Report from Dagstuhl Seminar 12512

Representation, Analysis and Visualization of Moving Objects

Edited by
Joachim Gudmundsson[1], Patrick Laube[2], and Emiel Van Loon[3]

1 **The University of Sydney, AU,** joachim.gudmundsson@sydney.edu.au
2 **Universität Zürich, CH,** patrick.laube@geo.uzh.ch
3 **University of Amsterdam, NL,** vanloon@uva.nl

Abstract

From December 16 to December 21, 2012, the Dagstuhl Seminar 12512 "Representation, Analysis and Visualization of Moving Objects" was held in Schloss Dagstuhl – Leibniz Center for Informatics. The major goal of this seminar was to bring together the diverse and fast growing, research community that is involved in developing better computational techniques for spatio-temporal object representation, data mining, and visualization of moving object data. The participants included experts from fields such as computational geometry, data mining, visual analytics, GIS science, urban planning and movement ecology. Most of the participants came from academic institutions but some also from government agencies and industry. The seminar has led to a fruitful exchange of ideas between different disciplines, to the creation of new interdisciplinary collaborations and to recommendations for future research directions. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper.

Seminar 16.–21. December, 2012 – www.dagstuhl.de/12512
1998 ACM Subject Classification E.1 Data Structures, F.2 Analysis of Algorithms and Problem Complexity, H.2.8 Database Applications
Keywords and phrases movement analysis, algorithms, visualization, geo-ecology, urban planning
Digital Object Identifier 10.4230/DagRep.2.12.89
Edited in cooperation with Georgina Wilcox

1 Executive Summary

Joachim Gudmundsson
Patrick Laube
Emiel Van Loon

The seminar brought together researchers and domain experts involved in developing and utilizing methods for knowledge extraction from massive amounts of data from moving objects. This knowledge is essential to substantiate decision making in public and private sectors, in application domains such as transportation modelling, urban planning, tourism, wildlife ecology, spatial epidemiology, location-based services, flight safety, and marine safety. Moving object data typically include trajectories of discrete spatial objects (e.g. humans, vehicles, animals, and goods), continuous phenomena (e.g. storms, ocean currents) as well as trajectories of abstract concepts (e.g. information flow, moving data points in attribute space) or even vectors of spreading diseases. Technologies for object tracking are low cost and increasingly reliable in terms of coverage and accuracy, hence movement records are nowadays generated in huge volumes on a routine basis, using diverse technologies such as radio frequency mapping, Global Navigation Satellite Systems, video sequences and Doppler radar.

The computational analysis of movement data has seen a successful first decade with progress made in capturing, preprocessing, storing, indexing and querying movement data, combined with promising results in visualizing movement and detecting movement patterns. However, whereas such basic progress in handling movement data was needed for establishing a new field and attracting funding, attention must now move on towards the extraction of useful information and process knowledge from tracking data.

In many application fields the need for analysing large sets of trajectories is evident and crucial; however, only very rudimentary automated analysis tools are available and anything more advanced is currently analyzed manually. As an example there are several companies tracking the movement of football players through video with a frequency of at least 25Hz and accuracy of approximately 10cm. But most of the analysis and the annotation (passes, throw-ins, goals etc) of the data are still made manually. Thus the analysis part has been neglected and in comparison with the image processing part it is technologically far behind in the development. A reason for this is the obvious lack of theoretical and practical solutions for many crucial fundamental problems.

For that reason, this seminar focussed on formalizing methods for algorithmic analysis, visual analytics, data mining and knowledge discovery, defined by a multidisciplinary team of researchers and practitioners.

Participants and Format

The seminar brought together researchers from several disciplines involved in developing and utilizing computational techniques for spatiotemporal object representation, data mining, and visualization. This community encompasses an interdisciplinary mix of methodologically oriented as well as application oriented researchers. The methods-oriented researchers are from fields such as theoretical computer science, spatial databases, knowledge discovery and data mining, visual analytics, and geographic information science. They were complemented by application scientists from various fields, especially behavioral ecology and urban planning.

Drawing upon positive experiences in previous seminars of this series, oral presentation sessions were complemented by special sessions dedicated to open research questions and project ideas, as well as to discussions in small, concurrent break-out groups focussing on a specific domain or case studies. Since the participants of the seminar came from quite different backgrounds, concise survey talks on the first two days were delivered on movement ecology (Ran Nathan), industry movement research (Ben Loke) and spatial data mining (Győző Gidófalvi).

A data challenge was organized for the participants prior to the seminar. This ensured that the participants were well aware of the application domain which was the focus of the seminar and it gave the domain experts a possibility to see the potential use of various different approaches. The data challenge included a bird migration data set provided by Emiel van Loon.

Many interesting research results were presented, demonstrating the progress in this field. The participants were highly satisfied with the quality of the seminar; especially the involvement of domain specialists from relevant application domains was highly appreciated.

2 Table of Contents

3 Overview of Talks

3.1 Towards an algorithmic framework for movement analysis with movement models

Kevin Buchin (TU Eindhoven, NL)

Joint work of Buchin, Kevin; Sijben, Stef; Arseneau, T. Jean Marie; Willems, Erik P.

In trajectory data a low sampling rate leads to high uncertainty in between sampling points, which needs to be taken into account in the analysis of such data. However, current algorithms for movement analysis ignore this uncertainty and assume linear movement between sample points. In this talk I present a framework for movement analysis using the Brownian bridge movement model (BBMM), that is, a model that assumes random movement between sample points. Many movement patterns are composed from basic building blocks like distance, speed or direction. I will show how to express their distribution over space and time in the BBMM and will demonstrate our framework by computing a "following" pattern. Our motivation to study the BBMM stems from the rapidly expanding research paradigm of movement ecology. To this end, I will present an application of the framework to the simultaneous movement of groups of wild and free-ranging primates.

3.2 Segmenting geese tracks by movement states

Maike Buchin (TU Eindhoven, NL)

Joint work of Buchin, Maike; Kruckenberg, Helmut; Kölzsch, Andrea
Main reference M. Buchin, H. Kruckenberg, A. Kölzsch, "Segmenting Trajectories based on Movement States," in Proc. of the 15th Int'l Symp. on Spatial Data Handling (SDH), 2012.

We report on work that originated during the Lorentz Center 2011 workshop. Dividing movement trajectories according to different movement states of animals has become a challenge in movement ecology, as well as in algorithm development. In this study, we revisit and extend a framework for trajectory segmentation based on spatio-temporal criteria for this purpose. We adapt and extend the framework for segmenting by movement states of an individual. We implement the framework and evaluate it using the example of two spring migration tracks of white-fronted geese. Manual and automatic classifications are compared and extensions discussed.

3.3 Movement in cordon structured spaces

Matt Duckham (The University of Melbourne, AU)

Joint work of Duckham, Matt; Bleisch, Susanne; Both, Alan; Lyon, Jarod; Laube, Patrick; Wark, Tim
Main reference A. Both, M. Duckham, P. Laube, T. Wark, J. Yeoman, "Decentralized monitoring of moving objects in a transportation network augmented with checkpoints," The Computer Journal, in press.

Traditional definitions of a "trajectory" (e.g., "set of n moving points whose locations are known as t consecutive time steps,") conjure up movement in space through time: the position of a moving object recorded at successive times [1]. However, a complementary view of movement is to record the time at which an object moves past known checkpoints, called cordons. These cordon-structured spaces and movements are common in a wide range of applications, from e-tolling and public transport smart cards; through resource tracking for emergency response; to environmental monitoring of tagged animal movements. Trajectory and cordon-structured movements are to some extent inter-convertible: both are representations of the same underlying phenomena. However, two important differences exist. First, the often lower, but more importantly variable spatial and temporal granularity of motion tracking in cordon-structured spaces makes many trajectory analysis tools not applicable to cordon structure movement data (such as path shape and movement velocity derivation). Second, the location of cordons is typically salient in the application. Cordons are placed in locations that provide a meaningful qualitative structure to the space. Unlike trajectories, observations of spatial location in cordon structured movement are not arbitrary.

References

1 Gudmundsson, J., van Kreveld, M., and Speckmann, B. Efficient detection of patterns in 2D trajectories of moving points. *GeoInformatica* 11, 195–215, 2007
2 Both, A., Duckham, M., Laube, P., Wark, T., Yeoman, J. (2013) Decentralized monitoring of moving objects in a transportation network augmented with checkpoints. *The Computer Journal*, in press.

3.4 Brief survey of trajectory prediction

Győző Gidófalvi (KTH – Stockholm, SE)

Joint work of Győző Gidófalvi; Fang Dong
Main reference G. Gidófalvi, F. Dong, "When and Where Next: Individual Mobility Prediction," in Proc. of MobiGIS, 2012.

Over the past decade, several technological and social trends enabled the large scale collection of moving object trajectories. The potential utility of likely regularities in these trajectories has spawned a new sub-field in data mining: trajectory data mining. After a brief review of frequent sequential pattern mining, the talk presents a brief survey and taxonomy of trajectory prediction—an important trajectory mining task—and illustrates how different approaches perform trajectory prediction at different spatio-temporal scales and under different application settings [1, 2, 3].

References

1 G. Gidófalvi, M. Kaul, C. Borgelt and T. Bach Pedersen. *Frequent Route Based Continuous Moving Object Location- and Density Prediction on Road Networks*. In Proc. of ACM-GIS, pp. 381-384, 2011.

2 G. Gidófalvi and F. Dong. *When and Where Next: Individual Mobility Prediction*. In Proc. of MobiGIS, 2012.

3 A. Monreale, F. Pinelli, R. Trasarti and F. Giannotti. *WhereNext: a Location Predictor on Trajectory Pattern Mining*. In Proc. of KDD, pp. 637-645, 2009.

3.5 The classification of passes in football

Joachim Gudmundsson (University of Sydney, AU)

Joint work of Estephan, Joël; Gudmundsson, Joachim

Studies have shown that coaches are less than 45% correct in their post-game assessment of what occurs during the 90 minutes of a football match [1]. As a consequence there has recently been a strong research interest in automatically analyst the performance of a player or a team. The time and money lately being exhausted on sports analysis only begins to exemplify its importance to the sporting community.

In this talk we explore the automatic classification of the quality of passes in team sports. We first generated an extensive data set of football passes and had them manually classified by experts. Models were then built using a variety of different feature selection methods and supervised learning algorithms. Four models were selected, two classifying the risk of a pass and the others the worth of a pass. The results were quite compelling/positive/encouraging with the accuracy rates ranging from 71% to 96%.

The developed models have been added to the ever growing set of tools that provide coaches and players with an in-depth analysis of the game.

References

1 I.M. Franks and G Miller. *Eyewitness testimony in sport*. Journal of Sport Behaviour, 9:39–45, 1986.

3.6 Analyzing and visualizing movement data with head/tail breaks

Bin Jiang (University of Gävle, SE)

Main reference B. Jiang, "Head/tail Breaks: A New Classification Scheme for Data with a Heavy-tailed Distribution," arXiv:1209.2801 [physics.data-an].
URL http://arxiv.org/abs/1209.2801

This paper summarizes a novel method for efficiently and effectively analyzing and visualizing movement data. The novel method, called head/tail breaks, was initially developed as a classification method for data with a heavy tailed distribution. Since most movement data exhibit heavy tailed distributions, head/tail breaks can be an important means for data analysis and visualization.

3.7 Trajectory grouping structure

Marc van Kreveld (Utrecht University, NL)

Joint work of Buchin, Kevin; Buchin, Maike; van Kreveld, Marc; Speckmann, Bettina; Staals, Frank

The collective motion of a set of moving entities such as a set of people, birds, or other animals, is characterized by groups arising, merging, splitting, and ending. Given the trajectories of these entities, we define and model a structure that captures all of such changes using the Reeb graph, a concept from topology. The trajectory grouping structure has three natural parameters that allow more global views of the data in group size, group duration, and entity inter-distance. We prove complexity bounds on the maximum number of maximal groups that can be present, and give algorithms to compute the grouping structure efficiently. We also study how the trajectory grouping structure can be made robust, that is, how brief interruptions of groups can be disregarded in the global structure, adding a notion of persistence to the structure. Furthermore, we showcase the results of experiments using data generated by the NetLogo flocking model and from the Starkey project. The Starkey data describes the movement of elk, deer, and cattle. Although there is no ground truth, the experiments show that the trajectory grouping structure is plausible and has the desired effects when changing the essential parameters. Our research provides the first complete study of trajectory group evolvement, including combinatorial, algorithmic, and experimental results.

3.8 Noldus InnovationWorks moving objects research themes

Ben Loke (Noldus Information Technology BV, NL)

In this talk we present an overview of the current Noldus InnovationWorks research themes.

Movement tracking: the AnyTrack project, a tool that should track any number of objects in any context with any type of sensor. With examples of use for the wild-life tracking project E-Track.

Automatic behavior recognition: from behavior recognition for rodents with a video based tracking tool, EthoVision XT, to people tracking, and gesture, pose and hand tracking techniques.

System integration: trend to analyze multi-sensor data from different systems into one integration tool The Observer XT for detailed synchronized analysis.

3.9 M-Atlas: A platform for mobility knowledge discovery

Mirco Nanni (ISTI-CNR – Pisa, IT)

Joint work of Giannotti, Fosca; Nanni, Mirco; Pedreschi, Dino; Pinelli, Fabio; Renso, Chiara; Rinzivillo, Salvatore; Trasarti Roberto
Main reference F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, S. Rinzivillo, R. Trasarti, "Unveiling the complexity of human mobility by querying and mining massive trajectory data," in VLDB Journal (VLDBJ), Special Issue on "Data Management for Mobile Services", Vol. 20, Issue 5, pp. 695–719, 2011.
URL http://dx.doi.org/10.1007/s00778-011-0244-8

Real applications involving the analysis of moving objects require a complex combination of tasks of several different kinds, including data mining (or other analysis) tools. In this talk I presented an analytics platform for supporting such kind of process, thought for — but not limited to — vehicle traffic trajectories, highlighting both the basic analytics tools provided by the system and its ability to build derived analysis functionalities and whole knowledge extraction processes.

3.10 Movement ecology

Ran Nathan (The Hebrew University of Jerusalem, IL)

Main reference R. Nathan, W.M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, P.E. Smouse, "A movement ecology paradigm for unifying organismal movement research," in Proc. of the National Academy of Sciences of the United States of America, Vol. 105, No. 49, pp. 19052–19059, 2008.
URL http://dx.doi.org/10.1073/pnas.0800375105

Understanding and predicting the dynamics of complex ecological and human-dominated systems are best accomplished through the synthesis and integration of information across relevant spatial, temporal and thematic scales. Recent advances in mechanistic modeling, data analysis techniques and tracking technology have enriched our capacity to disentangle the key parameters affecting dispersal, migration, foraging and other movement processes and to accurately quantify movement patterns. In lieu of this favorable background, movement ecology has recently emerged to facilitate the unification of movement research. Movement ecology aims at investigating the explicit links between the internal state, the motion and the navigation capacities of the individual, and the external (biotic and abiotic) environmental factors affecting its movement. Therefore, it provides a natural platform for examining the mechanisms underlying movement processes and patterns, their causes and consequences in changing environments. In this talk I presented the vision and basic principles of the movement ecology approach [1], and illustrated its application to the study of foraging of vultures [2] and navigation of fruit bats [3].

References

1 Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz and P. E. Smouse. 2008. *A movement ecology paradigm for unifying organismal movement research.* Proceedings of the National Academy of Sciences of the United States of America, 105: 19052–19059.
2 Nathan, R., O. Spiegel, S. Fortmann-Roe, R. Harel, M. Wikelski and W. M. Getz. 2012. *Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: Gen-*

eral concepts and tools illustrated for griffon vultures. Journal of Experimental Biology, 215: 986–996.
3 Tsoar, A., R. Nathan, Y. Bartan, A. Vyssotski, G. Dell'Omo and N. Ulanovsky. 2011. *Large-scale navigational map in a mammal.* Proceedings of the National Academy of Sciences of the United States of America, 108: E718–E724.

3.11 Workshops, data challenges and some thoughts

Ross Purves (University of Zurich)

In my talk I presented my take on the process of moving from our discussions in 2010 at Dagstuhl, to a data challenge on gull data at the Leiden Centre in 2011 and a workshop on Progress in Analysis of Real Movement Data in autumn 2012. I emphasized the utility of challenges in bringing us together to discuss common themes, even if direct results in terms of publications are less obvious. However, links made have resulted in new publications and it is clear that data challenges have great utility. I concluded by suggesting that data challenges should meet three criteria:

1. Be realistic (i.e. use real data and ask real questions)
2. Go beyond simple primitives
3. Allow us to compare high level methods (for example analysis not query/ derivation)

3.12 Semantic enrichment of trajectory data

Chiara Renso (ISTI-CNR – Pisa, IT)

Joint work of Renso, Chiara; Vania Bogorny
Main reference V. Bogorny, C. Renso, A. Ribeiro de Aquino, F. de Lucca Siqueira, L. Otavio Alvares, "CONSTAnT – A Conceptual Data Model for Semantic Trajectories of Moving Objects," in Transaction in GIS, 2013, to appear.

Mobility of individuals is a complex phenomena. The last few years has seen a growing interest in representing analysing and understanding the mobility of individuals, basically relying on the concept of a trajectory to represent the history of evolving locations. However, the raw trajectory representation as a sequence of georeferenced points and the paired analysis techniques so far do not produce satisfying results in terms of understanding and proper deployment for the final application. The proposal of representing trajectories in a more meaningful way, namely, semantic trajectories, is a step towards this ambitious goal. We propose a conceptual model called CONSTAnT for semantic trajectory representation that enhances the well known "stop and move" including many semantic facets that should be combined with raw trajectories. Analysis methods aimed at (1) inferring semantics from data to annotate trajectories and (2) finding patterns from semantically enriched trajectories have therefore to be investigated.

3.13 Alpha-visibility

Jörg-Rüdiger Sack (Carleton University, CA)

Joint work of Ghodsi, Mohammad; Maheshwari, Anil; Nouri-Baygi, Mostafa; Zarrabi-Zadeh, Hamid

We study a new class of visibility problems based on the notion of α-visibility. A segment t is said to be α-visible from a point p, if p can see t with an angle at least α; that is, if there exists an empty triangle with one vertex at p and side opposite to p on t such that the angle at p equals α. This notion of visibility appears to be natural and has the potential to be of practical value. For example, the smallest angle we can observe directly is determined by the ratio of the wavelength of light to the diameter of the eye's pupil, which is lower bounded by a constant 10^{-4} radians. Also, some optical/digital imaging devices have similar limitations, quantified by their resolutions.

Our α-visibility model is inspired by those limitations, and may provide a more realistic alternative to the classical visibility models studied in the literature. In this talk, we present some of the first results for several variants of the point/segment visibility problems in the α-visibility model. We discuss open problems and anticipate that this model will be extended in further studies.

3.14 Towards an interdisciplinary analysis platform: the benefits of R as an open source analysis framework

Kamran Safi (MPI für Ornithologie, DE)

As ecology is heading towards BigData with rapid speed, computational ecology has emerged as an interdisciplinary discipline. But, besides the problems of storage and versioning that large data sets bring, the issue of efficient data analysis and progress in methodology needs addressing. Here, I present MOVE (http://cran.r-project.org/web/packages/move/index.html) an R library that provides an interface with Movebank, allowing simple data access from R. This library should allow the future to make advances in development of methods in movement analysis more straight forward. Finally, I presented an outlook of ICARUS an initiative to track animals from space beginning in 2014 on the international space station. ICARUS will in the near future provide orders of magnitude more data and with that challenges in data storage and analysis.

3.15 Data challenge: Learning Oyster catcher behavior

Daniel Weiskopf (Universität Stuttgart, DE)

In this talk, the results of our work on the Data Challenge are presented. With the Data Challenge, data on the movement of oyster catchers was provided to the participants before the seminar. We built a processing pipeline consisting of some custom code for data cleaning and preprocessing and subsequent supervised learning with RapidMiner and Weka. First, we reproduced the results previously reported in a paper on the Data Challenge data set and then compared different classification methods provided by RapidMiner. For our choice of data cleaning, W-SOM and Weka-LMT performed best, and slightly better than the decision tree of the original paper. Subsequently, we analyzed the time series of the recorded data to extend the feature vector. We chose to include additional information from the accelerometer. The extended feature vector led to further slight improvement of classification results. Finally, we briefly report on attempts to apply other visual analytics techniques to the Data Challenge: scatter/gather browsing of trajectories [1] and inter-active learning and visualization of classifiers [2].

References

1 B. Höferlin, M. Höferlin, D. Weiskopf, G. Heidemann. *Interactive Schematic Summaries for Exploration of Surveillance Video*. Proceedings of ICMR '11, 2011, Article No. 9. DOI: 10.1145/1991996.1992005.

2 B. Höferlin, R. Netzel, M. Höferlin, D. Weiskopf, G. Heidemann. *InterActive Learning of Ad-Hoc Classifiers for Video Visual Analytics*. Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST), 2012.

3.16 Visual analytics of eye-tracking trajectories

Daniel Weiskopf (Universität Stuttgart, DE)

Joint work of Andrienko, Gennady; Andrienko, Natalia; Burch, Michael; Weiskopf, Daniel
Main reference G. Andrienko, N. Andrienko, M. Burch, D. Weiskopf, "Visual Analytics Methodology for Eye Movement Studies," in IEEE Transactions on Visualization and Computer Graphics, 18(12):2889–2898, 2012.
URL http://dx.doi.org/10.1109/TVCG.2012.276

In this talk, methods for the spatio-temporal visual analytics of eye- tracking data are discussed [1]. The basic approach is to apply visual analytics methods previously developed for geospatial analysis. However, due to the characteristics of eye-tracking data, not all of the geospatial techniques are applicable. Appropriate analysis techniques are illustrated by running through an eye-tracking example data set acquired in the context of evaluating different type of tree visualization techniques [2, 3]. Finally, challenges and open questions of spatiotemporal analysis of eye- tracking data are discussed.

References

1 G. Andrienko, N. Andrienko, M. Burch, D. Weiskopf. *Visual Analytics Methodology for Eye Movement Studies*. IEEE Transactions on Visualization and Computer Graphics 18(12) 2889-2898, 2012. DOI: 10.1109/TVCG.2012.276

2 M. Burch, N. Konevtsova, J. Heinrich, M. Höferlin, D. Weiskopf. *Evaluation of Traditional, Orthogonal, and Radial Tree Diagrams by an Eye Tracking Study*. IEEE Transactions on Visualization and Computer Graphics 17(12):2440-2448 (2011). DOI: 10.1109/TVCG.2011.193

3 M. Burch, G. Andrienko, N. Andrienko, M. Höferlin, M. Raschke, D. Weiskopf. *Visual Task Solution Strategies in Tree Diagrams*. Proceedings of IEEE PacificVis 2013.

4 Working Groups

4.1 Data Challenge

As part of this Dagstuhl seminar, a data challenge was prepared. The challenge was on purpose kept small. It mainly aimed at familiarizing the participants with the topics of the seminar, establishing a harmonized language amongst a very diverse group of participants and ultimately shortening the warm-up phase at the seminar. The data challenge consisted of two distinct parts: one aimed at an analysis question and another aimed at interactive visualization. Both parts used the same data set of bird movement which has been published and described in [1]. The goal of the analysis-part of the data challenge was to find features and algorithms for segmentation, shape and movement analysis which could be successful in classifying foraging behavior. The goal of the visualization-part was to develop concepts and tools to visually explore differences in (model-predicted or observations of) bird movement and behavior. The task description and data were made available six weeks before the Dagstuhl seminar and all seminar participants were invited to join.

4.1.1 Results on the analysis challenge

Kevin Buchin, Maike Buchin, Rob Weibel, and Daniel Weiskopf

Rob Weibel and Daniel Weiskopf considered point-based (not considering segment shape or geographical relationships) classifiers and clustering techniques. A range of classifiers were compared under a variety of cross-validation schemes. The conclusion of this work was that it is worthwhile to cross-compare and integrate the various successful classifiers by the different groups and also to define features at different resolutions (in this way encapsulating different levels of spatio-temporal heterogeneity). The tools used for this part in the data challenge comprised RapidMiner (http://sourceforge.net/projects/rapidminer/), Weka (http://www.cs.waikato.ac.nz/ml/weka/) and some tailor-made C++ and R code.

Kevin and Maike Buchin considered methods to include temporal relations into the classification problem. Here a naive Bayes classifier was compared to a hidden Markov model. It appeared that considering time-dependencies did improve the correct classification, without even involving additional features. Hence considering time dependencies seems to be a fruitful avenue in enhancing the correct classification of movement data. This work has been implemented in R.

4.1.2 Results on the visualization challenge

Susanne Bleisch and Aidan Slingsby

Aidan Slingsby demonstrated a visualization interface. It visualizes the location of individuals (on top of Bing satellite imagery), with highlighting according to behavior and details per gps-point. Next to the geographic display the interface contains a timeline, showing the habitat, modeled behavior and observed behavior in parallel. And finally, the interface also shows barcharts of time spent in each habitat, modeled behavior and observed behavior per individual or jointly for the several individuals. All three panels are linked and the geographical display as well as the time-line can be used for zooming, panning or selecting a subset of the data.

Susanne Bleisch demonstrated an interactive visualization of the challenge data, comprising a set of spatial views with coloring by individual and subsetting by behavior. The data were shown at different scale levels and shown against a background that did show wrongly classified results. With both tools, an overview of the classification errors (e.g. the distribution over space or time, the individuals where errors are most prevalent) were quickly obtained; both tools were implemented in Java.

The contributions to this part of the data challenge demonstrated that interactive data visualization is of great value to effectively explore the properties of both movement data and prediction errors of that data. However, it also became clear that it is difficult to use generic tools to make in-depth analysis.

4.1.3 Future directions and panel discussion

The discussion started with a review and summary of previous activities related to data challenges in the area. Initial ideas about data challenges and benchmarking in movement analysis came out of the predecessor Dagstuhl seminars Nr. 08451 in 2008 and Nr. 10491 in 2010. A similar yet different group of experts at the GIScience 2010 workshop on Movement Pattern Analysis in Zurich further stressed the need for community initiatives facilitating a direct comparison of methods. Finally, at seminar Nr. 10491 in 2010 a group of researchers around Bettina Speckmann, Emiel van Loon and Ross Purves initiated a follow-up event hosted by the Lorentz Center in Leiden, NL [2].

In the Lorentz data challenge, the task focused on calculating a number of simple movement statistics that are generally reported in animal ecology (like trip duration, distance traveled and average speed). It appeared that in spite of the detailed guidelines and the relatively complete data sets, relatively large differences arose in the calculated summary statistics by the different participants. The conclusion of this data challenge was that it is dangerous to conduct meta-analysis on studies that only report summary statistics of movement; at the same time the need for better definitions or generally accepted and robust algorithms to compute the required statistics was emphasized. Furthermore, the importance of storing complete analysis workflows (i.e. including all data pre-processing steps) to enable reproducible research was demonstrated through this data challenge.

A workshop on Progress in Movement Analysis hosted at University of Zurich took a related yet different approach. This workshop included explicitly only reports of methodological contributions when they emerged multidisciplinary collaborations with domain experts (for example bringing together both ecological and algorithmic expertise). Integrating the experiences made at Leiden, at PMA10 and MPA12, and the mini-challenge stages at the current seminar, the discussion then turned to the contributions such data challenges might offer the community.

The group quickly agreed that the challenges led to moderate progress but fell short of achieving the hoped for general comparability of methods. Several reasons were given for this shortcoming. First, comparing to, for instance, the VisMaster challenge in the rather coherent information visualization community, the movement analysis community is rather diverse featuring researchers with very different research cultures. Second, it was stressed that all initiatives targeting methodological progress are organized and hosted by methods experts, repeatedly struggling attracting large numbers of domain experts and hence often falling short of being truly multidisciplinary.

The group then discussed pros and cons of data challenges, potentially aiming at establishing a long-lasting series of challenges co-hosted with a perennial conference.

Pros:
- Challenges are a good way to get students involved
- Challenges bring knowledge together and can promote the exchange of knowledge on how to apply established techniques

Cons:
- Challenges rather encourage practice of established methods, rather than innovation. Authors of innovative methods will rather publish their work somewhere else with more prestige.

The group concluded with the insight that even though the so far staged challenges significantly contributed to more transparency in methodological movement analysis and have propelled interdisciplinary work, the potential is now exploited as the small group has had ample opportunities for exchange. Before organizing new challenges, the community agreed on making more accessible what has already been achieved (Leiden Challenge). This could mean, (a) putting all results and code in the public domain, or (b) putting effort into developing techniques and tools to a high standard such that is can be easily reused (for example as *R*-libraries). Future challenges could receive more attention when co-hosted with an established conference (GIScience, ACM SIG Spatial) and when offering a more formal incentive (access to really interesting data sets). Finally, the group discussed funding opportunities for larger scale challenges. Initial ideas came from the movement ecology community, around possible challenges with multi-sensor, combined location-acceleration data.

References

1 J. Shamoun-Baranes, R. Bom, E.E. van Loon, B.J. Ens, K. Oosterbeek, W. Bouten. *From sensor data to animal behaviour: an oystercatcher example*. PLoS One, 7(5), e37997, 2012. http://dx.doi.org/10.1371/journal.pone.0037997

2 J. Shamoun-Baranes, E.E. van Loon , R.S. Purves, B. Speckmann, D. Weiskopf, C.J. Camphuysen. *Analysis and visualization of animal movement*. Biol Lett 8:6–9, 2012. http://dx.doi.org/10.1098/rsbl.2011.0764

4.2 Defining the theoretical core of movement science

A first series of working group sessions aimed in parallel for a definition of the theoretical core of movement science. The groups addressed the problem from three different perspectives. The first working group assembled an inclusive list of research questions emerging an eclectic set of domains studying movement (bottom up). The second group took a top-down perspective and worked on a table of contents for a planned inclusive textbook on movement analysis. The last working group studied a concrete use case, aiming at identifying the fundamental questions and challenges through a systematic walk through the use case.

Rather doing this than assembling yet another research agenda. List/TOC can serve as a starting point for potential review articles.

4.2.1 Collection of application driven questions

Andrea Kölzsch, Ran Nathan, Ross Purves, Roeland Scheepens, Aidan Slingsby, Bettina Speckmann, Egemen Tanin, Emiel Van Loon, Georgina Wilcox, and Jo Wood

The working group assembled a bank of classified application questions. The list may help in seeing how methods related to moving objects could be transferred from one domain to another, or to see if there are certain types of recurring questions for which there are not sufficient methods to address them. What started at the seminar is now an ongoing initiative where researchers are still encouraged to maintain and further develop the list and classification. The only conditions for questions is that they must be related to application domains and not directly about methods or specific data sets.

4.2.2 Table of contents of a Movement Science book

Susanne Bleisch, Maike Buchin, Györö Gidófalvi, Joachim Gudmundsson, Patrick Laube, Ben Loke, Mirco Nanni, Chiara Renso, Rodrigo I. Silveira, Egemen Tanin, and Rob Weibel

This group aimed at narrowing down and defining the core of movement science by assembling a table of content for a planned text book on the subject. The working group attracted a diverse set of researchers from the major areas advancing movement science, including geographical information science, computational geometry, computer science and database research. For the below listed sections, the group discussed and assembled the relevant methods according the following criteria: (i) Gather methods the group members have directly worked on, (ii) complement this list with relevant methods the group members have not directly worked on but consider relevant, (iii) exclude methods considered outside the focus of the planned book.

The characteristic of every method or methods category to be included were specified according to the following roster. This exercise help distinguishing and categorizing methods and finally positioning these methods in the table of contents.

- **Why** and **what for** would one use that method?
- **What** is processed? Does the method process fixes (observation points), subtrajectories (segments), trajectories, single moving point objects, or groups of moving point objects?
- **How** does the method actually work? What is the mathematical, statistical or algorithmic underpinning of the method?
- Are there unifying **criteria** that help categorizing the methods?

The tentative core of movement science mirrored in an table of contents for a text book are structured as follows.

1. **Conceptual Models of Movement.** Conceptual data models for movement space (Euclidean vs. network vs. Voronoi spaces), models for trajectories, other models for representing movement.

2. **Data sources, preprocessing, cleaning.** Data quality and uncertainty, map matching and filtering. Clearly, many preprocessing methods can be considered proper analysis methods.
3. **Management of movement data**. Storing, databases, moving object databases, indexing, querying, streams.
4. **Analysis.** Deriving movement parameters, outliers, similarity, clustering, segmentation, sequence analysis, movement patterns, trajectory simplification, smoothing, *aggregation*, sampling issues, movement simulation, visualization and mapping, visual analytics, semantic annotation, data integration.
5. **Evaluation.** Sensitivity experiments, Monte Carlo simulations, well-documented case studies and benchmarking.
6. **Tools.** Practical advice, tools that do a fraction, include a matrix techniques x tools, e.g. MoveBank.

The example for the task *aggregation* shall illustrate the chosen procedure. **Why** is aggregation of trajectory data required? Aggregation copes with information overload and allows an overview of large trajectory data sources, it also compensates for location uncertainty of individual fixes. **What** is processed? Typical aggregation methods act on fixes of trajectory or fixes of many trajectories of always the same or a group of moving point objects. **How** do aggregations methods work? Approaches include the selection of a representative trajectory, a median trajectory, flow maps, convex hulls, kernel density estimations, heat maps, and 3D shapes and density maps. What are **criteria** useful for a categorization? Methods can be separated depending on the nature of the results (entities or fields?), interrelations can be found to *clustering.*

Instead of listing a plethora of methods, the group worked out the contribution of the present disciplines (as above geographical information science, computational geometry, computer science and database research). The main contribution of the group can be seen in putting together and structuring the methods with respect to specific ways of addressing a problem and generic properties of the discussed methods. Clearly, there are many different ways of organizing methods into sections representing categories or steps along the analytical process. It is furthermore little surprising that the group identified many interrelations between methods. For instance, *aggregation* can be carried out based on *clustering* methods, or *segmentation* of trajectories can be done using *clustering* of fixes. Nevertheless, the discussion converged towards a conclusive categorized table of contents.

4.2.3 Interactive annotation: interfacing expert opinion and automated annotation tools

Kevin Buchin, Urska Demsar, Andrea Koelzsch, Kamran Safi, Emiel van Loon, and Georgina Wilcox

Movement analysis often requires to annotate trajectories with additional information. In movement ecology, for example, different stages of movement, characterized by different types of behavior, inform ecologists about the amount of time spent on different activities. Such activity budgets, including the precise onset and offset, are an essential part of life history strategies and can provide valuable information necessary to understand selective processes and conservation of populations. However, with the advent of cheap GPS devices and the ability to gather information with better spatial and temporal resolution, the method of manually annotating trajectories faces a logistic problem. It is therefore crucial to reconstruct the decisions experts make to annotate trajectories and automatically assign behavior to trajectories with a high reliability. Such a procedure requires an interface between the expert who is doing the annotation and the computer that evaluates these annotations to eventually provide automated annotation tools.

The group discussed the value of a thorough understanding of the workflow required for an efficient interface. Obtaining the best results both from domain experts and the chosen software is clearly a function of ease of use and information provided. Thus, increasing the willingness of experts to devote time to train a supervised machine learning algorithm requires some interaction with a system that needs to provide the necessary information in an intuitive way. At the same time, such an interface would need to provide the experts with clear definitions of behavior that they should assign and potentially involve the iterative presentation of important layers after previous assessment in addition to information such as satellite maps and topography or landuse.

Investigating the inter- and intra-expert variance would allow researchers to re-evaluate the definitions of behavioral classes and reduce ambiguities in their definitions. It would also allow a better understanding of the processes and decisions that experts take in the process of decision making. Also a comparison of mismatches between experts (or the median expert assessment) with the machine based assessment will allow to better understand the necessary information, that both the machine and the experts would require, for a more reliable assessment of behavioral patterns. Ultimately, the iterative annotation of trajectories is a problem that has implications well beyond annotation of trajectories for the assessment of time budgets. The principles can be applied to a wide variety of data that require annotation and where the amount of the data asks for an automated assessment, such as acceleration data or other types of movement data.

The conclusions of the group were to pursue a two step approach. In a first step a study would be conducted in the field of movement ecology where experts are asked to annotate behavioral stages to a set of water fowl data. These expert opinions would then be modeled using different statistical and machine learning procedures with previously environmentally enriched data to identify the important contextual layers that correlate with the different behavioral stages. This study can be published as such in an ecological frame work, however, along this project the workflow and the processes involved in the course of this project would be closely monitored. The experts who would do the annotations would be questioned as to their experience and suggestions in improving the workflow for them, in terms of visualization and feedbacks from the machines. This consideration of the workflow and reported suggestions for requirements would allow the group in a second step to formulate well defined challenge for computer science groups to provide tools and methods that substantially ease the process and provide efficiency gain in iterative annotation. Such a tool can then provide help bridging the gaps between man and machine where with every new annotated trajectory new and better models could be generated allowing to replicate the decisions that experts take. It would be also possible with such a system to investigate differences between experts and lay audience and the potential for crowd sourcing evaluated.

Participants

- Susanne Bleisch
The University of Melbourne, AU
- Kevin Buchin
TU Eindhoven, NL
- Maike Buchin
TU Eindhoven, NL
- Urska Demsar
School of Geography & Geosciences, GB
- Matt Duckham
The University of Melbourne, AU
- Gyözö Gidófalvi
KTH – Stockholm, SE
- Joachim Gudmundsson
The University of Sydney, AU
- Bin Jiang
University of Gävle, SE
- Andrea Kölzsch
NIOO-KNAW, NL
- Patrick Laube
Universität Zürich, CH
- Ben Loke
Noldus Information Technology BV, NL
- Mirco Nanni
ISTI-CNR – Pisa, IT
- Ran Nathan
The Hebrew University of Jerusalem, IL
- Ross Purves
Universität Zürich, CH
- Chiara Renso
ISTI-CNR – Pisa, IT
- Jörg-Rüdiger Sack
Carleton Univ. – Ottawa, CA
- Kamran Safi
MPI für Ornithologie, DE
- Roeland Scheepens
TU Eindhoven, NL
- Rodrigo I. Silveira
UPC – Barcelona, ES
- Aidan Slingsby
City University – London, GB
- Bettina Speckmann
TU Eindhoven, NL
- Egemen Tanin
The University of Melbourne, AU
- Marc van Kreveld
Utrecht University, NL
- Emiel Van Loon
University of Amsterdam, NL
- Robert Weibel
Universität Zürich, CH
- Daniel Weiskopf
Universität Stuttgart, DE
- Georgina Wilcox
The University of Sydney, AU
- Jo Wood
City University – London, GB