MSc ENR IUNR ZHAW
n.MA.ENR.PTED

Specialized MSc in GIScience
GEO 880

T3 – Segmentation & Similarity

Zurich University
of Applied Sciences

Patrick Laube
Geoinformatics Research Group

Institute of Natural Resource Sciences
ZHAW Zurich University of Applied Sciences
Grüental, Postfach, CH-8820 Wädenswil

patrick.laube@zhaw.ch

University of Zurich[UZH]

Patrick Laube
Privatdozent Geographic Information Science

Department of Geography
The University of Zürich
Winterthurerstrasse 190, 8057 Zürich

Thread trailing of turtles: methods for evaluating spatial movements and pathway structure

Dennis L. Claussen, Michael S. Finkler, and Meghan M. Smith

Abstract: Although many authors have used thread trailing to monitor movements of turtles and other vertebrates, most have evaluated only a portion of the information contained in these pathways. We describe ways of extracting information from thread trails by determining length of trail, mean distance moved per unit time, net displacement, area encompassed by the trail, mean turning angle, and mean vector length. We further describe and compare four measures of trail sinuosity, the ratio of greatest distance to length of path (d/L), the square root of the ratio of area to path length ($\sqrt{A/L}$), fractal dimension (D), and sinuosity (S); we then applied these methods to the trails of a population of ornate box turtles (*Terrapene ornata*) from the Nebraska Sand Hills. Though preliminary, these data suggest that vegetation cover affects the sinuosity of turtle trails (based on D), that temperature affects the mean distance moved per day, that *T. ornata* ordinarily show little or no directional bias, and that injury can affect distance moved, trail sinuosity, and mean turning angle. Among these measures of trail sinuosity, D and S are highly correlated and provide similar information. Though less closely correlated with D and S, or with each other, $\sqrt{A/L}$ and d/L also reflect trail sinuosity; of these indices, however, d/L best reflects trail directionality. This analysis suggests that a more comprehensive evaluation of pathways at this fine scale will provide new insights into the routine movements of animals.

Résumé : Plusieurs auteurs ont utilisé la méthode du fil attaché pour suivre le trajet de tortues ou d'autres vertébrés, mais la plupart n'ont évalué qu'une partie de l'information contenue dans les itinéraires des déplacements. On trouvera ici la description de moyens de tirer des informations de données recueillies par la méthode du fil attaché, longueur du trajet, distance moyenne parcourue par unité de temps, déplacements nets, surface couverte par le trajet, angle moyen de virage et longueur moyenne du vecteur. Nous décrivons et comparons également quatre mesures de la sinuosité du trajet (rapport entre la distance maximale et la longueur du trajet (d/L), rapport entre la racine carrée de la surface et la longueur du trajet ($\sqrt{A/L}$), dimension fractale (D) et sinuosité (S)); nous avons utilisé ces méthodes dans l'étude d'une population de tortues *Terrapene ornata* des Sand Hills du Nebraska. Bien que préliminaires, les données indiquent que la couverture végétale affecte la sinuosité des trajets empruntés par les tortues (d'après la valeur de D), que la température affecte la distance moyenne parcourue en une journée, que les tortues ne montrent ordinairement pas de préférence particulière pour une direction donnée et enfin, que les blessures peuvent influencer la distance parcourue, la sinuosité du parcours et l'angle moyen de virage. Parmi les mesures de la sinuosité du parcours, D et S ont une forte corrélation et donnent les mêmes informations. Bien que reliées moins fortement à D ou S ou l'une à l'autre, les mesures $\sqrt{A/L}$ et d/L reflètent également la sinuosité du parcours; cependant, parmi tous ces indices, c'est la mesure d/L qui reflète le mieux la tendance du trajet vers une direction donnée. Cette analyse indique qu'une évaluation plus exhaustive des trajets à petite échelle peut élargir nos connaissances des déplacements habituels des animaux.
[Traduit par la Rédaction]

Introduction

We cannot fully appreciate the geographic distribution of a motile animal population until we have documented the movement patterns of its individual members. Countless studies, mostly focusing on mammals, birds, or insects, have investigated the areas within which animals move (i.e., their home ranges), the dispersal of animals, or their migrations (for example, see the review by Swingland and Greenwood 1983). Nevertheless, we know very little about animal movements during routine activities (Hailey 1989).

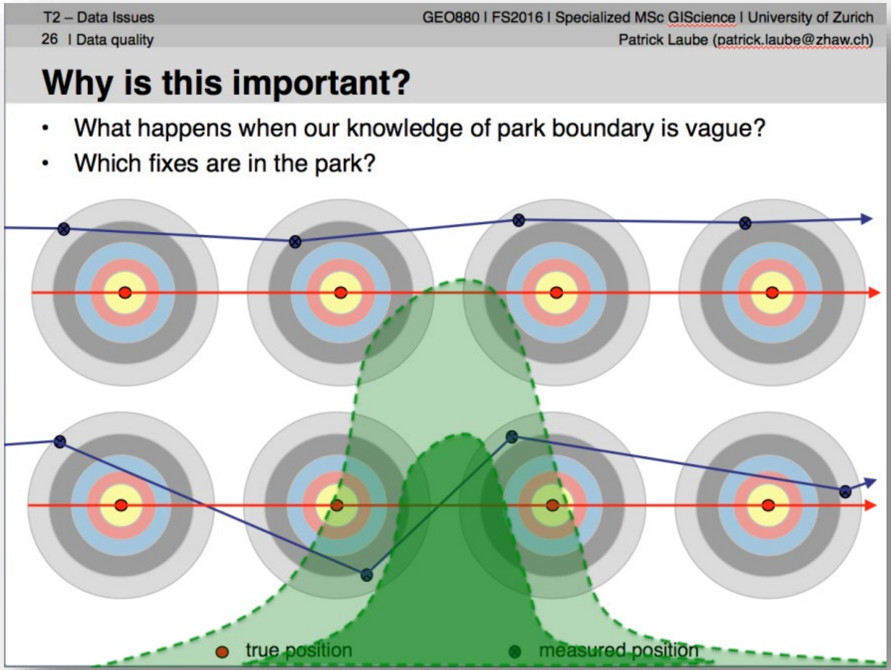
Most studies of the movements of animals have been based on mark and recapture methods, radiotelemetry, or, more rarely, thread trailing (also known as the spool and line method). The mark and recapture method is low-cost and useful for home-range investigations and for monitoring population dynamics, but provides no information about the path taken by an animal between point A and point B and little useful information about the velocity of animal movements. Also, the rate of recapture is low for many species and the increments between successive captures are often highly variable or even unpredictable. Radiotelemetry is a high-cost method that allows animals to be located at any time. This method typically can provide many more location points for each individual than the mark and recapture approach, and can be used to monitor long-term and large-scale movements of animals. It can provide estimates of mean minimum velocity between successive location fixes. However, except for the most sophisticated tracking systems, fine-scale information about the specific paths taken by individual animals is not available (Hailey 1989). Thread trailing, a low-cost method developed by Breder (1927) for

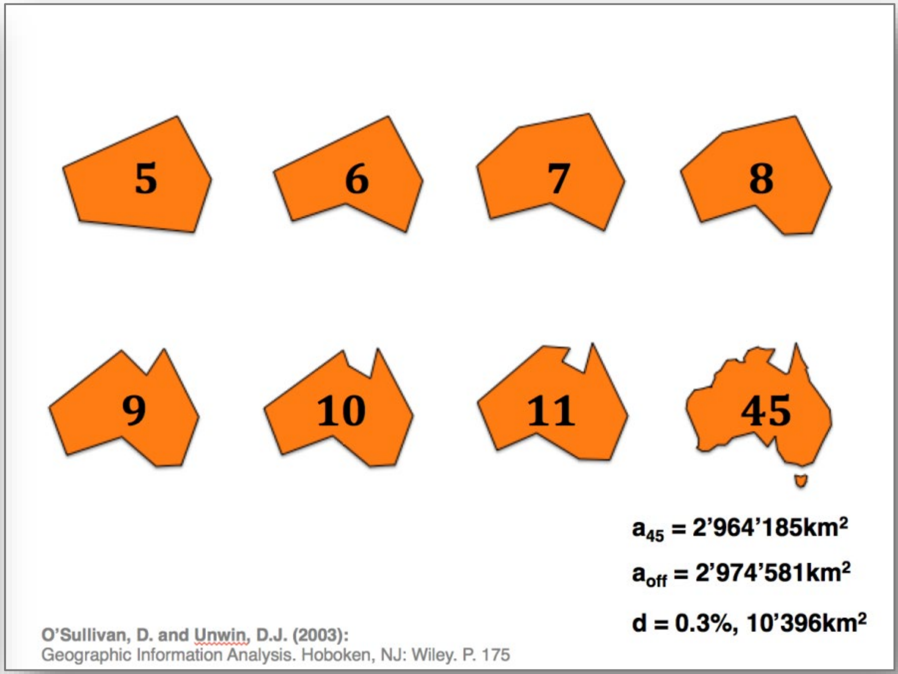
Received January 28, 1997. Accepted July 3, 1997.

D.L. Claussen,[1] M.S. Finkler, and M.M. Smith. Department of Zoology, Miami University, Oxford, OH 45056, U.S.A.

[1] Author to whom all correspondence should be addressed (e-mail: claussen@muohio.edu).

Data quality, scale, map matching





Today’s schedule

Approx. times	What	Form
08:00 – 08:40	Theory recap and Q&A	Patrick
	Break	
08:50 – 10:00	**Gr3** Similarity measures	Group work
	Break	
10:15 – 11:50	**E3** Segmentation & Similarity	R exercises
11:50 – 12:00	Wrap-up	

Today‘s learning objectives

- ✓ You understand the crucial role of segmentation in CMA, especially for separating stops from moves
- ✓ You can explain at least three different trajectory similarity measures in your own words and by drawing simple sketches.
- ✓ You can assess the suitability of a trajectory similarity measure for different types of given movement data.
- ✓ You build up R skills for segmenting trajectories and computing similarities

Reading assignments

- **R3.1** R3.1 Tao et al., (2021). A comparative analysis of trajectory similarity measures, Int. J. GISs & Remote Sensing
- **R3.2** Toohey, K., & Duckham, M. (2015). Trajectory similarity measures. SIGSPATIAL Special, 7(1), 43-50.

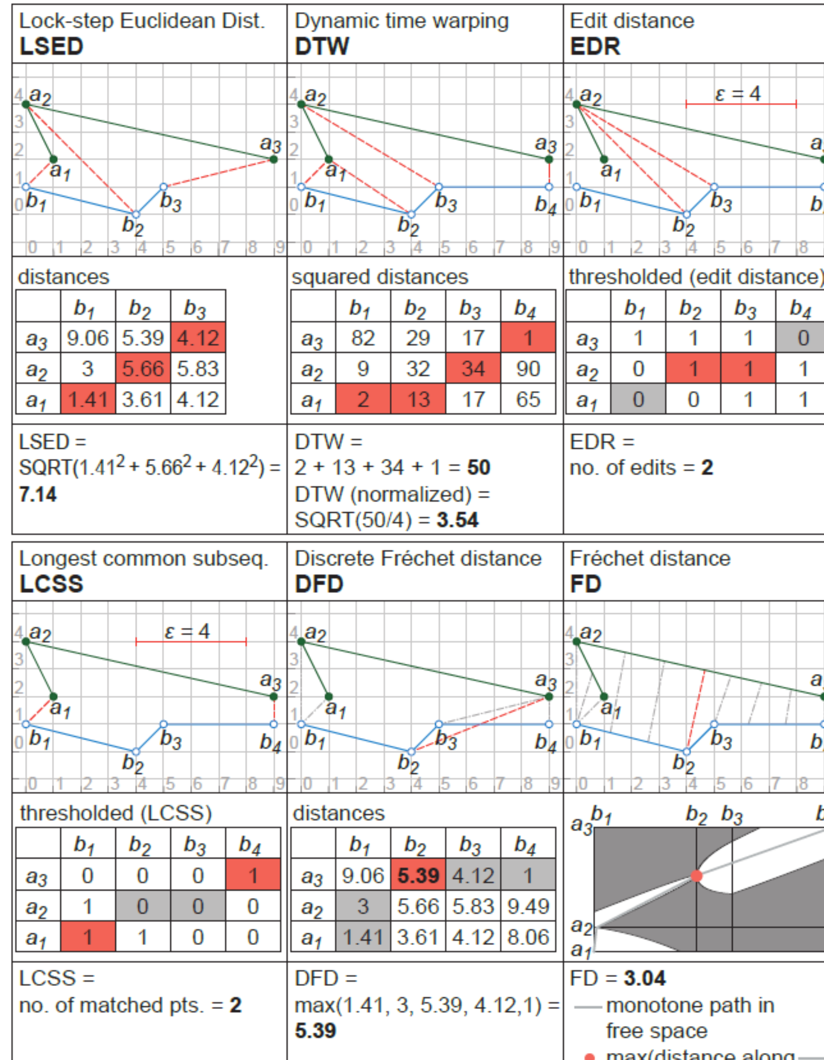


Table of contents: Trajectory operations

1. **Segmentation**
   - Stops and moves
   - Semantic annotation
2. Similarity
   - Why compute trajectory similarity?
   - Why is it difficult?

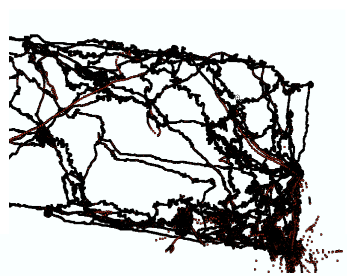
Segmentation

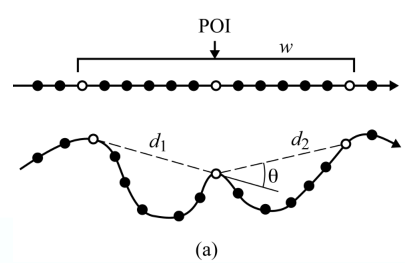
- *Segmentation* refers to the process
  - partitioning movement data into multiple segments, with the goal of **simplifying** or changing the representation of the trajectory into something that is more meaningful or easier to analyze.
  - partitioning a trajectory into a (typically small) number of pieces, where the obtained segments have **uniform characteristics**.
- **preprocessing step**, aiming at reducing noise and condensing the signal for a given analytical task, or
- …**main analysis** task, when the goal is the identification of moves or corridors (E4 Segmentation)

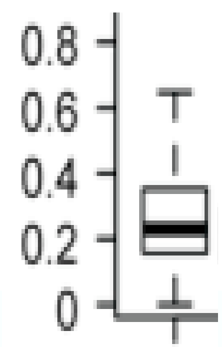
Segmentation 101: Stops & Moves

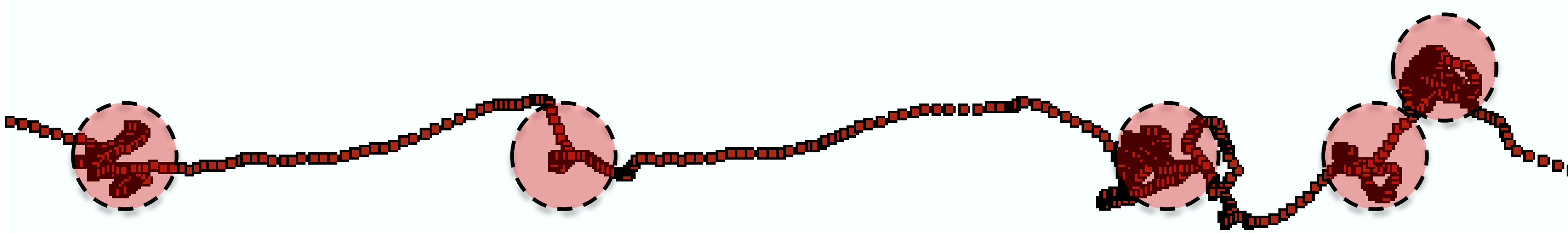
- cows sit around a lot, “pseudo movement”
- From 970000 fixes down to 390000 fixes

?







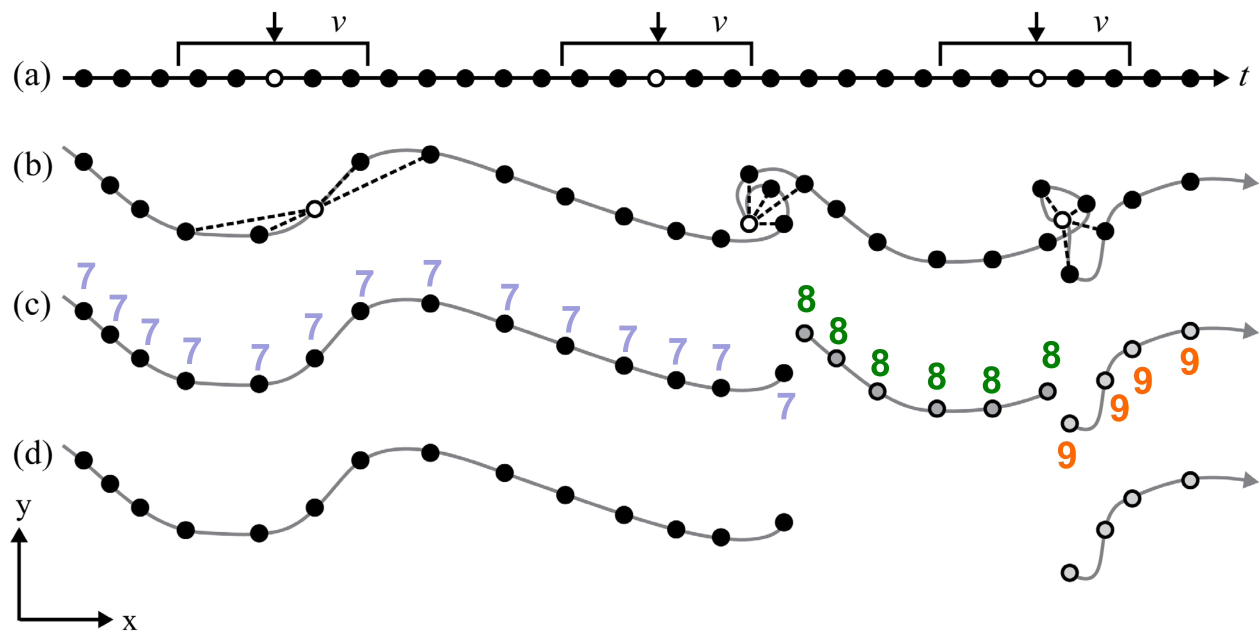


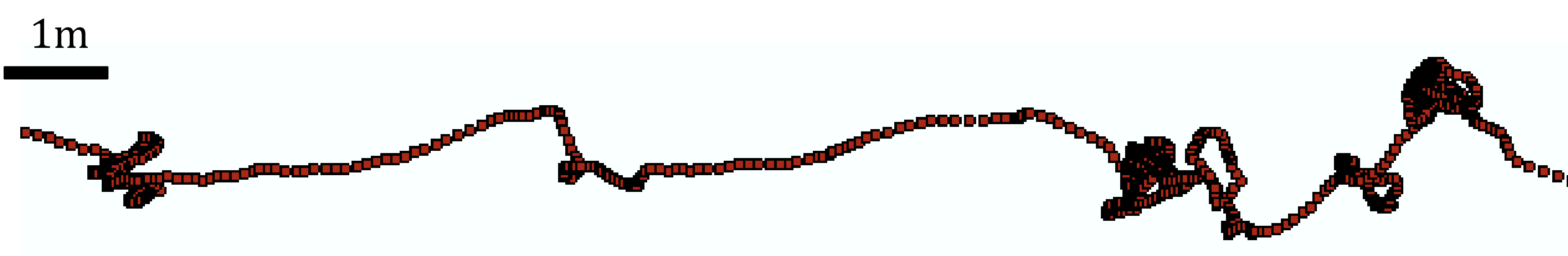
➲ Laube, P., Purves, R. (2011). How fast is a cow? Cross-scale Analysis of Movement Data, *Transaction in GIS*. 15(3), 401-418

Segmentation 101: Stops & Moves

- Segmentation: ave.distance *d* = 3m, temporal window *v* = 300s
- Filtering: 60s

Distance calculation from one point
stop will be removed --> NA
segments get grouped by





➲ Laube, P., Purves, R. (2011). How fast is a cow? Cross-scale Analysis of Movement Data, *Transaction in GIS*. 15(3), 401-418

131 shades of grey

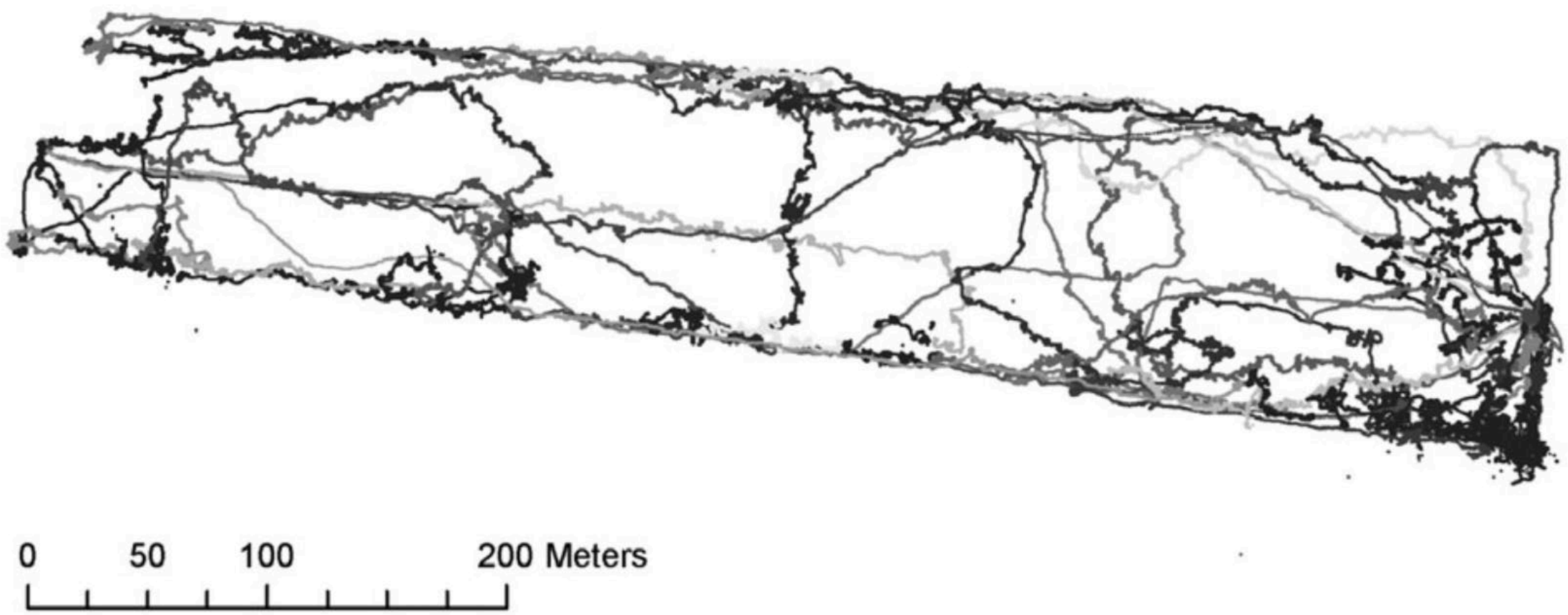
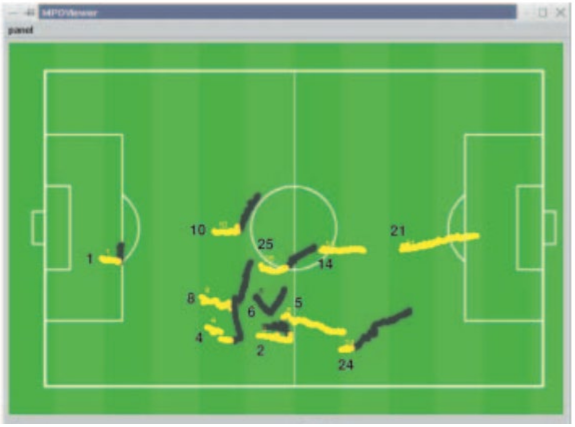
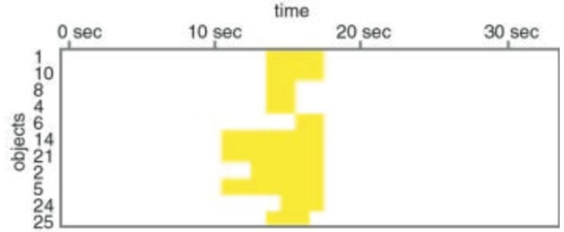


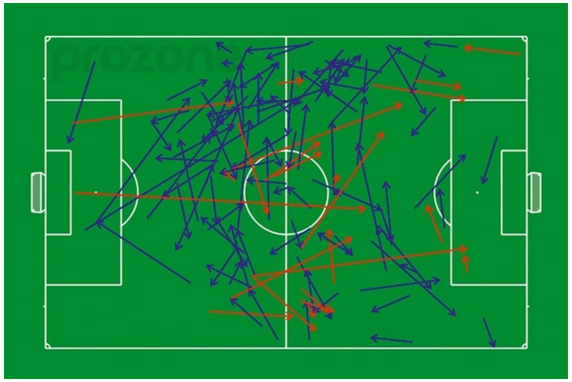
Figure 5 Trajectory of cow (#404020) after filtering and segmentation, each of the 131 unique segments has a unique shade of grey

Alternative segmentation criteria

- Trajectory segmentation can be based on
  - Stops & moves
  - Derived **movement properties** (speed, azimuth, e.g. REMO framework)
  - Geometry, shape of the trajectories

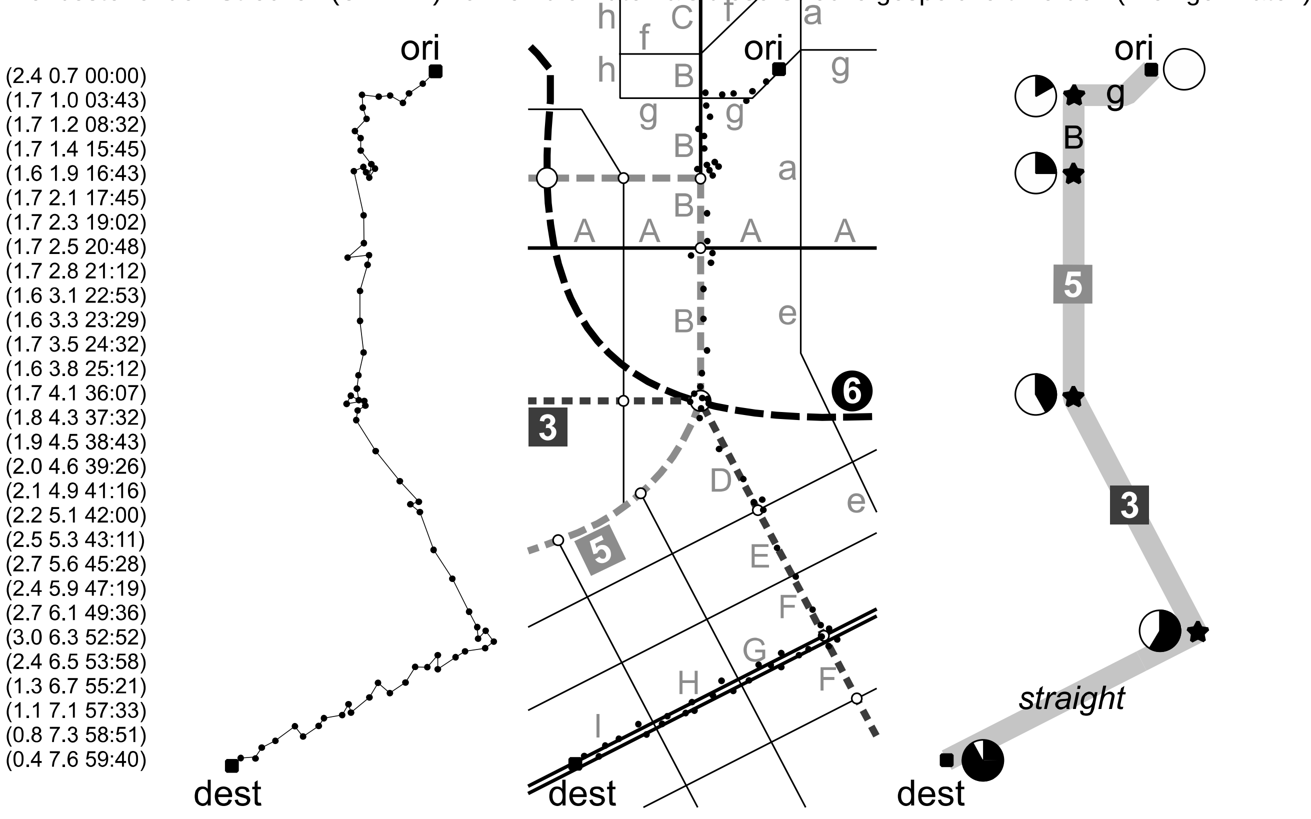






Semantic annotation of trajectories

Bei bestehenden Strecken (UBAHN) können die Daten als diese Strecke gespeichert werden (Weniger Daten)



➲ Richter, K. F., Schmid, F., & Laube, P. (2012). Semantic trajectory compression: Representing urban movement in a nutshell. *Journal of Spatial Information Science*, *2012*(4), 3-30.

Semantic annotation of trajectories

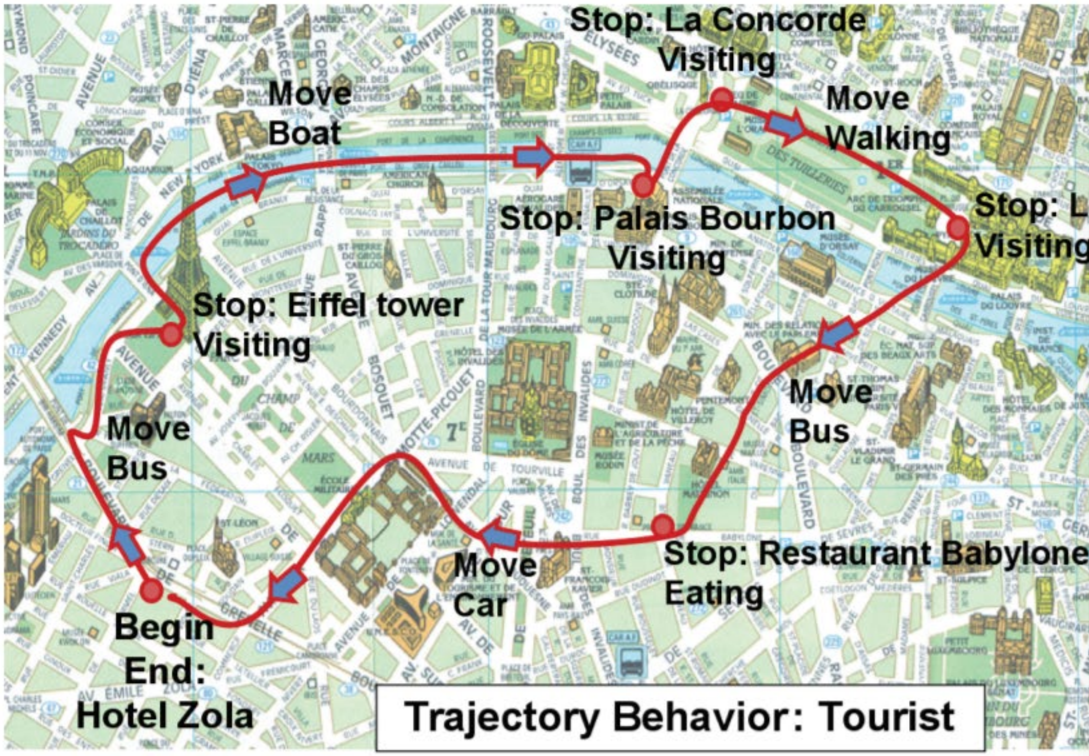


Fig. 6. A semantic tourist trajectory with annotated stops and moves (map from Mapp
unknown).

42

Semantic Trajectories Modeling and Analysis

CHRISTINE PARENT, University of Lausanne
STEFANO SPACCAPIETRA, Swiss Federal Institute of Technology (EPFL)
CHIARA RENSO, ISTI-CNR
GENNADY ANDRIENKO and NATALIA ANDRIENKO, Fraunhofer Institute IAIS
VANIA BOGORNY, Federal University of Santa Catarina (INE/UFSC)
MARIA LUISA DAMIANI, University of Milan
ARIS GKOULALAS-DIVANIS, IBM Research-Zurich
JOSE MACEDO, Federal University of Ceará
NIKOS PELEKIS and YANNIS THEODORIDIS, University of Piraeus
ZHIXIAN YAN, Swiss Federal Institute of Technology (EPFL)

Focus on movement data has increased as a consequence of the larger availability of such data due to current GPS, GSM, RFID, and sensors techniques. In parallel, interest in movement has shifted from raw movement data analysis to more application-oriented ways of analyzing segments of movement suitable for the specific purposes of the application. This trend has promoted semantically rich trajectories, rather than raw movement, as the core object of interest in mobility studies. This survey provides the definitions of the basic concepts about mobility data, an analysis of the issues in mobility data management, and a survey of the approaches and techniques for: (i) constructing trajectories from movement tracks, (ii) enriching trajectories with semantic information to enable the desired interpretations of movements, and (iii) using data mining to analyze semantic trajectories and extract knowledge about their characteristics, in particular the behavioral patterns of the moving objects. Last but not least, the article surveys the new privacy issues that arise due to the semantic aspects of trajectories.

Categories and Subject Descriptors: H. [**Information Systems**]: General; H.2.0 [**Database Management**]: General

General Terms: Algorithms, Design, Legal Aspects, Management

Additional Key Words and Phrases: Movement, mobility tracks, tracking, mobility data, trajectories, trajectory behavior, semantic enrichment, data mining, activity identification, GPS

This work is supported by the EU, FET OPEN, 2009-2012 Programme, Coordination Action type Project MODAP (Mobility, Data Mining, and Privacy) http://www.modap.org.
A. Gkoulalas-Divanis is currently affiliated with IBM Research-Ireland
Authors' addresses: C. Parent, Université de Lausanne HEC-ISI, CH-1015 Lausanne, Switzerland; S. Spaccapietra, EPFL-IC-LSIR, Station 14, CH-1015 Lausanne, Switzerland; C. Renso (corresponding author), ISTI-CNR, Via Moruzzi 1, 56010, Pisa, Italy; email: Renso.chiara@gmail.com; G. Andrienko and Natalia Andrienko, Fraunhofer Institute IAIS, Schloss Birling-hoven, Sankt-Augustin, D-53754 Germany; V. Bogorny, UFSC-CTC-INE, CEP 88040-900-Campus Universitario Cx.P. 476, Florianopolis S.C., Brazil; M. L. Damiani, Universita degli studi di Milano, Via Comelico 39, 20135 Milano, Italy; A. Gkoulalas-Divanis, IBM Research Ireland, Damastown Industrial Estate Mulhuddart, Dublin 15, Ireland; J. Macedo, Campus do PICI, Computer Science Department, Fortaleza, Brazil, CEP 60.L15-190; N. Pelekis, Department of Statistics and Insurance Science, University of Piraeus, Karaoli-Dimitriou 80, Piraeus, GR-18534, Greece; Y. Theodoridis, Department of Informatics, University of Piraeus, Kararoli-Dimitriou 80, Piraeus, GR-18534, Greece; Z. Yan, EPFL-IC-LSIR, Station 14, CH-1015 Lausanne, Switzerland.

DOI: http://dx.doi.org/10.1145/2501654.2501656

ACM Computing Surveys, Vol. 45, No. 4, Article 42, Publication date: August 2013.

➲ Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., ... & Theodoridis, Y. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR)*, *45*(4), 42.

Semantic annotation: Behaviours

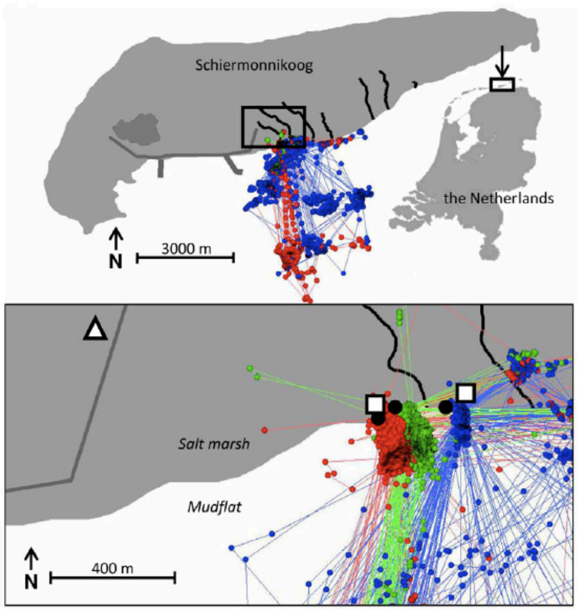
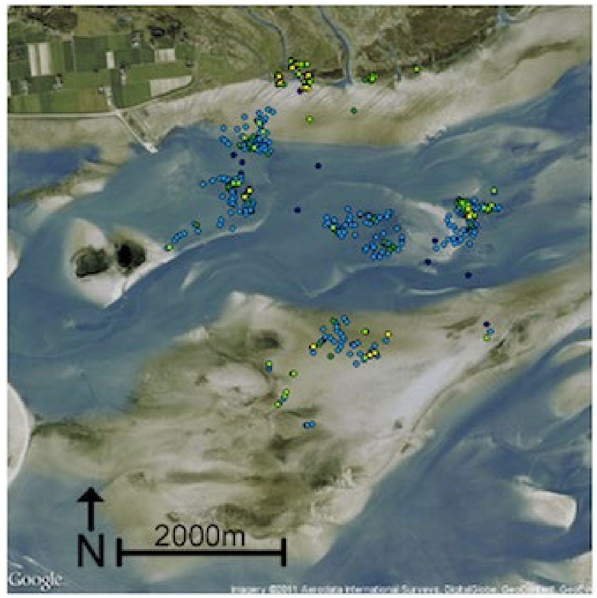
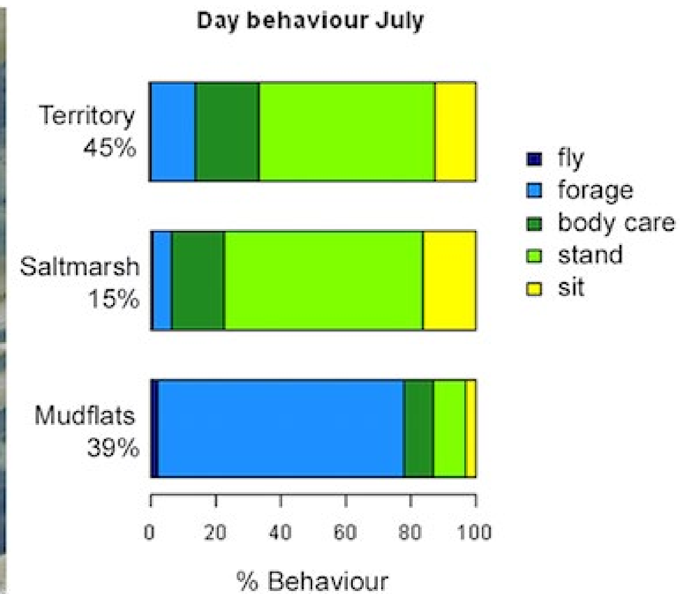


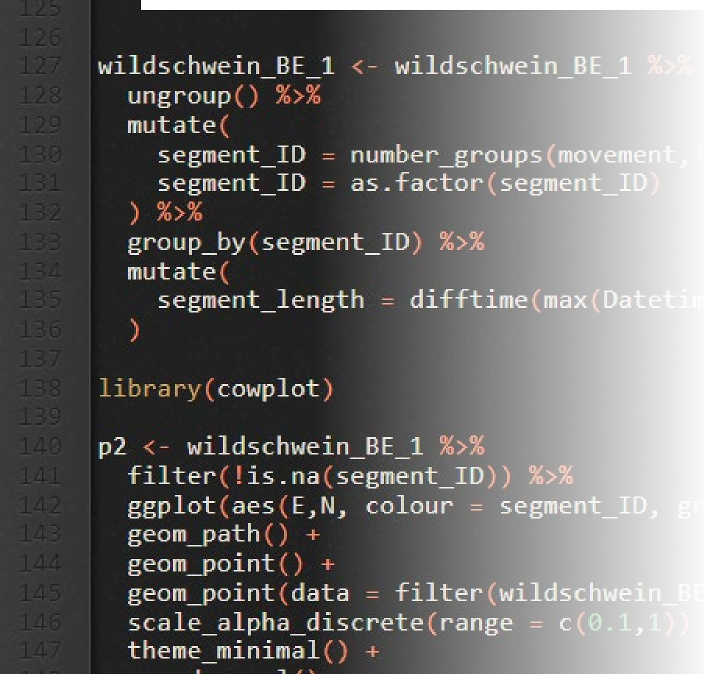
Figure 1. The study area on the island of Schiermonnikoog, the Netherlands (53.29°N, 06.10°E) at different spatial scales. The points represent GPS fixes of three oystercatchers (green – tag 166, red – tag 167, blue – tag 169; Table S1) from 1 July 2009 to 31 July 2009, with consecutive points connected by lines. The black circles are the nests of these birds. The locations of the observation towers are indicated by a square and the base station by a triangle. Black lines represents creeks, dark grey lines represent urban infrastructure.
doi:10.1371/journal.pone.0037997.g001

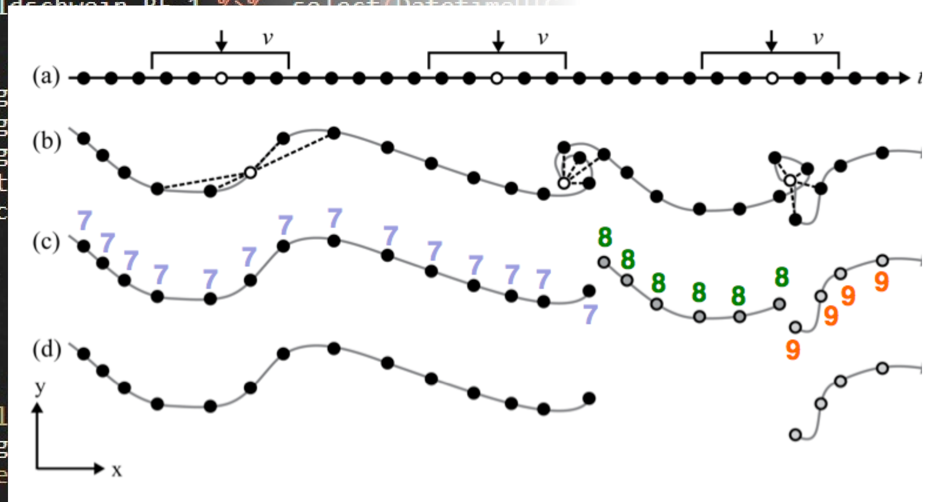


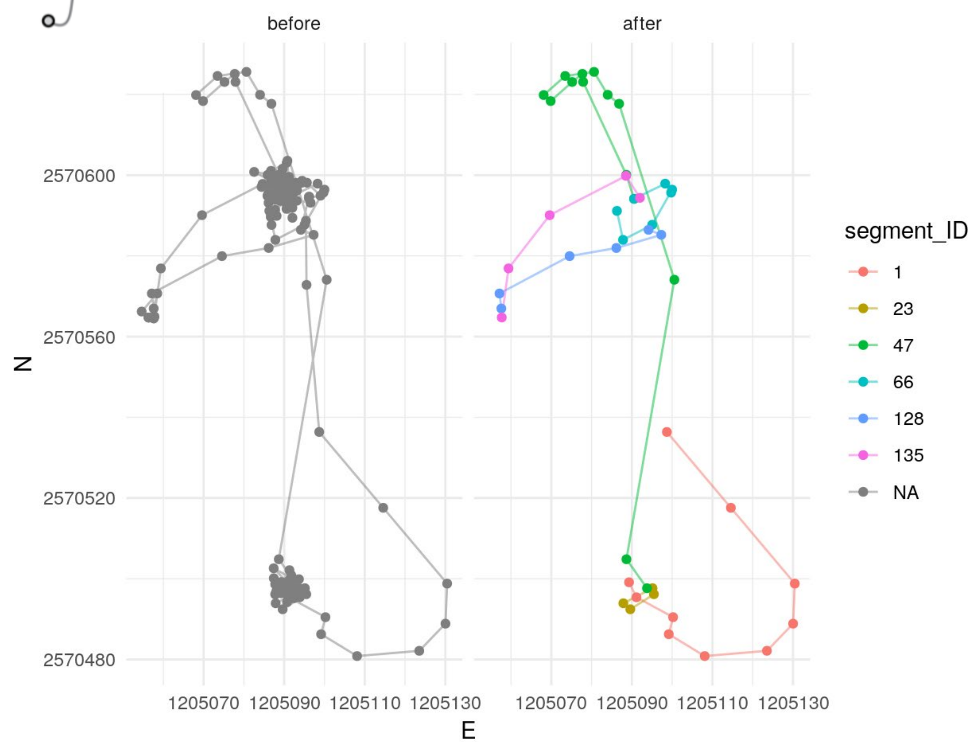


➲ SHAMOUN-BARANES, J., et al. 2012. From Sensor Data to Animal Behaviour: An Oystercatcher Example. PLoS ONE, 7(5), e37997.

R exercises


wildschwein_BE_1 %>% select(
p1
p1
if
} e
wildschwein_BE_1 <- wildschwein_BE_1 %>%
ungroup() %>%
mutate(
segment_ID = number_groups(movement,
segment_ID = as.factor(segment_ID)
) %>%
group_by(segment_ID) %>%
mutate(
segment_length = difftime(max(Datetim
)
library(cowplot)
p2 <- wildschwein_BE_1 %>%
filter(!is.na(segment_ID)) %>%
ggplot(aes(E,N, colour = segment_ID, gr
geom_path() +
geom_point() +
geom_point(data = filter(wildschwein_BE
scale_alpha_discrete(range = c(0.1,1))
theme_minimal() +


(a)
(b)
(c)
(d)
v
t
y
x


before
after
N
E
2570600
2570560
2570520
2570480
1205070 1205090 1205110 1205130
segment_ID
1
23
47
66
128
135
NA

Summary: Segmentation

- The most basic segmentation **separates stops from moves**. This can be an important **preprocessing** step when really only movement shall be studied.
- There are many different ways of segmenting a trajectory (parameter thresholds, geometric measures), choosing one must consider the **application context**.
- Segmentation can also be the main analysis task, as in the semantic **annotation** of trajectories

Table of contents: Trajectory operations

Trajectory similarity

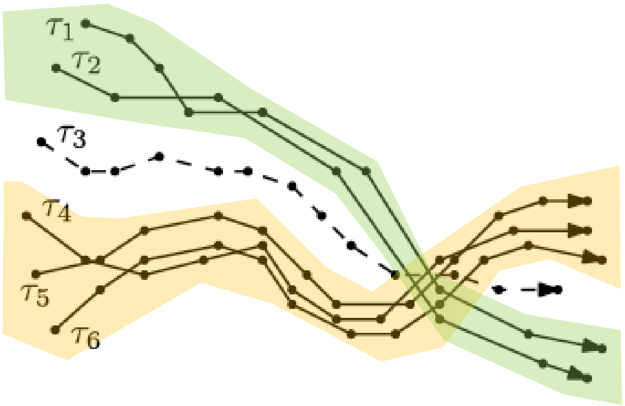


Fig. 2. Clustering six trajectories in two groups.

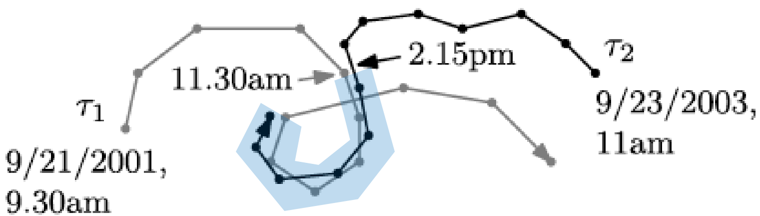


Fig. 3. The middle part of trajectory τ_1 is similar to the last part of trajectory τ_2.

➲ BUCHIN, K., et al. 2011. Finding long and similar parts of trajectories. *Computational Geometry*, 44(9), 465-476.

Why? Clustering!

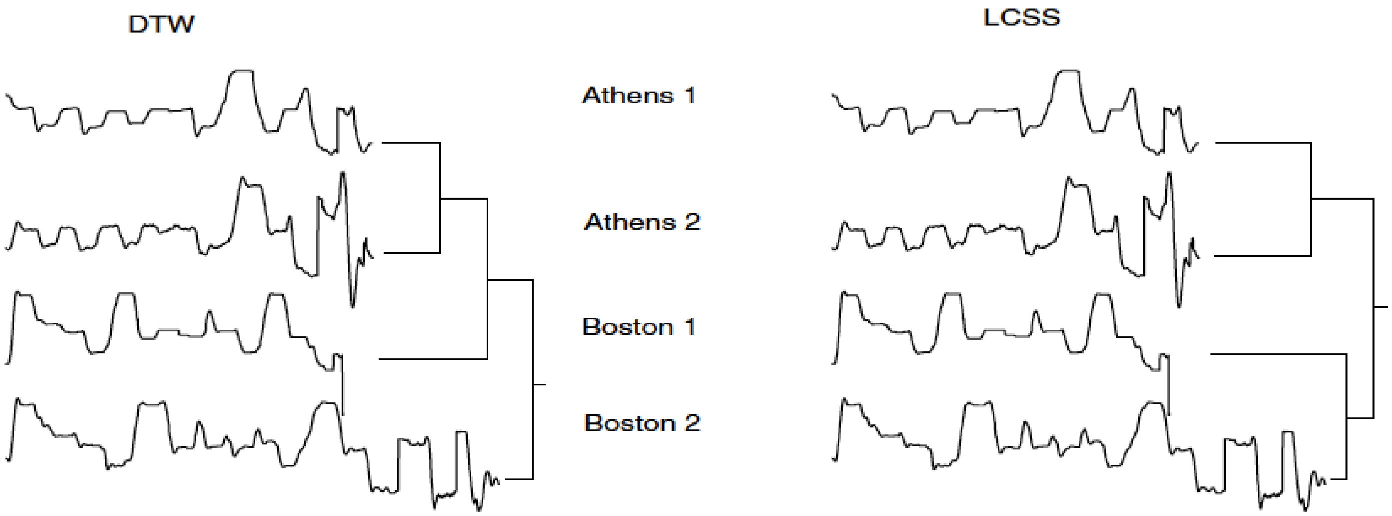


Figure 2. Hierarchical clustering of 2D series (displayed as 1D for clariry). *Left*: The presence of many outliers in the beginning and the end of the sequences leads to incorrect clustering. DTW is not robust under noisy conditions. *Right*: The $LCSS$ focusing on the common parts achieves the correct clustering.

➲ VLACHOS, M., KOLLIOS, G. and GUNOPULOS, D., 2002. Discovering Similar Multidimensional Trajectories. 18th International Converence on Data Engineering (ICDE'02).

Procedure

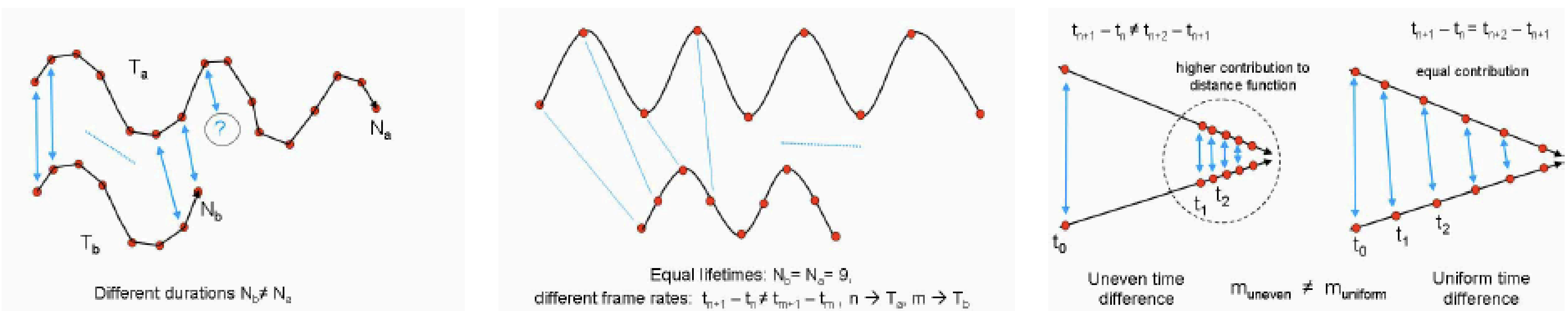
Procedure for assessing trajectory similarity and subsequent trajectory clustering

- Specify a **frame of reference**. What shall be compared? The entire lifeline of an object?
- Choose or define a **distance metric**. euclidian, manhatten
- Compute **similarity matrix** and apply a suitable clustering technique.

Why is assessing trajectory similarity difficult?

Trajectories can express:

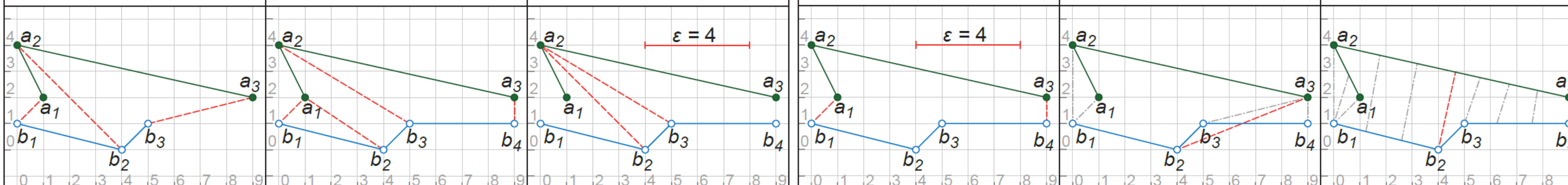
- outliers
- similarity in spatially (or temporally) disjoint sections
- very unequal duration
- unequal sampling
- heterogeneous sampling (frame rate)



➲ Figure: PORIKLI, F., Trajectory Distance Metric Using Hidden Markov Model based Representation. ed. IEEE European Conference on Computer Vision, Workshop on PETS, 2004 Prague.

Seminar S3: Similarity measures

Lock-step Euclidean Dist. **LSED**	Dynamic time warping **DTW**	Edit distance **EDR**	Longest common subseq. **LCSS**	Discrete Fréchet distance **DFD**	Fréchet distance **FD**



LSED – distances

	b_1	b_2	b_3
a_3	9.06	5.39	4.12
a_2	3	5.66	5.83
a_1	1.41	3.61	4.12

LSED = SQRT($1.41^2 + 5.66^2 + 4.12^2$) = **7.14**

DTW – squared distances

	b_1	b_2	b_3	b_4
a_3	82	29	17	1
a_2	9	32	34	90
a_1	2	13	17	65

DTW =
2 + 13 + 34 + 1 = **50**
DTW (normalized) =
SQRT(50/4) = **3.54**

EDR – thresholded (edit distance)

	b_1	b_2	b_3	b_4
a_3	1	1	1	0
a_2	0	1	1	1
a_1	0	0	1	1

EDR =
no. of edits = **2**

LCSS – thresholded (LCSS)

	b_1	b_2	b_3	b_4
a_3	0	0	0	1
a_2	1	0	0	0
a_1	1	1	0	0

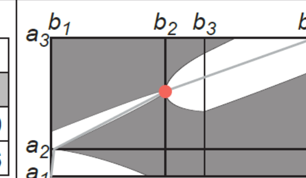
LCSS =
no. of matched pts. = **2**

DFD – distances

	b_1	b_2	b_3	b_4
a_3	9.06	**5.39**	4.12	1
a_2	3	5.66	5.83	9.49
a_1	1.41	3.61	4.12	8.06

DFD =
max(1.41, 3, 5.39, 4.12,1) =
5.39

FD



FD = **3.04**
— monotone path in free space
● max(distance along —)

Edit distance

- A bit like Scrabble…
- Borrowed from natural language processing
- a way of quantifying how dissimilar two strings
- String operations can have different costs
  - Insertions
  - Deletions
  - Substitutions
- Applied to trajectories: Trajectories are segmented, segments are treated as letters and their similarity is assessed

Z	ü	r	i	c	h

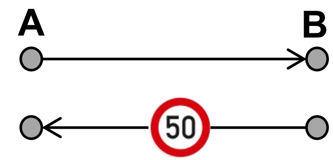
- **ü** *substituted by u, cost 0.25*
- **e** *inserted, cost 0.5*
- **i** *deleted, cost 0.5*
- **er** *inserted, cost 2*0.5*

total cost: 2.25

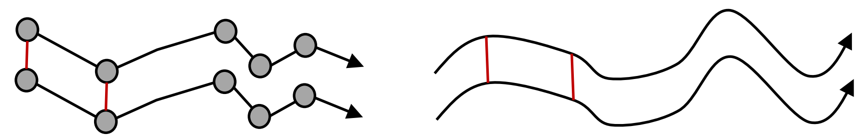
Z	u	e	r	c	h	e	r

Properties of similarity measures

- Metric vs. non-metric

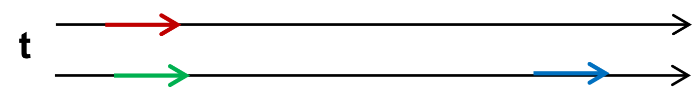


- Discrete vs. continuous

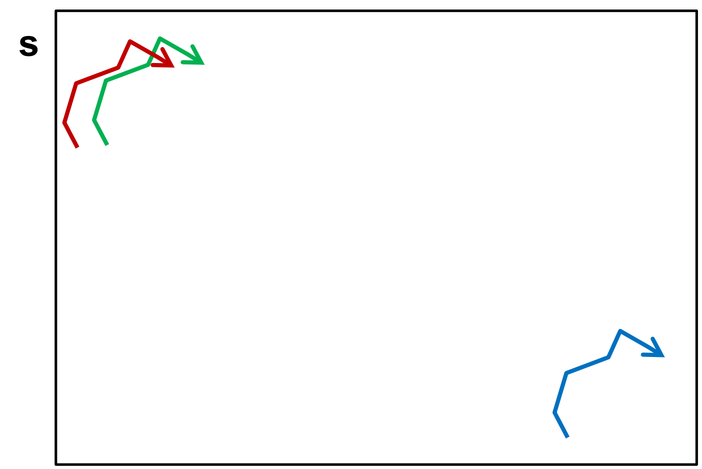


same time the same action or the same action different time

- Relative vs. absolute time

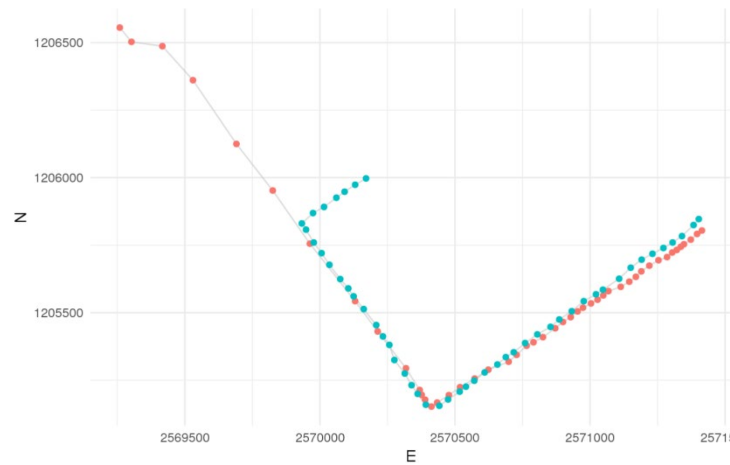


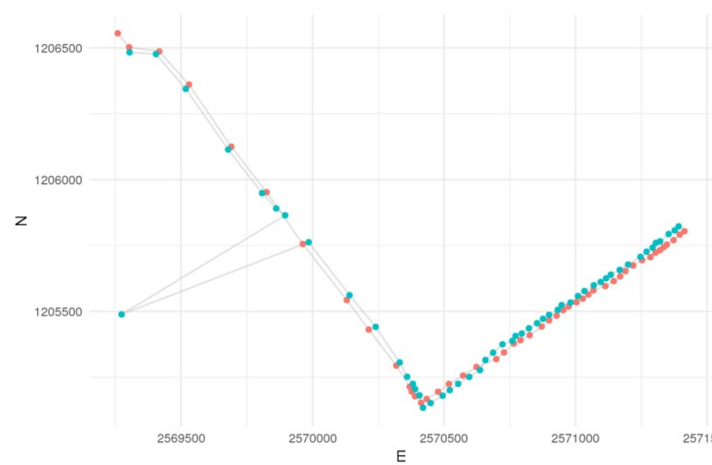
- Relative vs. absolute space



R Exercise

```r
wildschwein_BE_1 %>%  select(DatetimeUTC,

p1 <- wildschwein_BE_1 %>%
  ggplot() +
  geom_path(aes(E,N), alpha = 0.5) +
  geom_point(aes(E,N,colour = movement))
  theme_minimal() +
  coord_equal()

p1



if (knitr::is_html_output()){
  library(plotly)
  ggplotly(p1)
} else{print("Interactive map only availa


wildschwein_BE_1 <- wildschwein_BE_1 %>%
  ungroup() %>%
  mutate(
    segment_ID = number_groups(movement,
    segment_ID = as.factor(segment_ID)
  ) %>%
  group_by(segment_ID) %>%
  mutate(
    segment_length = difftime(max(Datetim
  )

library(cowplot)

p2 <- wildschwein_BE_1 %>%
  filter(!is.na(segment_ID)) %>%
  ggplot(aes(E,N, colour = segment_ID, gr
  geom_path() +
  geom_point() +
  geom_point(data = filter(wildschwein_BE
  scale_alpha_discrete(range = c(0.1,1))
  theme_minimal() +
```





Summary: Trajectory similarity

- Trajectories are very **complex** spatio-temporal objects and hence assessing their similarity is **difficult** and can be done in many different ways.
- Similarity measures are typically used for trajectory **clustering** or **subtrajectory** matching.

Gw4 (13.05.2022): Movement Patterns Bazaar

https://en.wikipedia.org/wiki/Grand_Bazaar,_Istanbul

Gw4 (12.05.2022): Movement Patterns Bazaar

- Choose a movement pattern
- Prepare 3-4 slides
- Upload onto Moodle **until 10.05.2023 23:59** (see L3 Seminar)
- Present your pattern in class

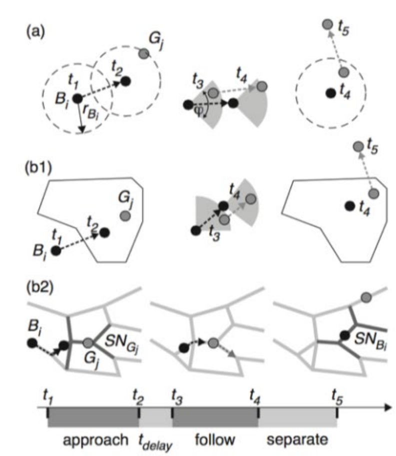


Figure 3.7: Pursuit and escape pattern (**P16**. Merki and Laube, 2012). A black actor approaches a grey reactor, they follow each other (front region φ), and finally they separate again. (a) Lagrangian perspective, (b1) Eulerian perspective, (b2) Eulerian perspective for network bound movement.

Laube, P. (2014). *Computational Movement Analysis*. Springer, p. 39.

https://en.wikipedia.org/wiki/Grand_Bazaar,_Istanbul