Recognizing patterns of movements in visitor flows in nature areas

Arend Ligtenberg, Ramona van Marwijk, Bart Moelans, Bart Kuijpers

Abstract — This paper presents some approaches for geo-spatial analysis of movement behavior of visitors of recreational areas. The approaches are bases on the use of moving object databases containing Temporary Annotated Sequences (TAS). The TAS result from the use of GPS or mobile phones for tracking visitors. Two examples are presented for a case study carried out in the Dutch National Park Dwingelderveld. About 461 visitors were tracked using a GPS device. Based on these GPS recordings their trajectories have been reconstructed. The relation between the type of landscape in terms of openness and the speed of movement have been analyzed. Additionally a similarity analyses based on Fréchet analysis shows clusters of movements.

Index Terms — GIS, Movement Behavior, Spatial Temporal Analyses

1 Introduction

It is well known that large numbers of visitors in nature area increased the burden on the management of these areas. Quality of these areas might be affected in several ways. Besides effects on the ecological quality of the area through disturbance by people also the recreational quality might be affected by crowdedness, noise, and differences in recreation behavior as found for example in situations where mountain bikers or horse riders meet walkers. Crucial to understanding the effects of visitor behavior on both ecological and recreational qualities is knowledge about the movement behavior of visitors and the relation with its current socio-spatial environment. Although established frameworks exist explaining these relations [1-3] aspect like behavior of individuals in relation with their physical environment, interaction amongst individuals and differences in behavior as a consequence of intention and activity are still difficult to measure and analyze.

The use of easily available GPS systems and positioning based on mobile phones enable for additional methods to analyze movement behavior compared to the more traditional methods applied in recreation research (predominantly surveys). The research presented here aims to explore the additional value of using the "moving objects data" resulting from tracking people with GPS or other location aware devices such as mobile phones.

As already pointed out by [4] the movement behavior of people is strongly influence by the various constraints imposed by their environment and their capabilities. Reconstructing movement behavior from moving objects i.e. spatial temporal recordings might provide additional information to recreation researchers as well as policy makers. Currently research efforts in this field are directed to visual analytics i.e. the exploration of visual (graphical) representations of data using interactive visualization tools [5] and Geographical Knowledge Discovery [6]

2 Theory

2.1 Temporary Annotated Sequences (TAS)

Movements of objects trough space generally are recorded using TAS. In general a TAS consist of the minimum tuple: <id, x, y, t>. The id is an unique identifier, x, and y are spatial coordinates in some coordinate system and t is the data timestamp of the recording [7] resulting from objects (visitors) moving through a geographical space. These visitors might either broadcast their Space Time (ST) positions (mobile phones), store their ST positions (GPS, navigation devices), or allowing their ST position to be monitored (for example RFID).

2.2 Trajectories

Based on TAS the paths followed by the visitors can be reconstructed. Such a reconstruction of followed paths is generally referred to as geo-spatial lifelines or trajectories [8].
A trajectory by definition is a subset of the total recordings describing a visitor's movement. Each trajectory knows a begin and end and mostly a number of stops (and starts). The definition of a stop mainly depends on the scale of analyses. The question if a temporary non-movement is the begin or end of a trajectory or just a stop depends on the goal of the analysis and the related spatial and temporal resolutions. Same counts for the question if a TAS is just a ST-point in the geo-spatial lifeline or a stop. For example the question if waiting for a crossing or a traffic light a stop or just a part of the trajectory can only be answered in the context of a crisp defined goal of analysis.

2.3 Characteristics of trajectories

Once trajectories have been constructed various characteristics of the movements of visitors can be analysed. Trajectory characteristics can be divided into characteristics based on a single trajectory and characteristics based on multiple trajectories. Single trajectory characteristic are, speed, acceleration, shape, number of stops and stop duration. Multiple trajectory characteristics are, amongst others: density in space and or time i.e. how many trajectories are within a certain time interval within a certain distance of each other, and interactions; how many trajectories interact in space and time. Interactions are for example crossing or convergent, divergent or parallel movements of trajectories within the same period and area.

3 DWINGELERVELD CASE

To demonstrate the use of TAS and trajectories a number of analyses have been carried out using a GPS-tracking dataset of visitors of the Dwingelderveld National Park (DNP). This Dutch nature area – containing 3700 ha and situated in Drenthe, a province in north eastern Netherlands – was chosen because of its recreational attractiveness and ecological quality. The area is ecologically important as it is the largest wet heath land area in Northwest Europe (1550 ha). The heath land area is bordered by forest (2000 ha). The DNP is also a Natura2000 area, which means it is part of a European network of important nature conservation areas. The DNP receives at least 1.6 million visitors yearly. It is a typical Dutch nature recreation area with an extensive recreational network for both short strolls (60 km marked trails that are each less than 7 km in length) and long walks, for cycling ("normal", racing, mountain biking) and for horse riding. During a study amongst 399 visitors of the DNP in August 2006 detailed tracks where recorded using Garmin GPS devices. Fig. 1 shows a small excerpt of the complete dataset of the total survey consisting of about 142000 TAS. Fig. 1 shows also some errors in the data due to false GPS recordings. Especially under forest or near tall buildings GPS positioning is

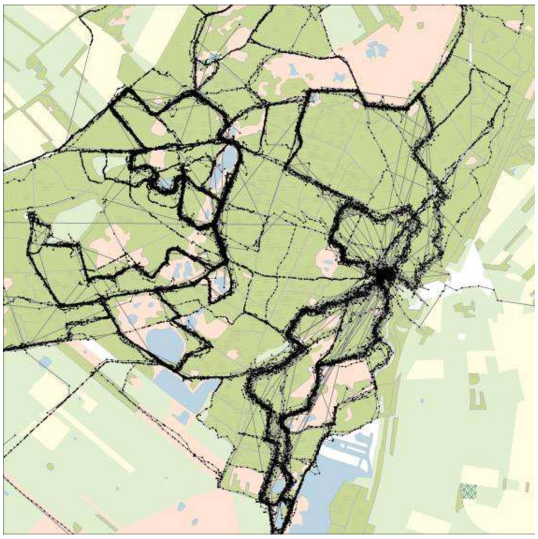

Fig 1: Part of the Dwingelderveld dataset showing the TAS en constructed trajectories

rather inaccurate and non-stable due to poor satellite reception and multipath reflectance. Of the 399 collected GPS tracks only 311 (78%) were complete. Additionally GPS outliers were removed by only considering recording within 25 m at each side of the paths in the path network. This accounts reasonable for the

4 Analyses

To demonstrate how the mentioned trajectory characteristics might help getting insight in the movement behaviour of visitors the following are

- *Arend Ligtenberg is with the Centre for Geo-information of Wageningen University. the Netherlands. E-mail arend.ligtenberg@wur.nl*
- *R. van Marwijk is with the Socio-Spatial Analysis and Forest & Nature Conservation Policy groups at Wageningen University, the Netherlands. E-mail: ramona.vanmarwijk@wur.nl*
- *Bart Moulans is with the theoretical computer science group at Hasselt University, Belgium. E-mail: bart.moelans@uhasselt.be*
- *Bart Kuijper is with the theoretical computer science group at Hasselt University, Belgium, E-mail: bart.kuijpers@uhasselt.be*

calculated:

1. Walking speed in relation with the openness the type of landscape
2. Similarities amongst various trajectories.

All analyses are carried out using the standard ESRI ArcGis 9.2 software. For reconstructing trajectories from the TAS an additional plug-in (HawthTools) was used for reasons of convenience

4.1 Walking speed

As most of the visitors (65%) tracks in DNP follow marked trails, we expect specific landscape preferences of minor influence. However, interesting to know would be the influence of the type of landscape on the visitors' movement behavior (in this case the walking speed). To be able to analyze this a dataset have been prepared showing for all walking paths in the DNP the visibility of landscape for each paths segment (Fig2) [9]. These visibility was classified in 4

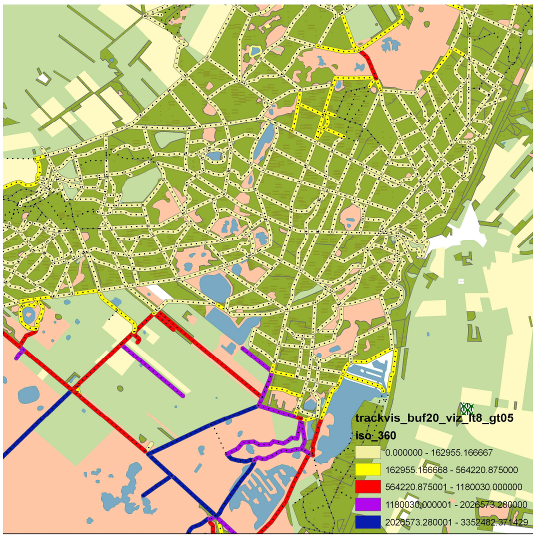

Fig 2: Part of the dataset showing the openness of the landscape based on the visible areas (m2) from a path (calculated at a 50 mtr. interval)

landscape types: closed, open, boundary and mixed. Based on the trajectories the speed between each TAS was calculated and combined with the classified paths network using GIS based procedures. The resulting dataset provides a detailed insight into the speed at all segments of the path network (see Fig. 3)

Next, using these two datasets the average speed for each path class was calculated showing the effect of the landscape difference upon the movement speed of the visitors (Fig. 4). Although the differences are not very high and the observed standard deviations rather high this type of analyses might provide insight in the effects of landscape upon movement behavior. At this moment the topographical map was used to generated the landscape typology. A more targeted analyses and validation of the dataset, with respect to the

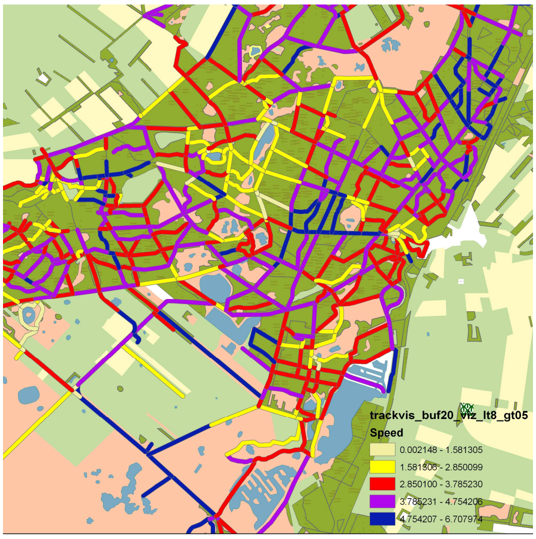

Fig 3: Part of the dataset showing the average speed at the followed paths

openess would probably lead to a better results. Additionally the used GPS device performed rather poor under forest cover leading to fluctuation in speed as a result of inaccurate positioning. Especially walking speeds as in this case this leads to relative high deviations.

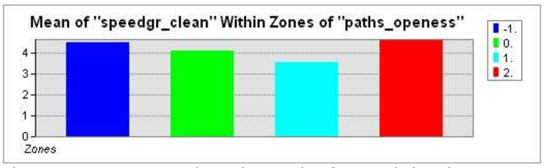

Fig 4: average speed on the paths for each landscape type (-1 = closed, 0 is boundary, 1 = open, and 2 = mixed). Stand. dev. respectively 1.64, 1.74, 1.93, 1.82

4.2 Datamining

A second type of analyses demonstrated here that allows for insight into the movement behavior of visitors is finding patterns in data using data mining techniques. Fig. 5 shows the result of a cluster analyses based on the Fréchet distance, a distance measure that accounts for the continuity of the trajectories.[10]. For the clustering we used k-medoids algorithm [11] with as similarity measure the discrete Fréchet distance. Many applications consider the Fréchet distance for curves as a good measure for similarity between polylines (i.e., traces of trajectories). Because of the high computational cost of this distance measure the ``less correct'' derivate for polylines, called the discrete Fréchet distance is used.

Clustering based on similarities can reveal various modes of use of an area in terms of frequently followed routes, and recreational pressure on certain parts of an area. The 4 clusters shown in Fig. 5 are based on visitors who stated not to follow a predefined route (browsers).

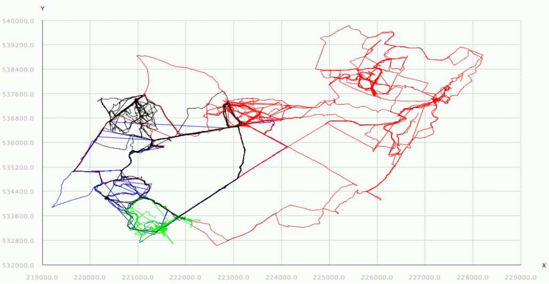

Fig 5: results of the discrete Fréchet similarity analysis (4 clusters) for routes followed by browsers (entire study area)

The clusters show links with the parking areas that define in most cases the begin en ends of the trajectories. This type of similarity analysis seems to be especially relevant for discovering patterns in movements of free moving visitors or movements over dense or large networks of paths.

6 DISCUSSION

The techniques shown in this paper are an illustration to show the use of moving object data for additional analyses of visitors behavior. It provides techniques to explicitly include environmental characteristics into analyses of movement behavior. There is however, the need to test and validate these type of analyzes for their practical applicability in recreation research. A complicating aspect at this moment seems to be the quality of current GPS recordings. They often are unstable, hampering the tracking of the subtle differences in movement behavior in individual trajectories. They easily get blurred in the noise caused by inaccurate measurement of GPS. The launch of the Galileo network (expected around 2011) will probably offer a better accuracy combined with a better performance in forest areas.

Data mining techniques like similarity matching and clustering offer additional insight in the general spatial and/or temporal patterns of movements cause by visitors. Managers and policy makers can use these patterns to increase their insight in the use of an area in different periods. However, before these techniques successfully can be applied additional research is needed into the relation between movement, environment and resulting patterns. There is a strong need for the development of concepts and methods that relate data-oriented models commonly applied in data-mining, the spatial modeling of geo-science and social models used in recreation research.

7 REFERENCES

1. Hartig, T., *Nature experience in transactional perspective.* Landscape an Urban Planning, 1993. **25**: p. 17-36.
2. Marwijk van, R., B.H.M. Elands, and J. Lengeek, *Experiencing nature: the recognition of the symbolic environment within researchnd management of visitors flows.* For. Snow Landsc. Rex, 2007. **81**(1/2): p. 59-79.
3. Lengkeek, J., J.W. Te Kloeze, and R. Brouwer, *The multiple realities of the rural environment. The significance of tourist images for the countryside*, in *Images and realities of rural life. Wageningen perspectives on rural transformations*, H. De Haan and N. Long, Editors. 1997, Van gorcum: Assen.
4. Hägerstrand, T., *What about people in regional science?* Papers of the Regional Science Association, 1970. **24**: p. 7-21.
5. Andrienko, G., et al., *Visual Analytics Methods for Movement Data*, in *Mobility, Data Mining and Privacy*, F. Giannotti and D. Pedreschi, Editors. 2008, Springer: Berlin Heidelberg. p. 375-408.
6. Wachowicz, M., et al., *Characterising the Next Generation of Mobile Applications through a Privacy-Aware Geographic Knowledge Discovery Process*, in *Mobility, Data Mining and Privacy*, F. Giannotti and D. Pedreschi, Editors. 2008, Springer: Berlin Heidelberg. p. 39-70.
7. Hornsby, K. and M.J. Egenhofer, *Modeling moving objects over multiple granularities.* Annals of Mathematics and Artificial Intelligence, 2002. **36**: p. 177-194.
8. Laube, P., S. Imfeld, and R. Weibel, *Discovering relative motion patterns in groups of moving point objects.* International Journal of Geographical Information Science, 2005. **19**(6): p. 639 - 668.
9. Weitkamp, G., et al. *Three Sampling Methods for Visibility Measures of Landscape Perception*. in *COSIT 2007*. 2007. Melbourne, Australia: Springer-Verlag.
10. Kuijpers, B., B. Moelans, and W.v.d. N. *Qualitative Polyline Similarity Testing with Applications to QuerybySketch, Indexing and Classification*. in *ACMGIS 06*. 2006. Arlington, Virginia AMC.
11. Tan, P.N., M. Steinbach, and V. V Kumar, *K-medoids: Introduction to Data Mining*. 2005, Boston, MA, USA: Addison-Wesley Longman Publishing Co.

A. Ligtenberg is an Ass. Prof. at the Centre for Geo-information of Wageningen university. He holds a BSc as an irrigation engineer and a MSc in environmental planning and design from Wageningen University majoring in spatial planning and GIS. In 2006 he obtained his PhD on research on the use of multi-agent simulation for interactive multi-actor spatial planning. His current research interests include geo-visualization, complex spatial systems, and agent based spatial modelling.

R. van Marwijk holds a BSc in Rural Development

Studies (2000) and a MSc in Land Use Planning (Recreation & Tourism, 2003). She started as junior researcher at Socio Spatial Analysis Group, Wageningen University until end 2004 and then started her PhD at Forest and Nature Conservation Policy Group, Wageningen University. Her MSc-thesis ("A small contribution? Small tourism entrepreneurs and sustainable development in Malindi, Kenya") received the Ed van Thijn Thesis Award in 2004. In 2006 her joint poster with Karolina Taczanowska ("Types of Typologies – From Recreationists & Tourists to Artificial Agents") was rated as best Poster at MMV3 and she received the Storm vd Chijs award (promising female PhD student of Wageningen University). Her current research interests are nature based recreation and tourism, sustainable tourism development, recreation quality, leisure experiences, place attachment and meaning, tourist types.

B. Moelans is a member of the theoretical computer science group at Hasselt University.
He holds a MSc in computer science from the Transnational University Limburg majoring in operations research.
His area of research are spatial databases, spatio-temporal databases and GIS.

B. Kuijpers is Prof. at the theoretical computer science group at Hasselt University.
In 1998 he obtained his PhD on topological properties of spatial databases in the polynomial constraint model, which was awarded with the IBM-Belgium Prize.
His area of research are spatio-temporal and constraint databases, GIS, database theory and logic **.**

Acknowledgments:
This work is funded by the Dutch Space for Geo-information program (RGI) and the EU Geopkdd project for funding this research. Special thanks to Gerd Weitkamp of the Centre for Geo-Information of Wageningen UR for carrying out the visibility analyses.