining without candidate maintenance. *IEEE Transactions on Knowledge and Data Engineering, 19*(8), 1042–1056.
Wood, Z., & Galton, A. (2009). Classifying collective motion. In B. Gottfried, & H. K. Aghajan (Eds.), *Behaviour monitoring and interpretation* (pp. 129–155). IOS.
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.