Module INM433 – Visual Analytics

Lecture 05

Analysis of mobility (movement data)

given by

prof. Gennady Andrienko and

prof. Natalia Andrienko

Content and objectives

- The lecture is dedicated to data representing trajectories of moving objects. We consider their structure and properties, which depend on the methods and technologies used for data collection. We explain the differences between quasi-continuous and episodic movement data and the implications for analysis.
- You will learn how to identify stops in trajectories and how to divide trajectories into trips based on the detected stops. You will also learn how to extract other movement events from trajectories.
- A method for spatial abstraction and summarisation of movement data will be introduced, with which a sets of trajectories can be compactly represented and also transformed into spatial time series.
- We show how trajectories can be analysed using density-based clustering with a set of specific distance functions.

Structure and properties of movement data

(trajectories of moving objects)

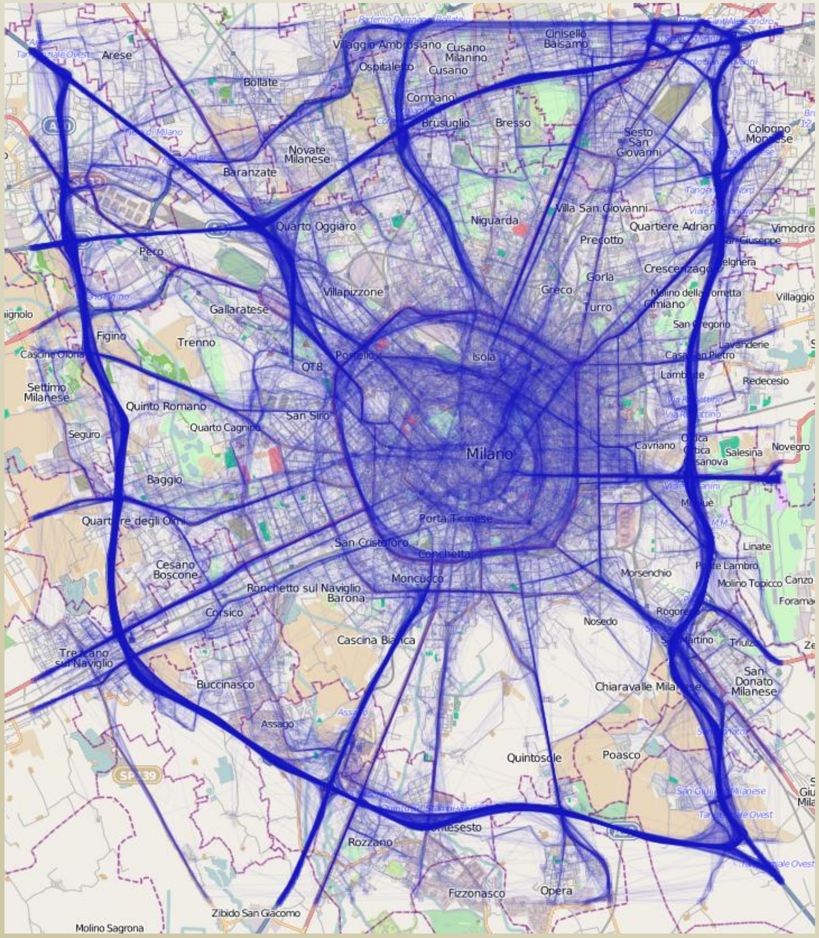
Structure of trajectory data

- A trajectory of a moving object (shortly: *mover*) is represented by a sequence of *position records*: (time, location, <thematic attributes>)
  - The records specify where the object was at different time moments.
- When a dataset contains trajectories of diverse moving objects, the position records must also contain object identifiers:
  - (object identifier, time, location, <thematic attributes>)
- The structure of a trajectory dataset: $O \times T \rightarrow S \{\times A\}$
  - I.e., object-referenced time series of spatial locations
- Besides, a trajectory by itself is a spatio-temporal object.
  - Spatial position: the path (line in space).
  - Existence time: the interval from the first to the last location.
  - A trajectory can be viewed as a line in the space-time continuum.

Example dataset: trajectories of cars in Milan

- GPS-tracks of 17,241 cars in Milan, Italy
- Time period: April 01-07, 2007 (Sunday to Saturday)
- Received from Octo Telematics www.octotelematics.com special thanks to Tina Martino
- Data structure:
  - Anonymised car identifier
  - Date and time
  - Geographic coordinates
  - Speed

OCTO The reliable way

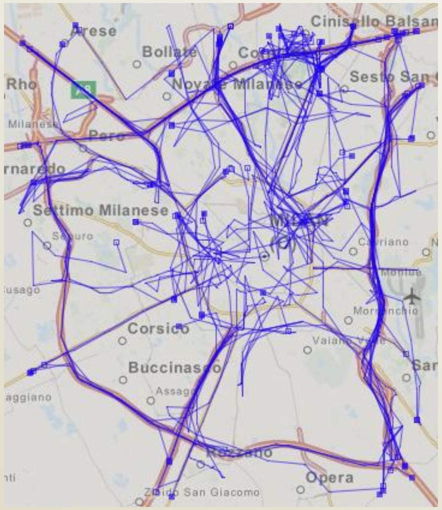


The trajectories from one day are drawn on a map with 5% opacity

Trajectories: data and visual representation

Sequences of position records

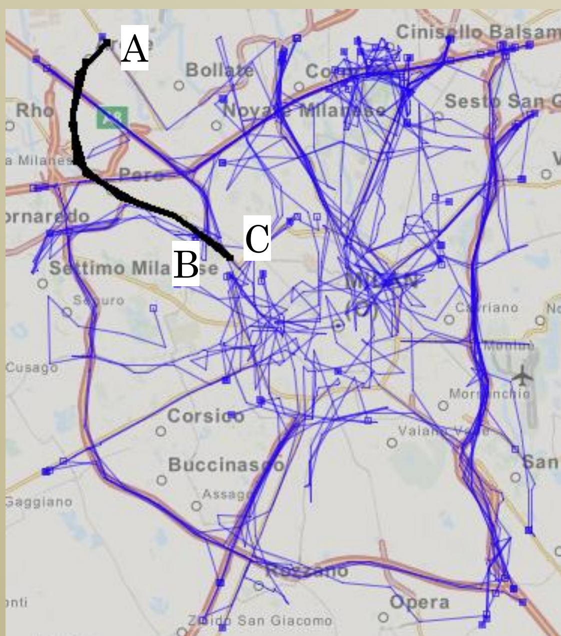
	Car id	point N	longitude	latitude	time
1	104876	1	9.119127	45.558304	04/04/2007 06:45:15
2	104876	2	9.142448	45.559753	04/04/2007 06:48:49
3	104876	3	9.156955	45.554962	04/04/2007 06:54:54
4	104876	4	9.156504	45.55017	04/04/2007 07:00:12
5	104876	5	9.156504	45.55017	04/04/2007 07:11:08
6	104876	6	9.156844	45.547703	04/04/2007 07:13:26
7	104876	7	9.156909	45.547688	04/04/2007 07:19:23
8	104876	8	9.162037	45.554867	04/04/2007 07:40:02
9	104876	9	9.167628	45.55907	04/04/2007 08:02:32
10	104876	10	9.172845	45.555725	04/04/2007 08:05:38
11	104876	11	9.172696	45.555492	04/04/2007 10:03:31
12	104876	12	9.166886	45.54498	04/04/2007 10:09:38
13	104876	13	9.163299	45.557983	04/04/2007 10:12:05
14	104876	14	9.162168	45.554855	04/04/2007 10:13:51
15	104876	15	9.162158	45.55487	04/04/2007 11:36:23
16	104876	16	9.162622	45.557976	04/04/2007 12:08:17
17	104876	17	9.16232	45.55496	04/04/2007 12:09:19
18	104876	18	9.162361	45.554943	04/04/2007 15:30:22
19	104876	19	9.122161	45.55825	04/04/2007 15:38:51
20	110800	1	9.266509	45.386322	04/04/2007 05:21:45
21	110800	2	9.261211	45.40307	04/04/2007 05:22:57
22	110800	3	9.247442	45.418125	04/04/2007 05:24:13
23	110800	4	9.254333	45.43362	04/04/2007 05:29:45
24	110800	5	9.257282	45.451492	04/04/2007 05:32:44
25	110800	6	9.252168	45.468708	04/04/2007 05:34:21
26	110800	7	9.251433	45.48671	04/04/2007 05:35:48
27	110800	8	9.258238	45.504066	04/04/2007 05:37:05
28	110800	9	9.260647	45.522255	04/04/2007 05:38:26
29	110800	10	9.278728	45.53516	04/04/2007 05:39:48
30	110800	11	9.274316	45.533176	04/04/2007 11:57:53
31	110800	12	9.261258	45.519493	04/04/2007 11:59:21
32	110800	13	9.256271	45.502003	04/04/2007 12:00:51
33	110800	14	9.250924	45.484287	04/04/2007 12:02:11
34	110800	15	9.251606	45.465813	04/04/2007 12:03:32
35	110800	16	9.258425	45.44919	04/04/2007 12:04:53
36	110800	17	9.251738	45.43149	04/04/2007 12:06:14
37	110800	18	9.251131	45.415356	04/04/2007 12:07:39
38	110800	19	9.260662	45.399742	04/04/2007 12:08:54
39	110800	20	9.270314	45.382904	04/04/2007 12:10:10
40	116291	1	9.234817	45.508648	04/04/2007 18:01:18
41	116291	2	9.257177	45.51305	04/04/2007 18:06:57
42	116291	3	9.255232	45.498352	04/04/2007 18:09:05

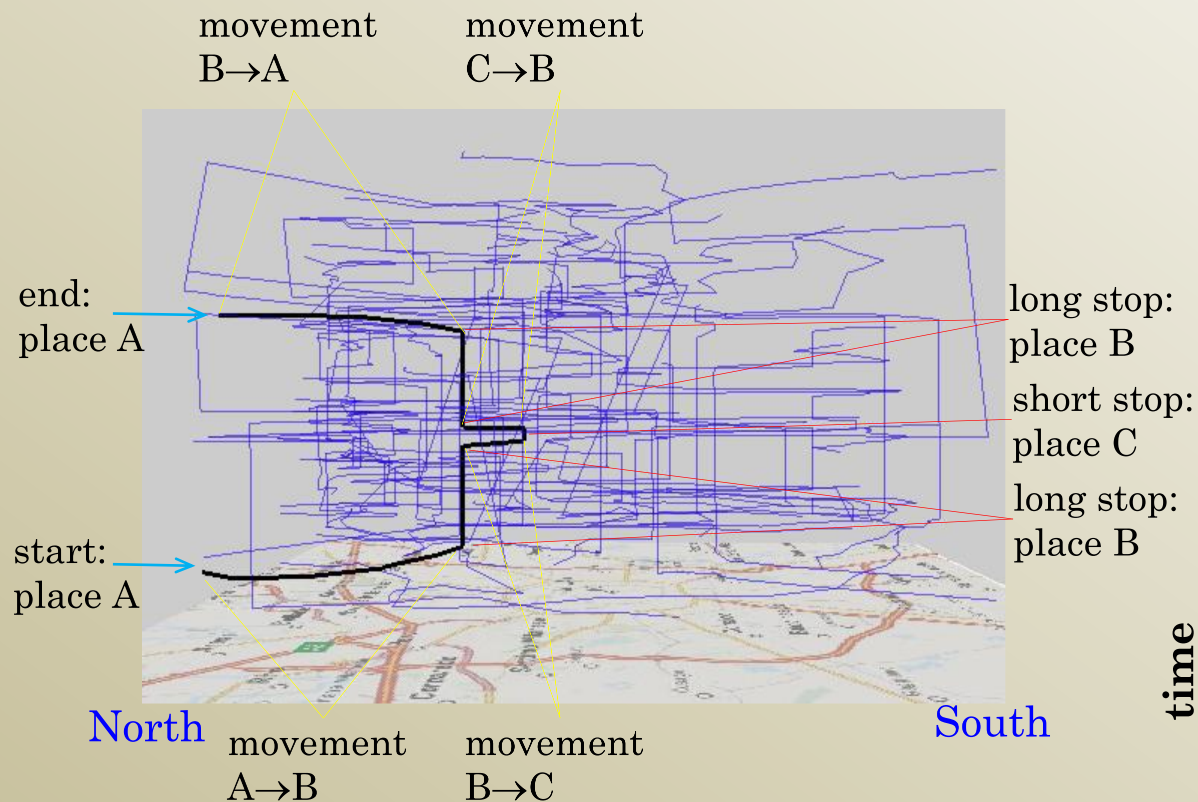


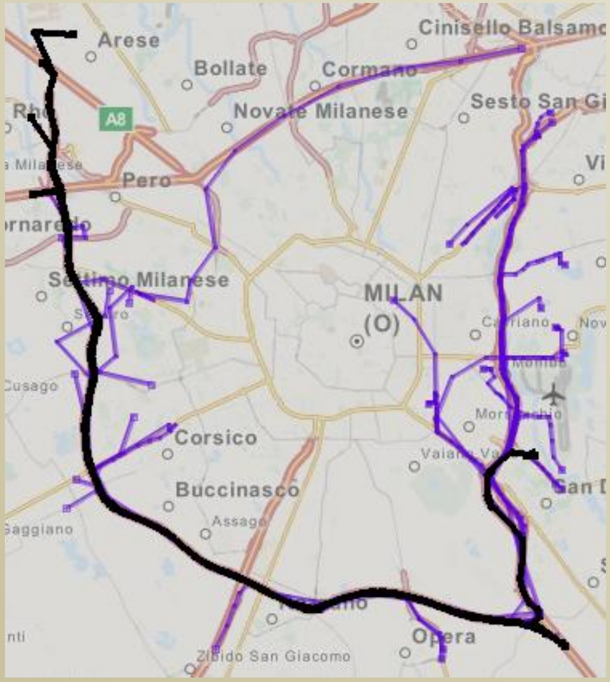
Space-time cube

Spatio-temporal view

spatial footprint



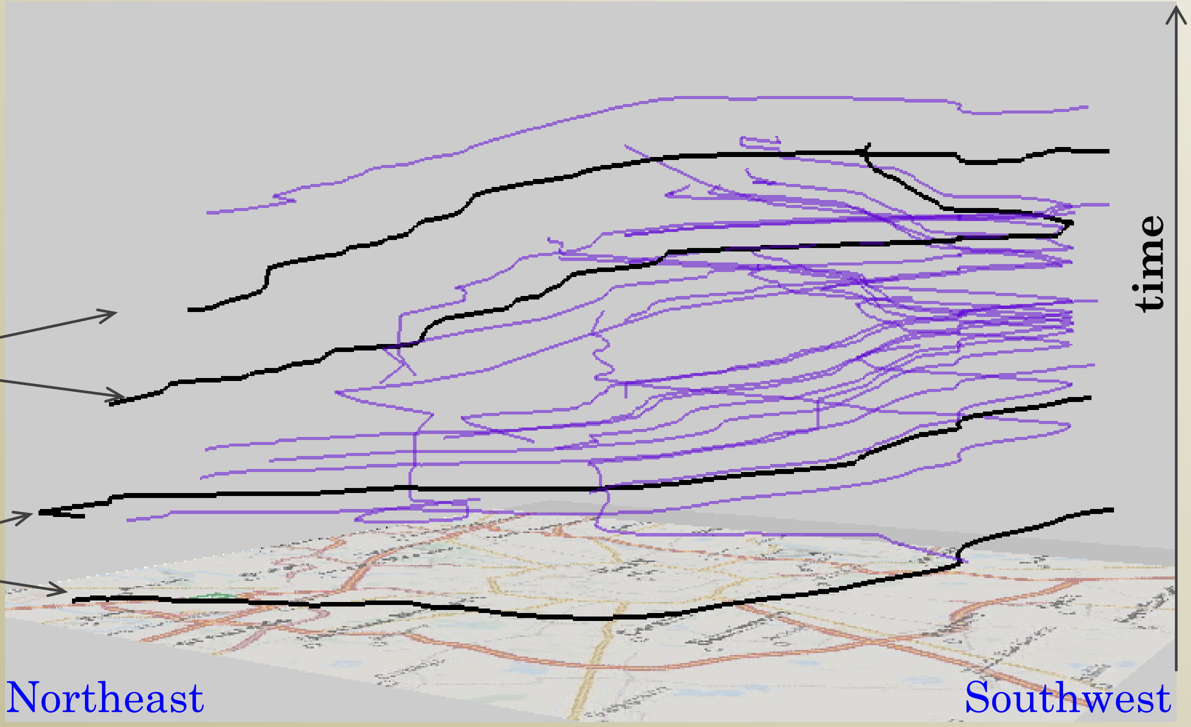




Spatio-temporal view

slow movement

fast movement



The interpretation of the line slope is the same as for OD moves.

Trajectories as objects

- As objects, trajectories may have various attributes.
- Static attributes: characterise the trajectory as a whole and do not vary over time
  - Path length, duration, total displacement (straight-line distance between the start and end locations), sinuosity (path length / displacement ratio), tortuosity (measure of zigzagginess), …
    - Can be computed from the position records
  - Other attributes can be attached: transportation means, trip purpose, …
- Time-variant (dynamic) attributes, i.e., time series: characterise the movement at different times
  - Spatial position
  - Speed, direction, acceleration *(can be computed from the position records)*
  - Other attributes: transportation means, physical condition of the mover, …

Methods of collecting trajectory data

- Time-based: positions of movers are recorded at regularly spaced time moments.
- Change-based: a record is made when mover's position, or speed, or movement direction differs from the previous one.
- Location-based: a record is made when a mover enters or comes close to a specific place, e.g. where a sensor is installed.
- Event-based: positions and times are recorded when certain events occur, in particular, when movers perform certain activities
  - mobile phone calling, sending an SMS, posting a Twitter message with coordinates, taking a photo with a GPS-enabled device, …
- Combinations, e.g., time-based position measurement but change-based recording (a position is not recorded if no change have occurred).

Technologies for collecting or reconstructing trajectories

- GPS tracking
  - "A GPS tracking unit is a device that uses the Global Positioning System to determine the precise location of a vehicle, person, or other asset to which it is attached and to record the position of the asset" (Wikipedia).
- RFID tracking (radio-frequency identification)
  - Movers wear RFID chips (tags) containing electronically stored data.
  - RFID readers (radio transmitters-receivers) send signals to tags and read their responses. The tag data and time are recorded.
  - A trajectory of a tag carrier can be reconstructed based on the spatial positions of multiple readers the carrier has passed and the recorded times.

Technologies for collecting or reconstructing trajectories (continued)

- Bluetooth sensing
  - Bluetooth-enabled devices (e.g., mobile phones) carried by movers are registered when they come into the range of a static Bluetooth sensor.
  - The sensor records the time and the MAC address (media access control address) of a device, which uniquely identifies the device.
  - Trajectories of the devices can be reconstructed based on records from multiple sensors analogously to RFID.
  - Various problems: a mover may have several devices → multiple tracks of the same mover; the Bluetooth may not always be enabled → missing position records; …
- Reconstruction from data collected not for tracking purposes
  - Mobile phone use events: user id + event time + antenna id (can be replaced or extended by the antenna's coordinates)
  - Social media posts containing coordinates: Twitter, Flickr, YouTube, …

Privacy issues

- Movement data are usually anonymised, so that the identifiers contained in the data cannot be associated with concrete movers.

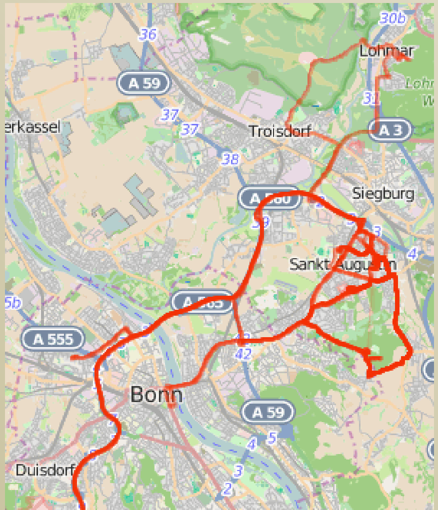
☹However, this is not sufficient!

  - Frequently visited places of a person can be easily extracted from movement data.
  - Knowing the places and visit times, someone can identify the person.
- Intensive research on protecting location privacy
  - E.g., by distorting the data
  - No ideal solution yet
- Conclusions:
  - Movement data need to be carefully protected (⇒ hard to get for research ☹)
  - Be cautious in sending geo-located posts to social media!
    - Do not send such posts from your home and work or study places!

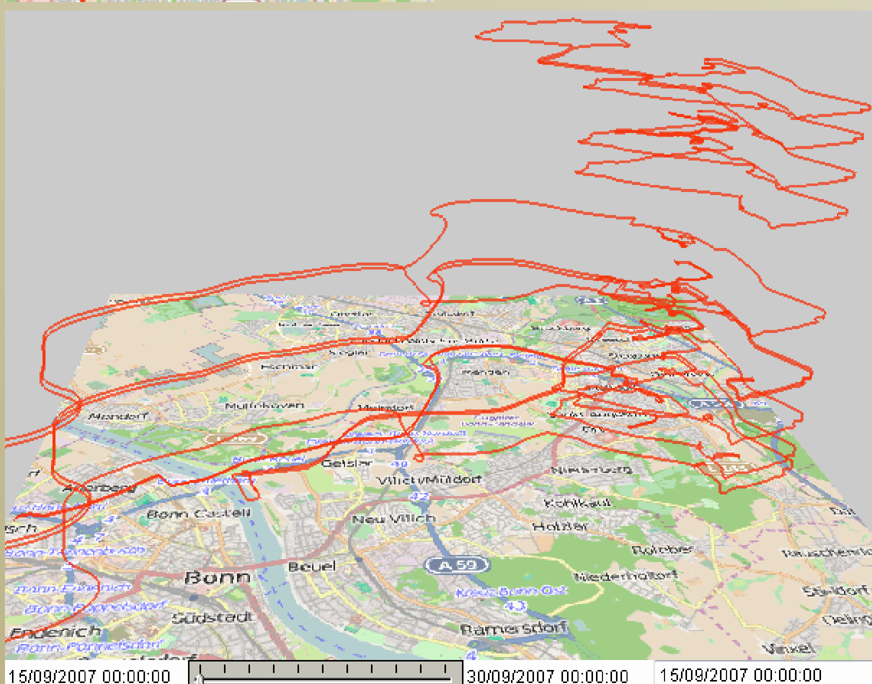
Quasi-continuous and episodic trajectories

- Glossary:
  - *Temporal resolution* = length of the time intervals between the position records (small → fine resolution, large → coarse resolution).
  - *Spatial resolution* = the minimal change of mover's position that can be reflected in the data
    - GPS tracks: fine; mobile phone data: coarse (positions = cells); RFID and Bluetooth: depend on the spatial density of the sensors; usually coarse
  - *Interpolation*: determining intermediate positions of a mover between recorded positions
- <u>Quasi-continuous</u> trajectories:
  - fine temporal and spatial resolution; interpolation is possible
- <u>Episodic</u> trajectories:
  - low temporal or spatial resolution or frequent temporal or spatial gaps between records; interpolation is not valid

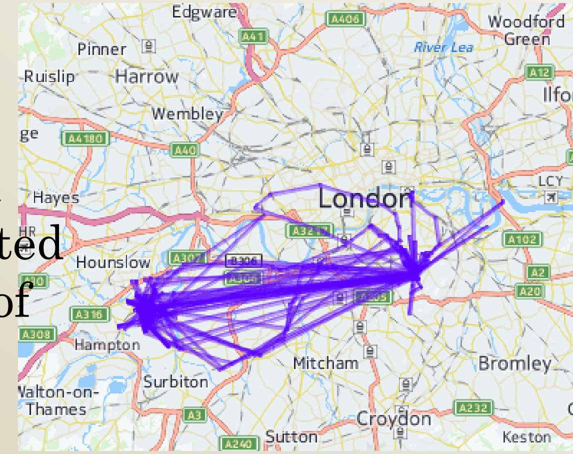
Examples

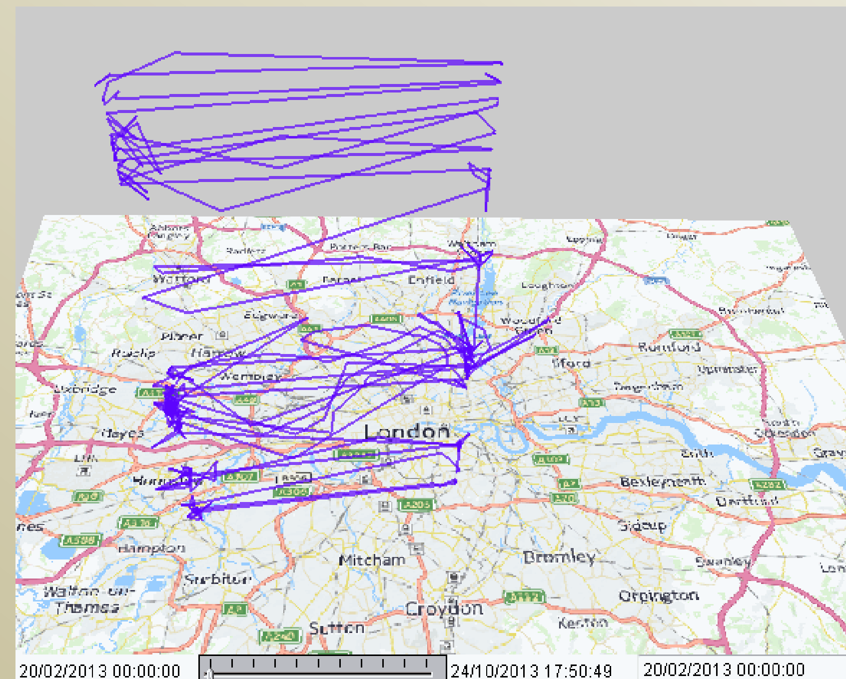


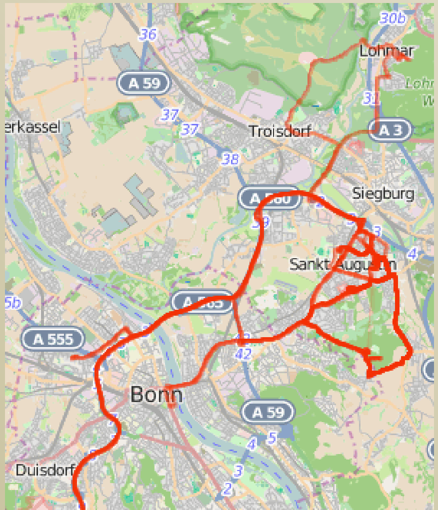
Quasi-continuous: a GPS track of a car



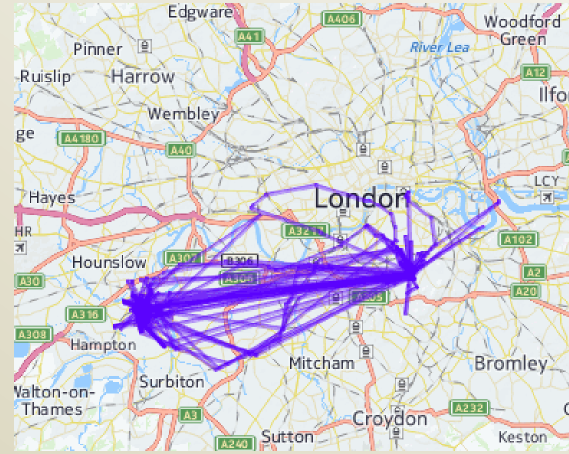
Episodic: a reconstructed trajectory of a Twitter user

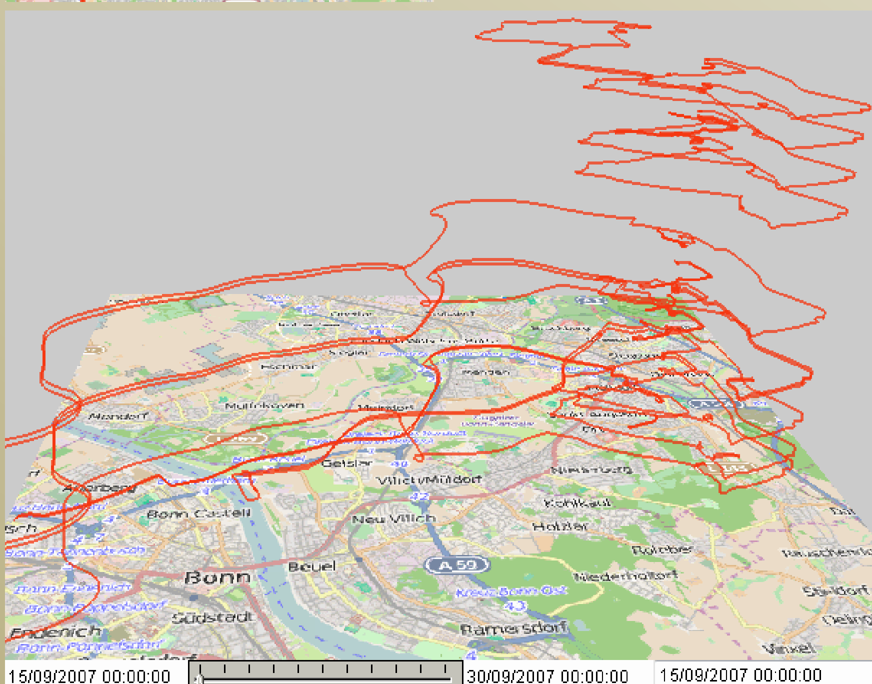


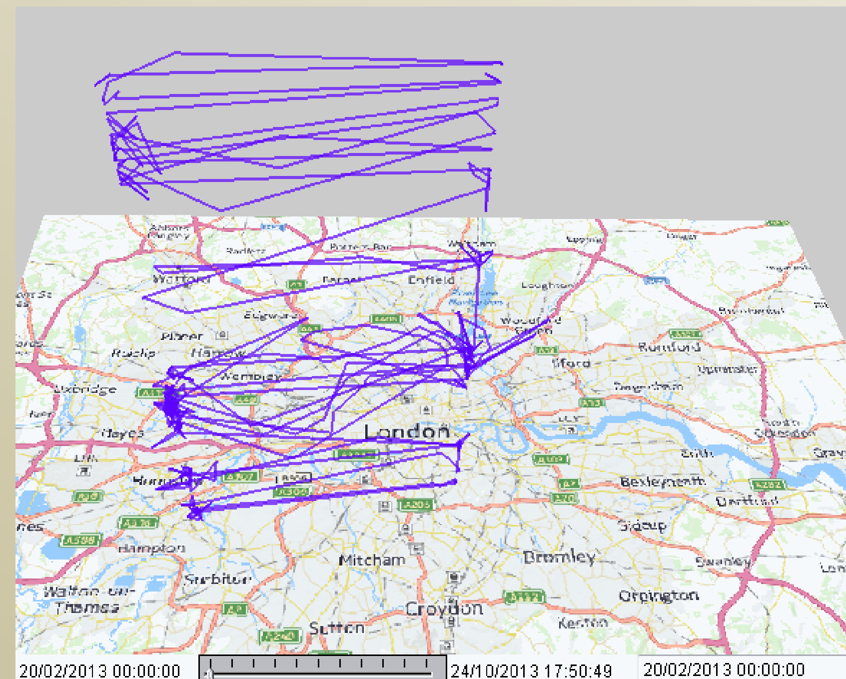




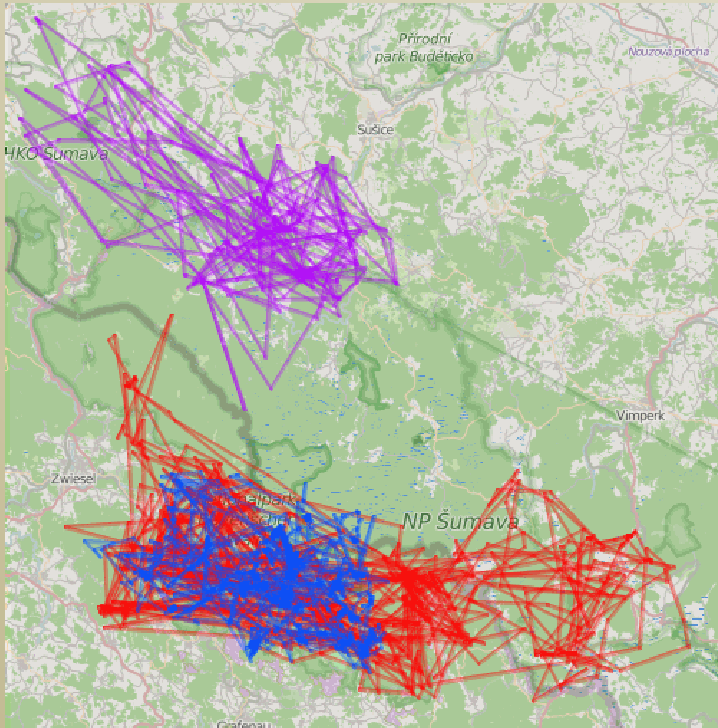
Note that repeated movements and repeated visits of the same places are present in both examples.
Hence, the privacy concerns refer to both quasi-continuous and episodic trajectories.

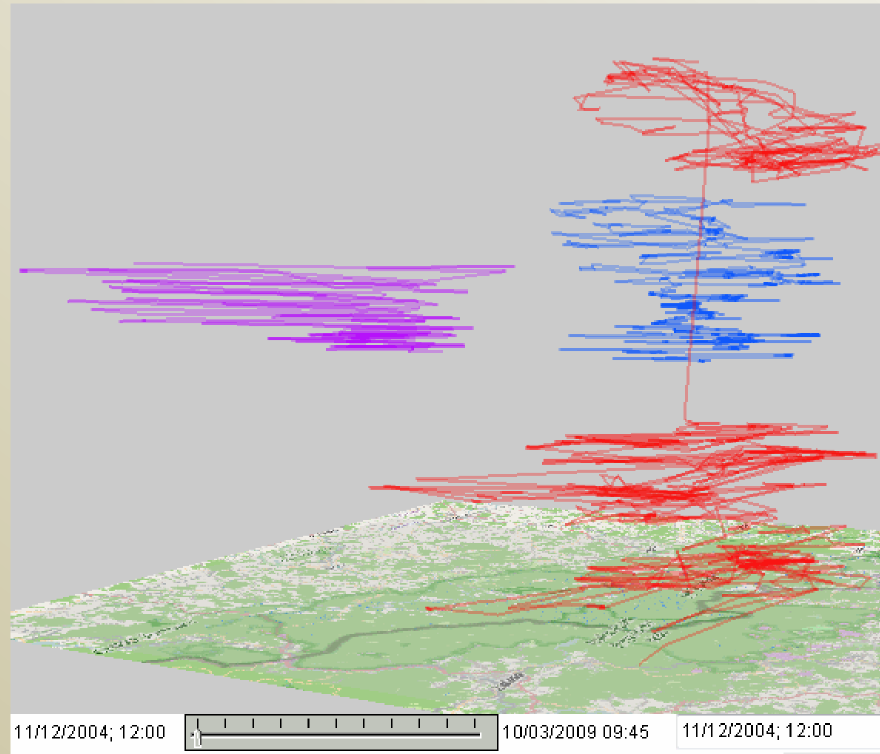






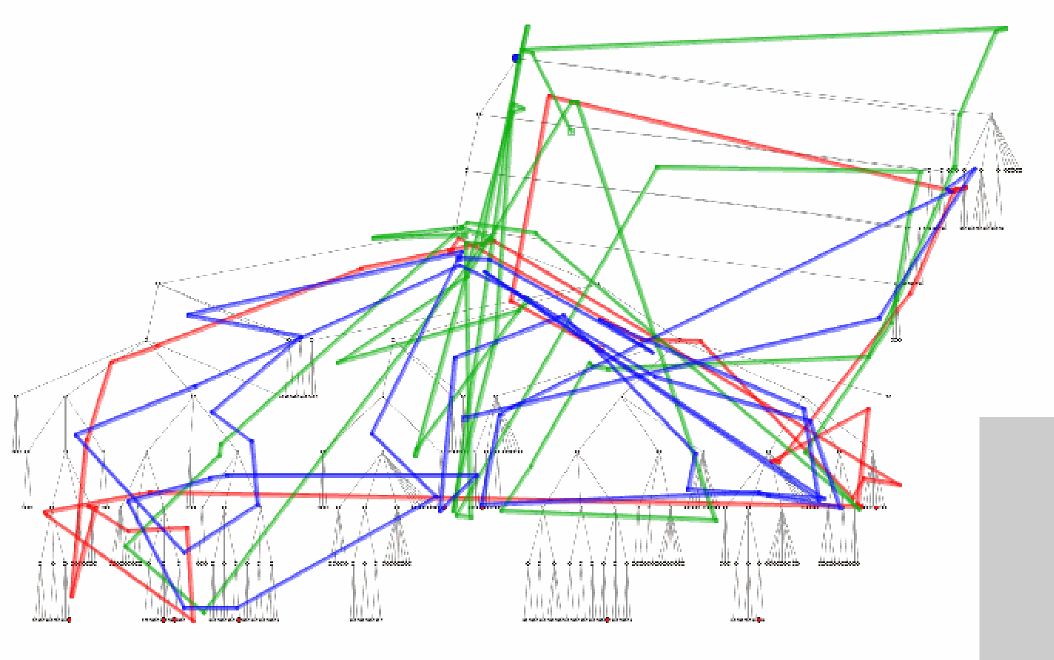
Not all GPS tracks are quasi-continuous





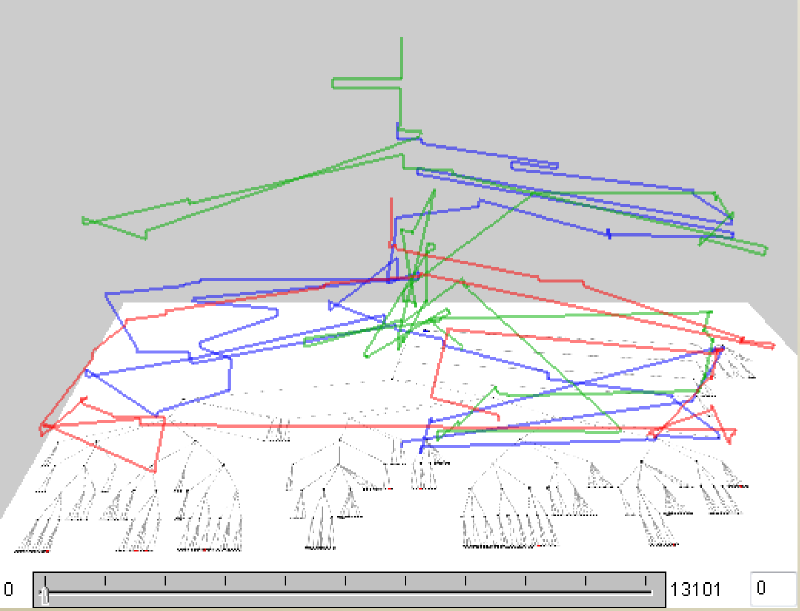
The frequency of measuring and recording positions may be intentionally reduced, e.g., for extending the battery life when tracking animals.

Non-geographic movement data



Example: eye movement tracks of 3 participants of an experiment on understanding graphs.

Eye movement data: despite a very fine temporal resolution, large spatial gaps occur between records, as in episodic trajectories. The gaps correspond to eye jumps (saccades). Interpolation is not meaningful.



Trajectories and trips

- Most often, movement data concerning a mover is a mere sequence of records (mover id, time, position) covering the whole period of observation.
- The mover might not continuously move all that time but could make stops.
- The stops and trips (movements between the stops) are not explicit in the data.
- When required for analysis purposes, the stops and/or trips need to be extracted from the trajectories.

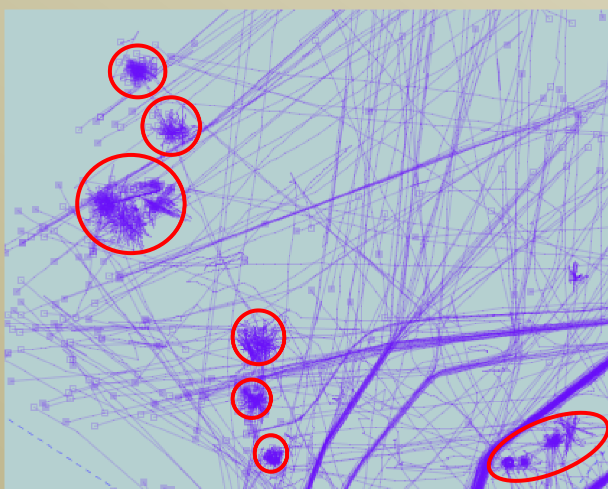
Finding stops in trajectories

- Based on time gaps: if position recording was done only during movement, stops are signified by time gaps between records.
  - E.g., a car tracking GPS device switches off when the car motor is off.
- Based on speed: speed = 0 *(during a time interval)* $\Rightarrow$ stop
  - Problem: mover's positions recorded during a stop may differ due to measurement errors $\Rightarrow$ the speed may never be 0.
- Based on a bounding box: the spatial bounding box of a sequence of positions is small $\Rightarrow$ stop
  - Requires choosing the maximal box size threshold
    - May require multiple trials when the range of positioning errors is not known in advance.
- In all cases, a minimal stop duration need to be chosen (= minimal duration of stillness that can be considered as a significant stop).

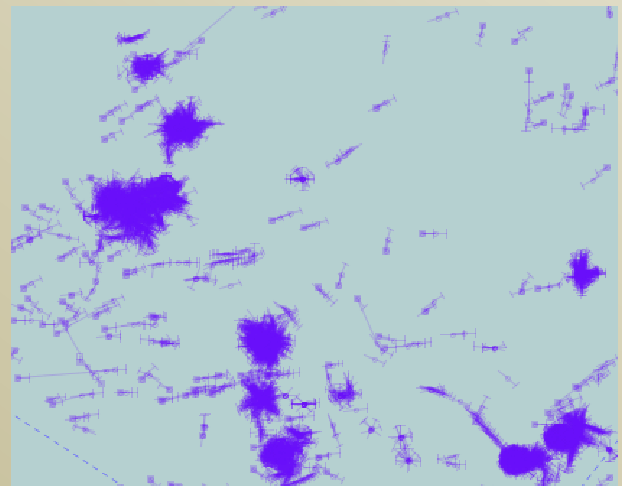
Example: interactive extraction of stops from vessel trajectories

The vessel positions were recorded also when the vessels were anchored.

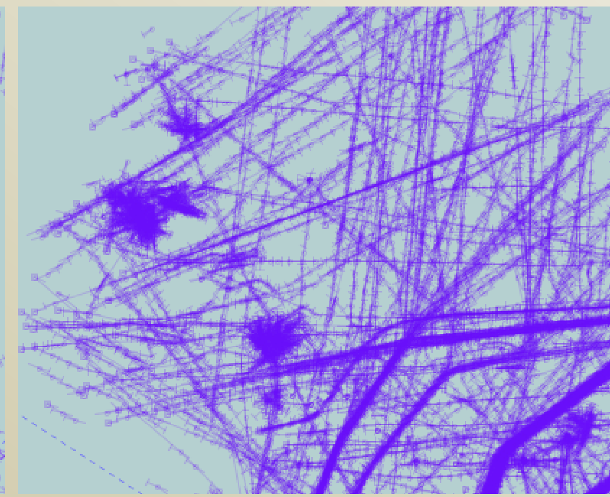
The stops of the vessels appear on a map like these tangles:



Position filter: bounding box diagonal (BBD) in 1 hour is below 3 km.

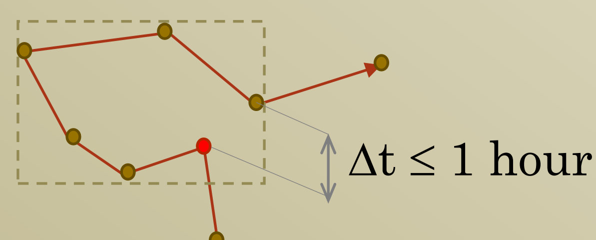


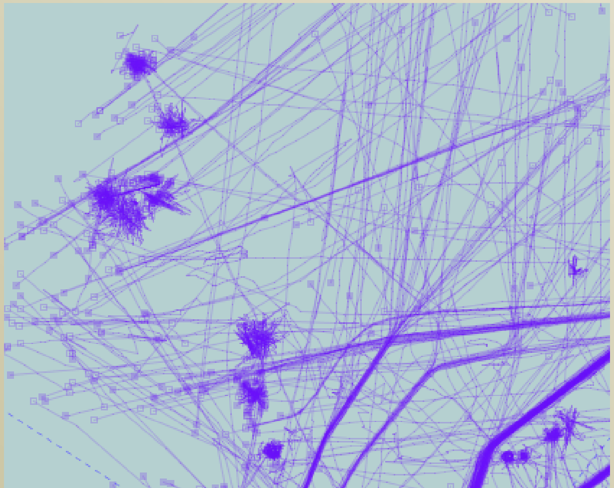
The inverse filter: BBD in 1 hour is >= 3 km. Tangles still appear.



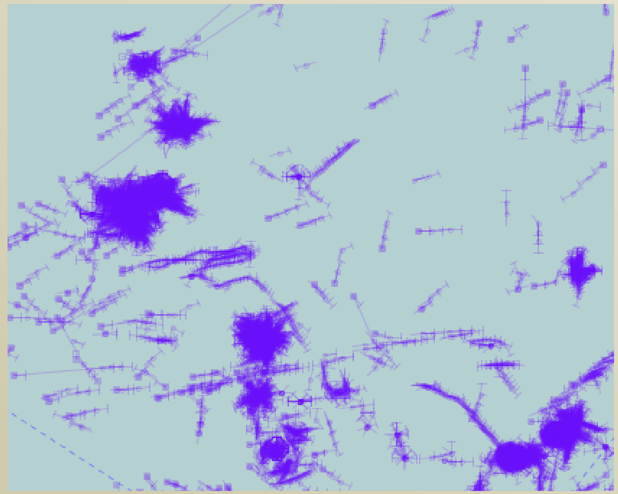
Bounding box in 1 hour:

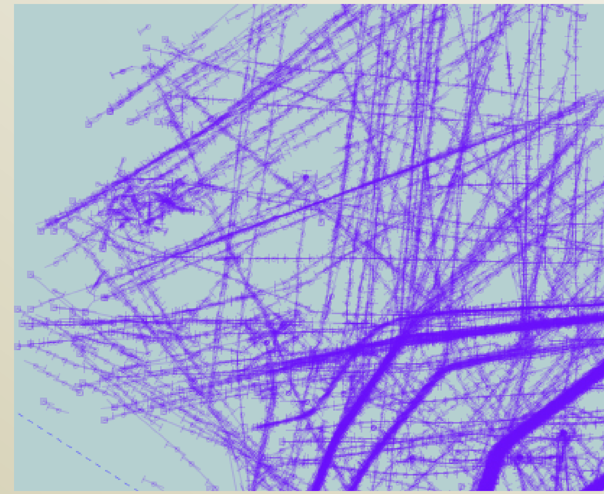
(computed for each point of a trajectory)



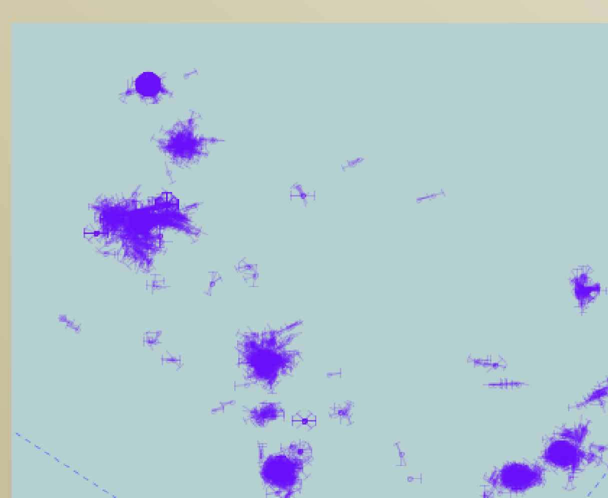


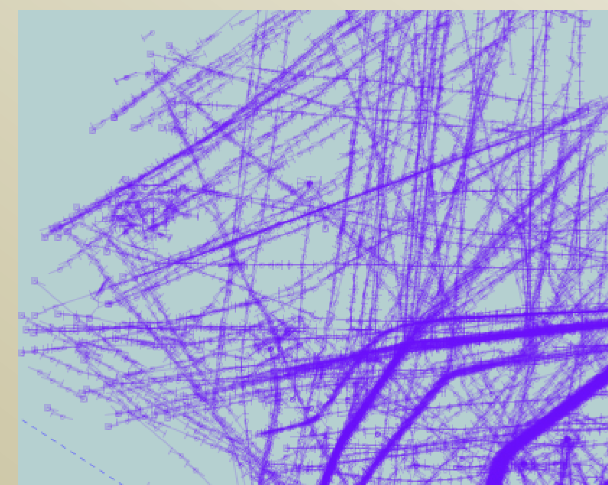
Position filter: BBD in 1 hour < 6km. The result of the inverse filter is OK, but the direct filter selects not only stops but also slow movements.





Additional position filter: sinuosity in 1 hour >= 1.5. The combination of two filters gives sufficiently good results.

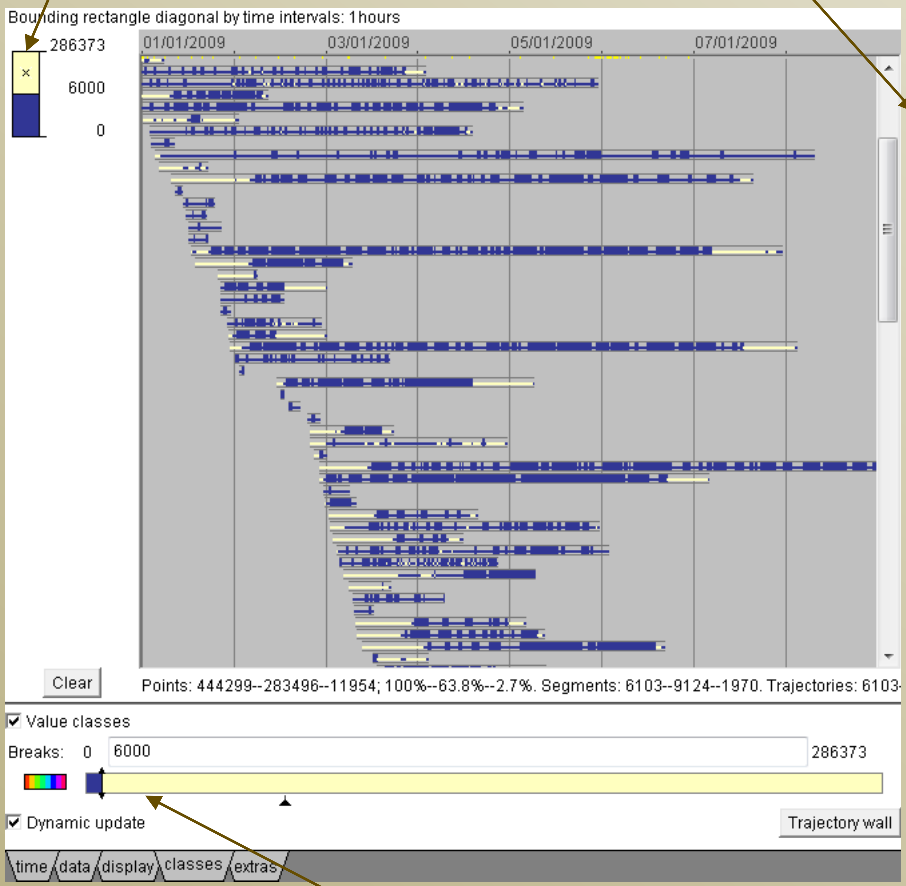


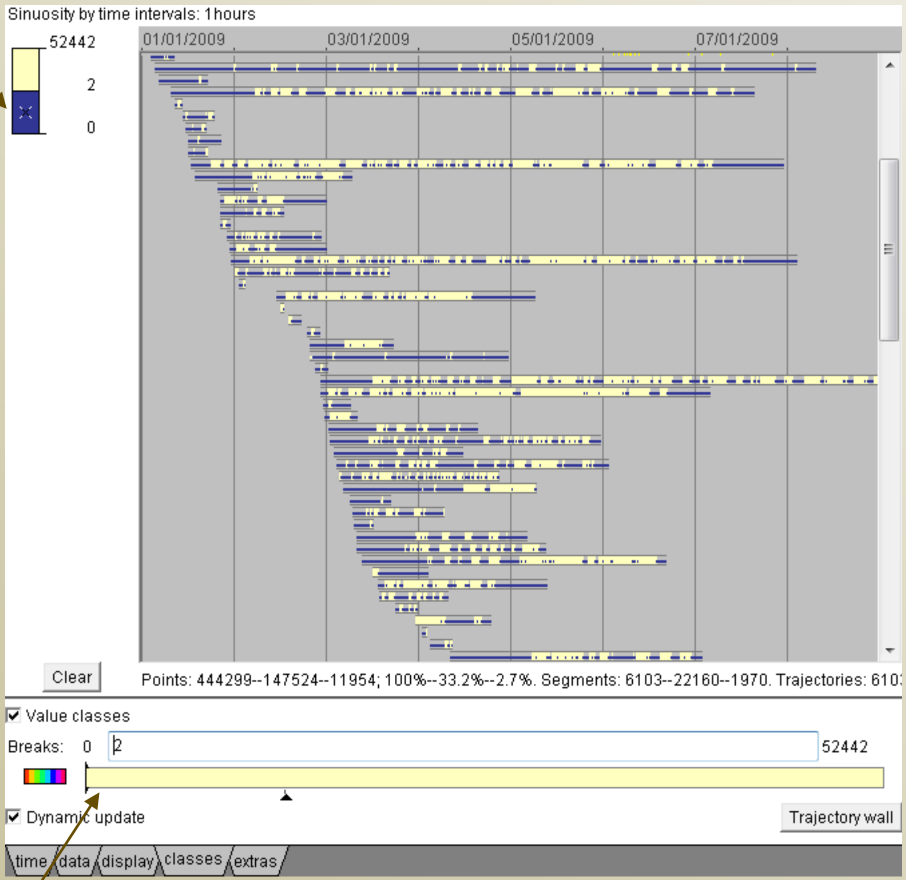


Interactive position filtering

Trajectory timeline display (Gantt chart)

Deselecting (hiding) class intervals

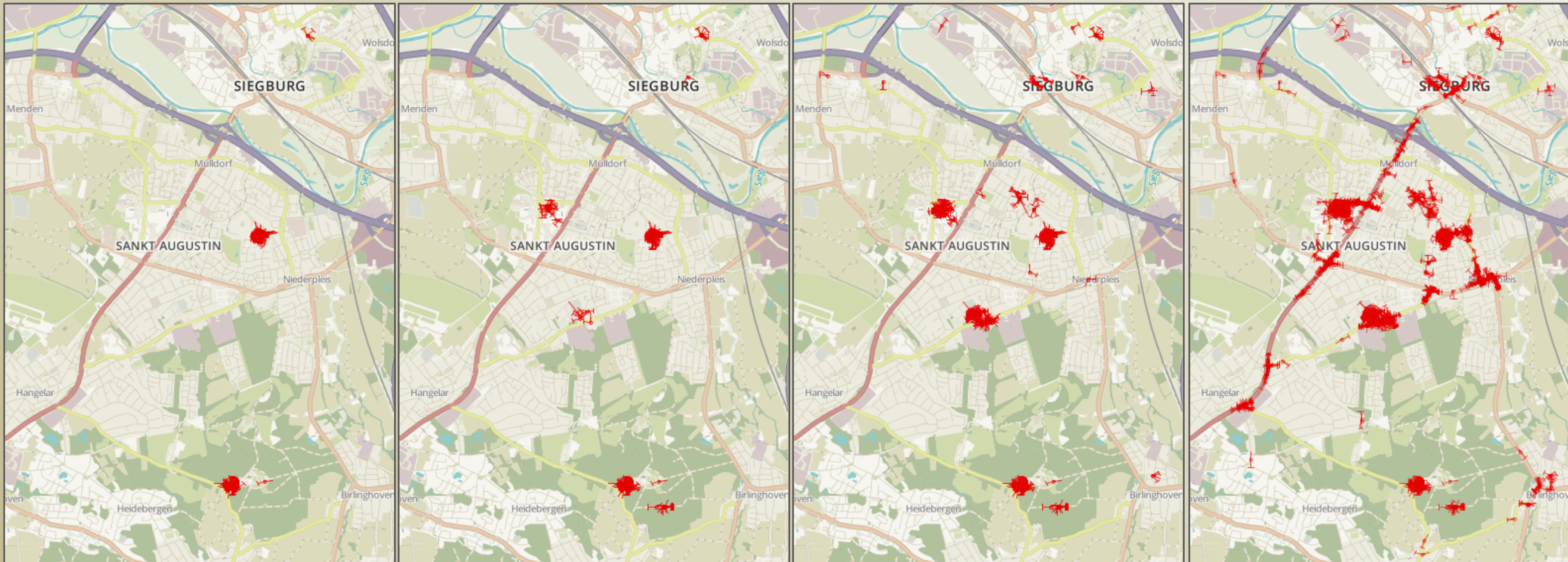




Setting class intervals

Different stop durations

Stops in a year-long trajectory of a private car



≥ 5 hours | ≥ 1 hour | ≥ 10 minutes | ≥ 1 minute

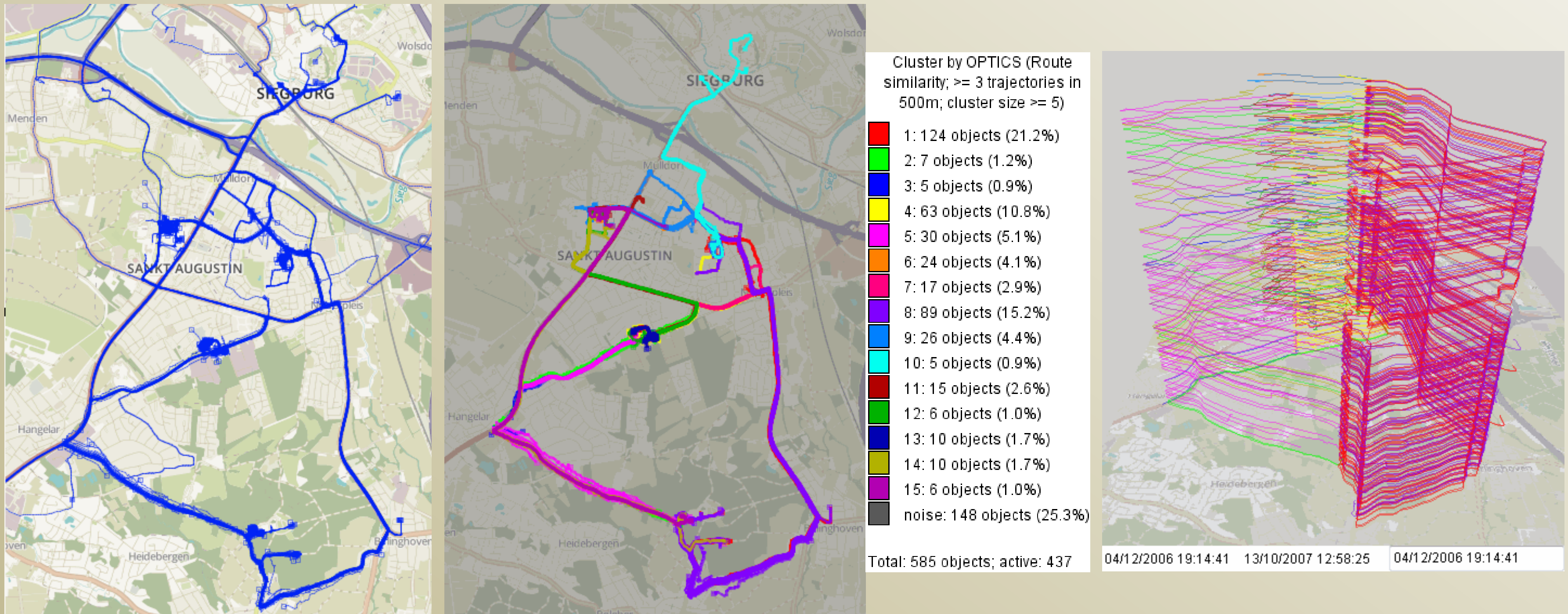
The locations of the stops with different durations have different meanings:

- Long stops: the most important places (home, work, …)
- Medium stops: important places (shopping, sports, health care, …)
- Short stops: traffic lights, traffic obstruction, …

Division of a trajectory

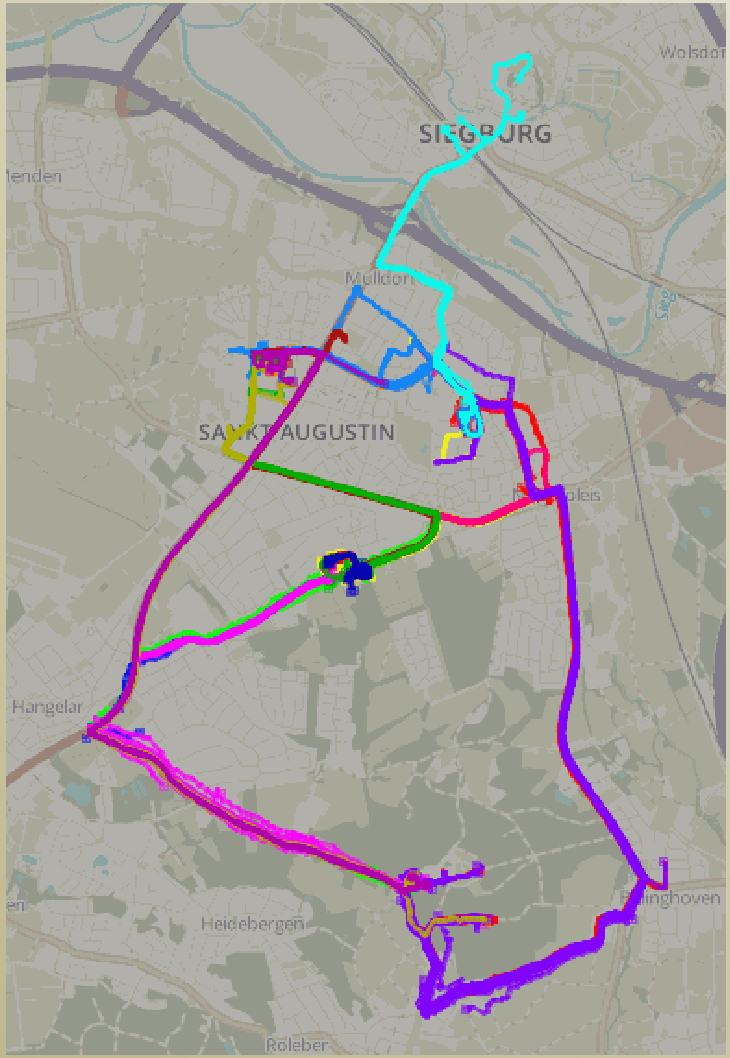
- It may be useful for analysis to divide a long sequence of position records of a mover into sub-sequences according to various criteria.
  - The sub-sequences are also called trajectories. Each (partial) trajectory gets an additional identifier to be distinguished from other trajectories.
- <u>Division into trips</u>
  - Find and mark stops of a suitable duration; then select the sub-sequences between the stops as trajectories representing trips.
  - Enables analysing the routes between the trip origins and destinations and the variation of movement characteristics on the same route.
- <u>Division based on a time cycle</u>
  - Choose an appropriate time cycle (daily, weekly, seasonal, …); choose some position within the cycle; break the trajectory in all places where the chosen cycle position falls between two consecutive points.
  - Enables analysing regular movements.

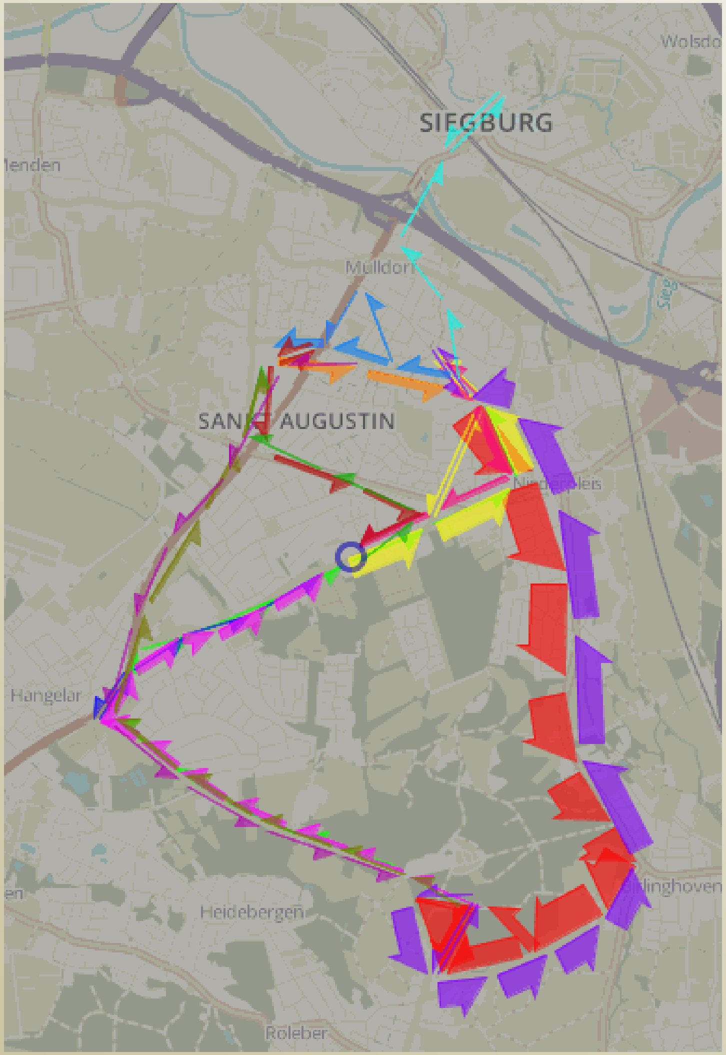
Example: division into trips

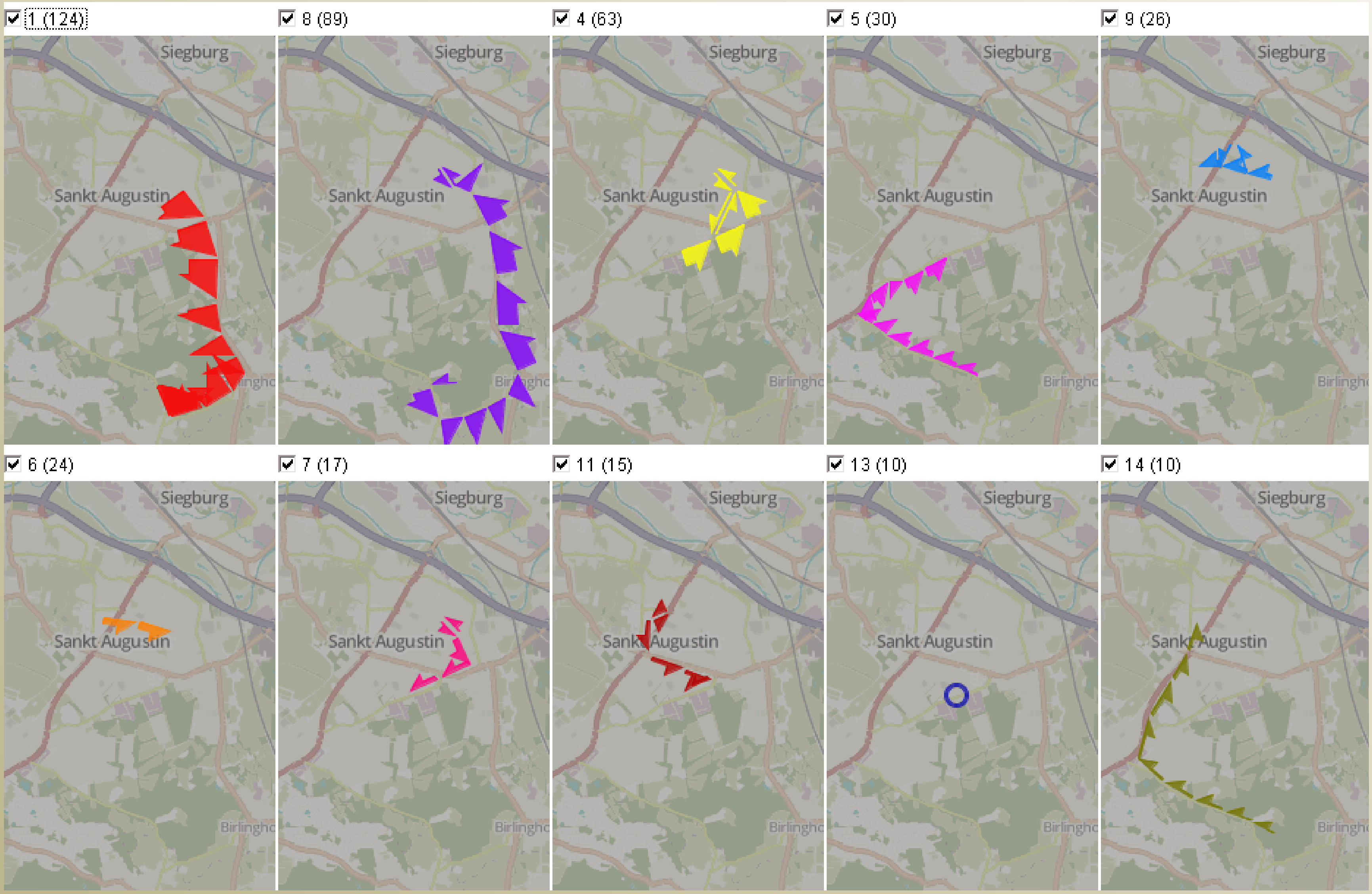


The year-long trajectory of a car has been divided into 585 trip trajectories by stops with duration ≥ 15 minutes. To make the trajectories distinguishable, we have clustered them with DBC by route similarity (using a corresponding distance function; to be discussed later). The noise (25.3%) is hidden.

A summarised representation of the trip routes:

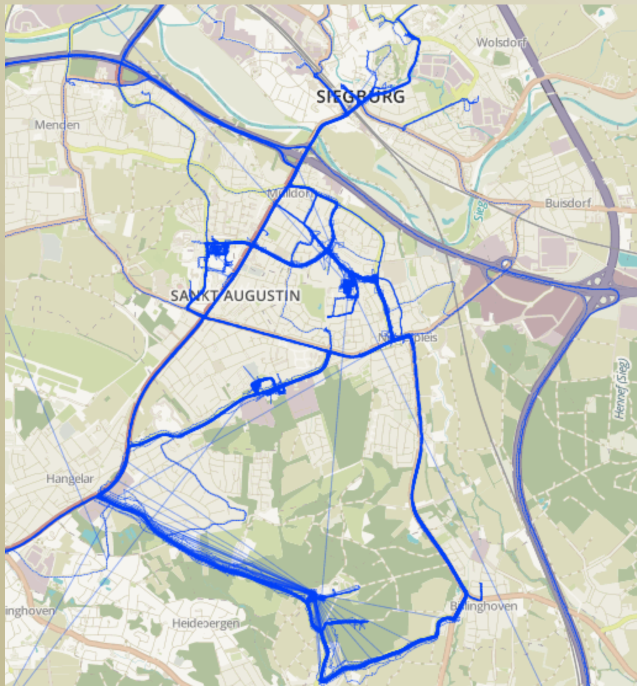


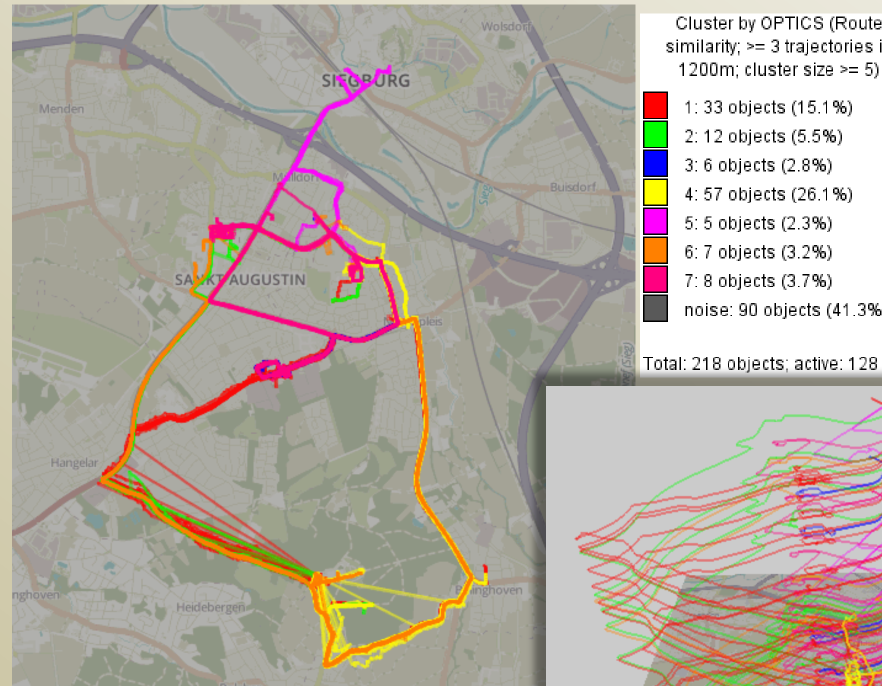




10 most frequent routes

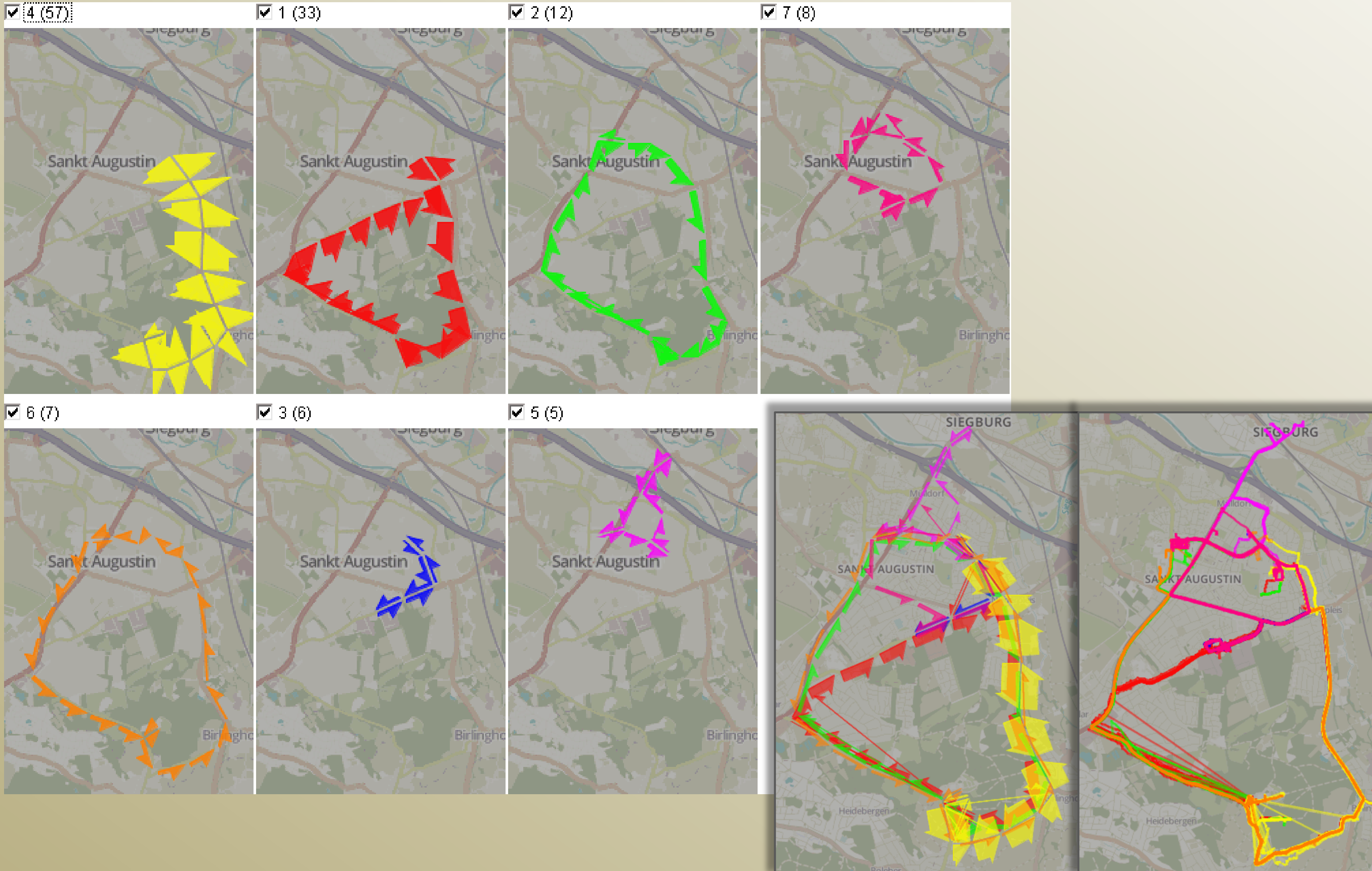
Example: division by time cycle (daily)





The year-long trajectory of a car has been divided into 218 daily trajectories by the daily cycle position 04:00. For distinguishability, we have clustered the trajectories with DBC by route similarity. The noise (41.3%) is hidden.

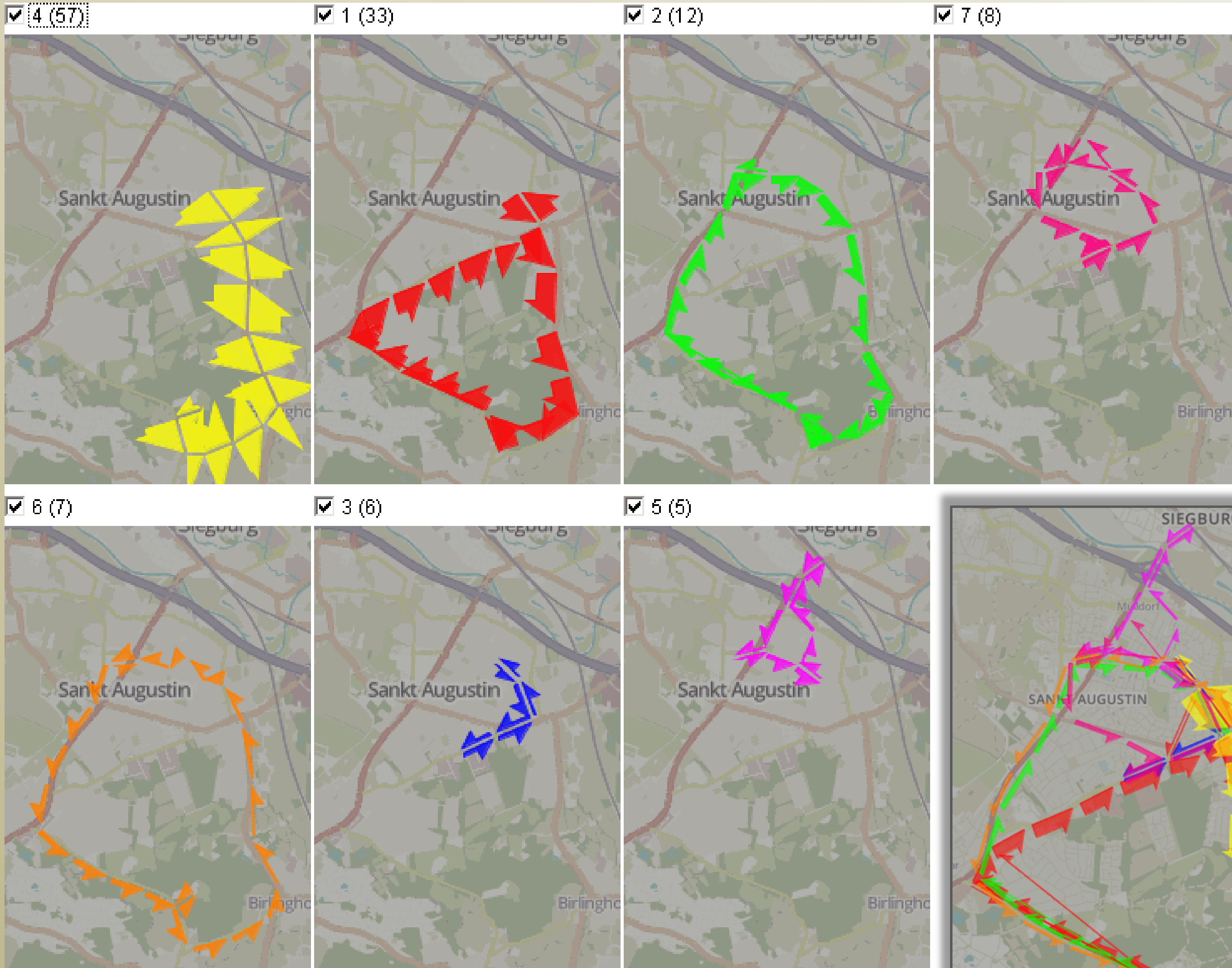
A summarised representation of the regular daily mobility behaviours:

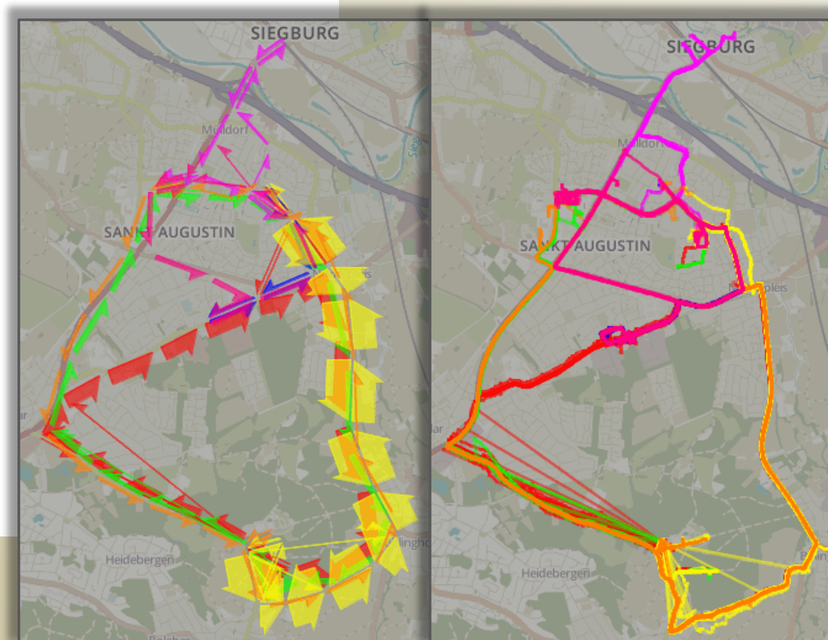


Questions?

Structure and properties of movement data

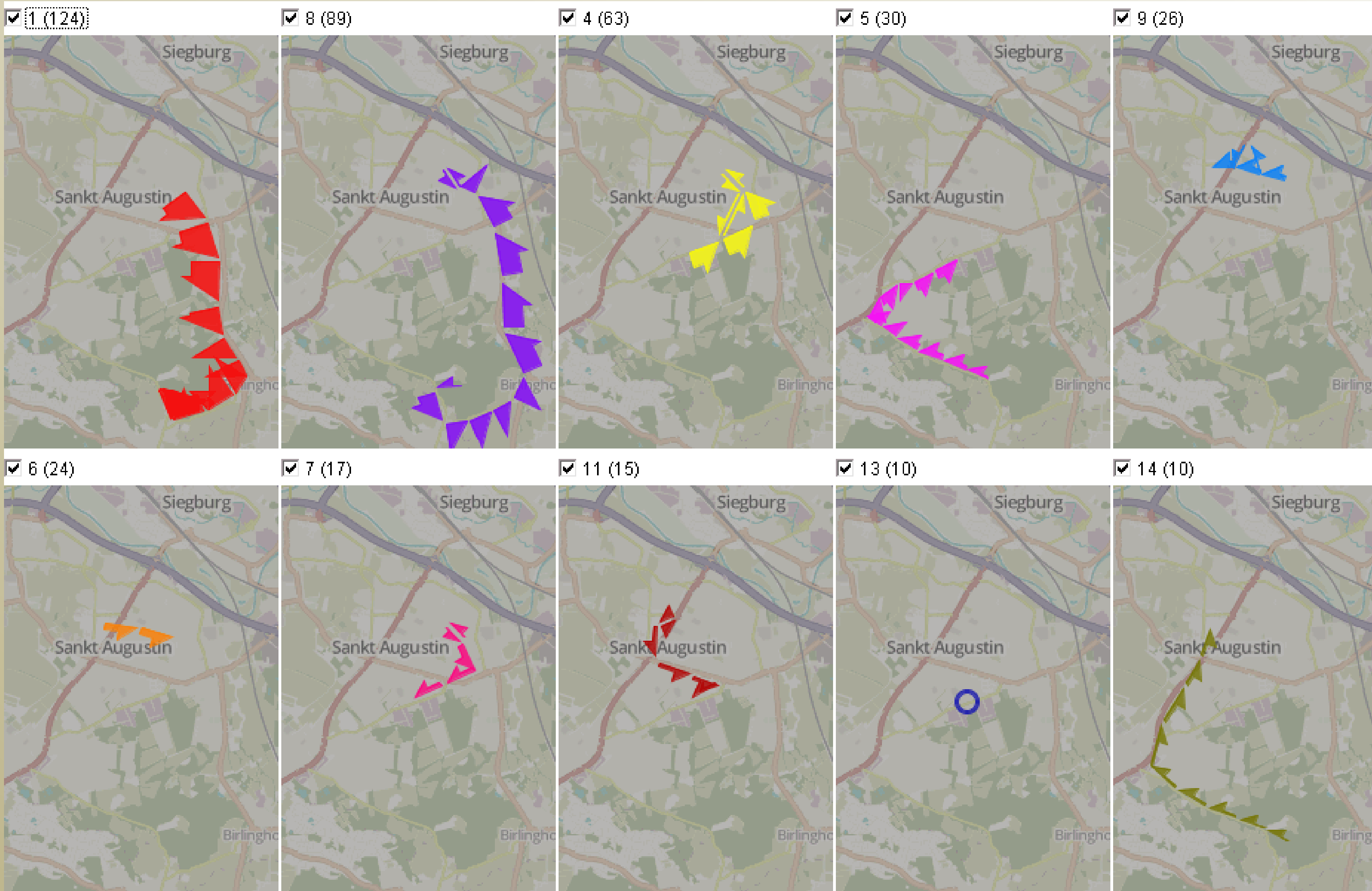
Spatial abstraction and summarisation of trajectories



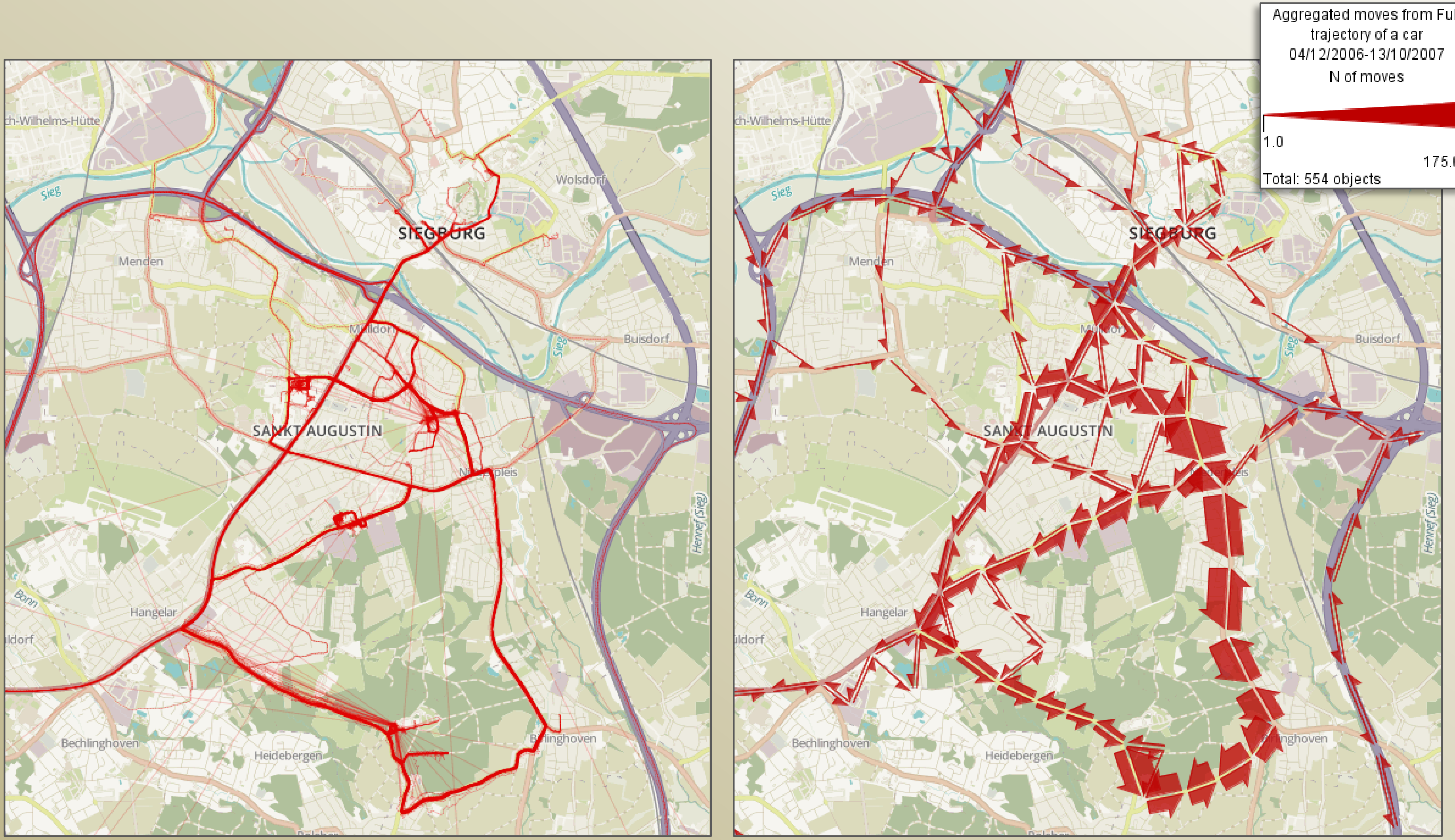


Spatial summaries enable a convenient overall view of multiple trajectories.

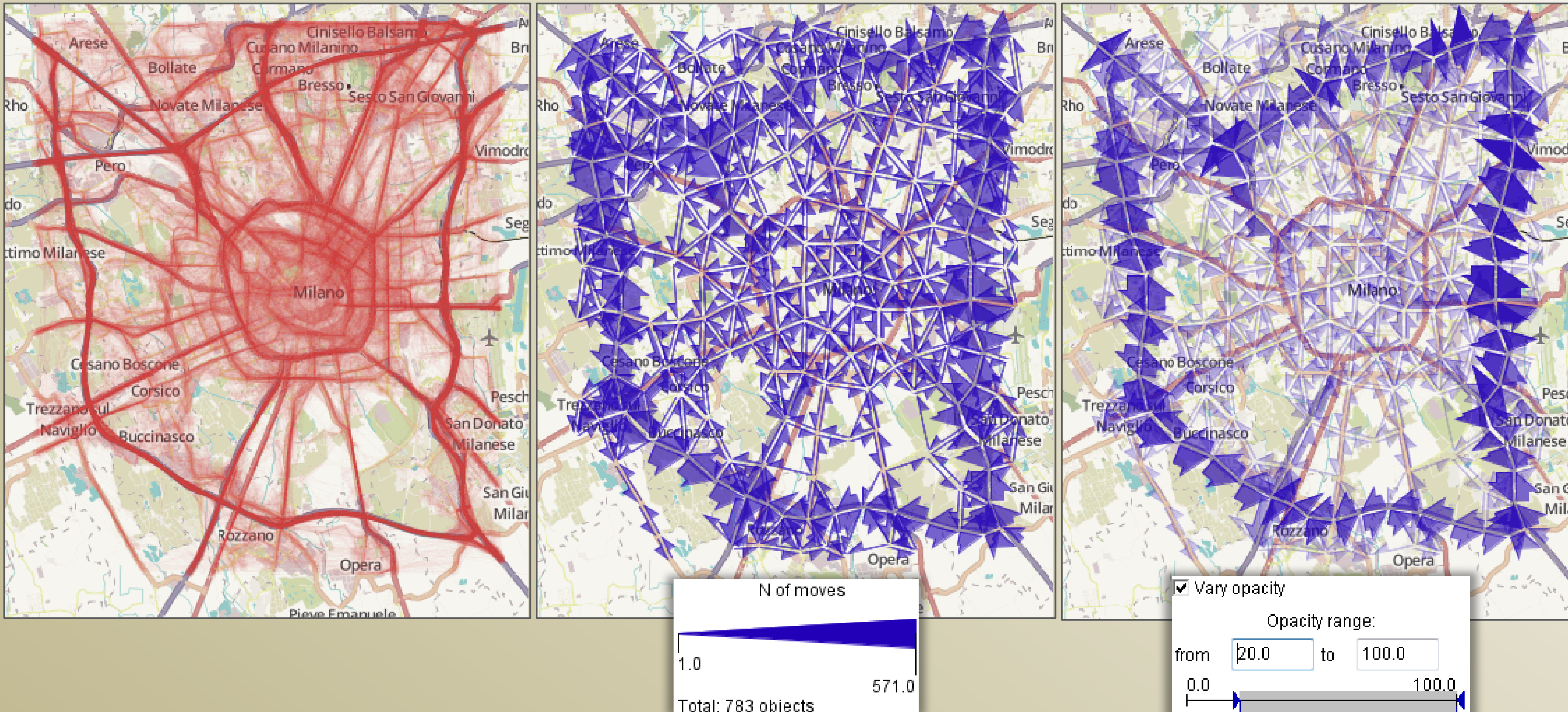
We can easily see movement directions and relative frequencies of movement in different places, which are not visible in maps with individual trajectories.



Spatial abstraction and summarisation gives compact representation of groups (clusters) of trajectories.

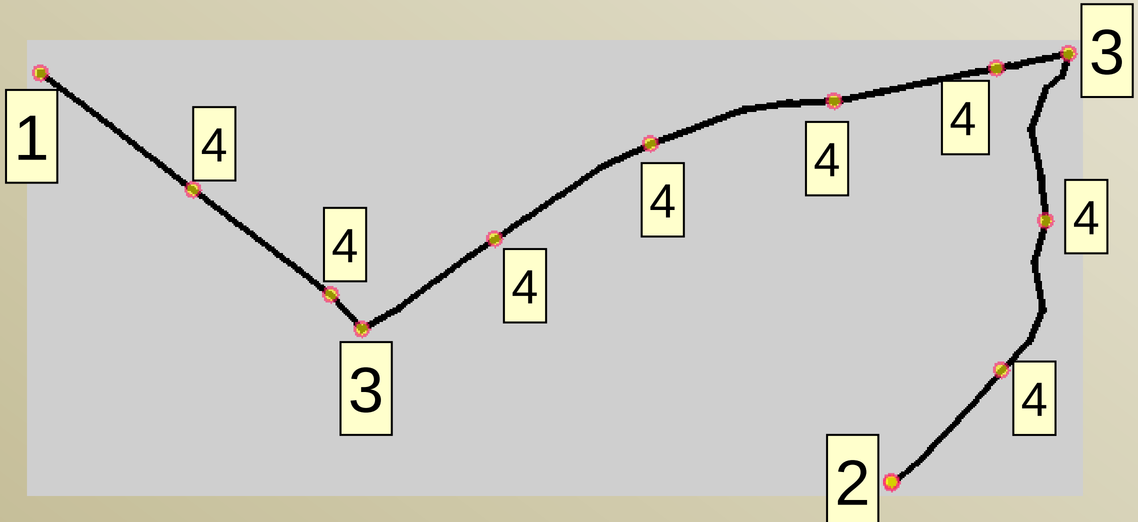


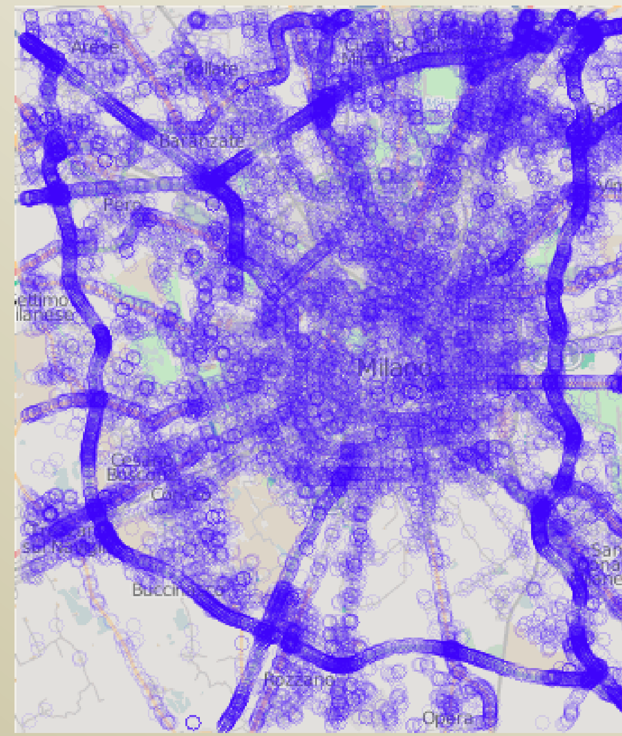
Spatial abstraction and summarisation also provides an overview of a whole set of trajectories of any size.



When the movements over a territory are very dense, even a summarised view may look cluttered. Variation of the opacity of the flow symbols makes it better readable. This example: a summarised view of 6,731 daily car trajectories.

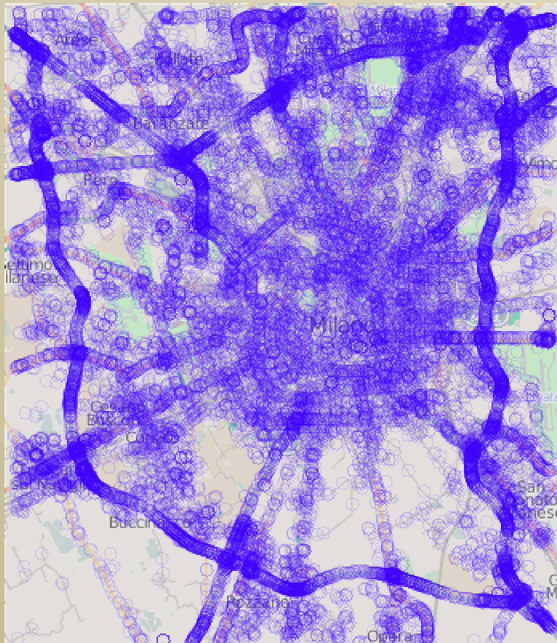
Spatial abstraction of trajectories: how?

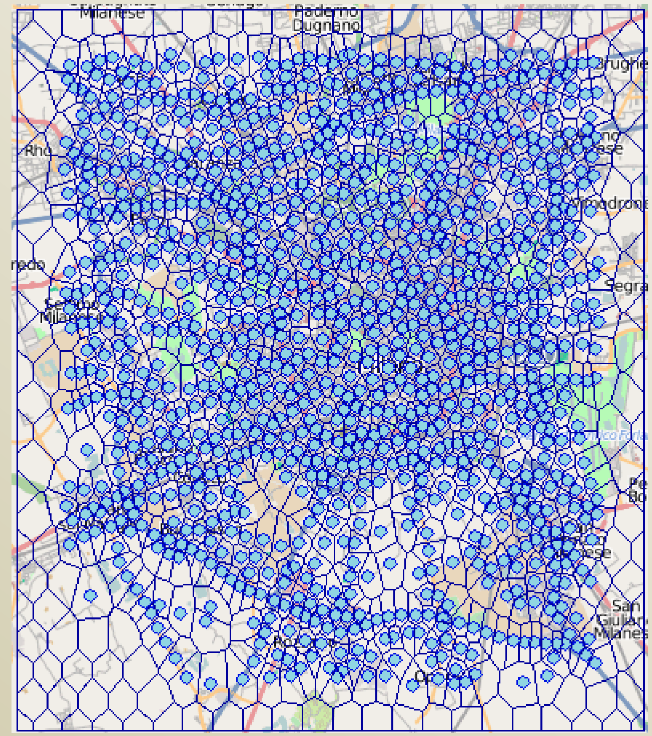




1. Extract characteristic points from the trajectories. Characteristic points include starts (1), ends (2), points of significant turns (3), points of significant stops, and representative points from long straight segments (4).

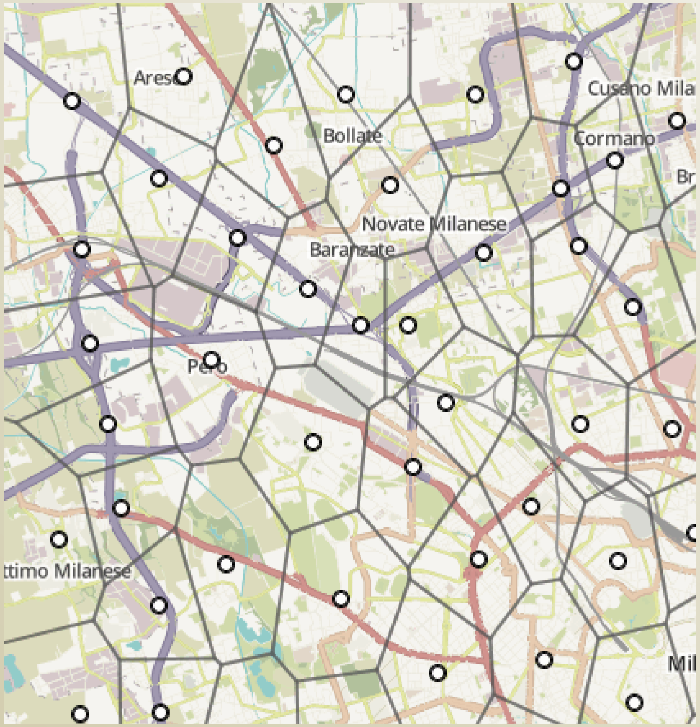
Division of the territory



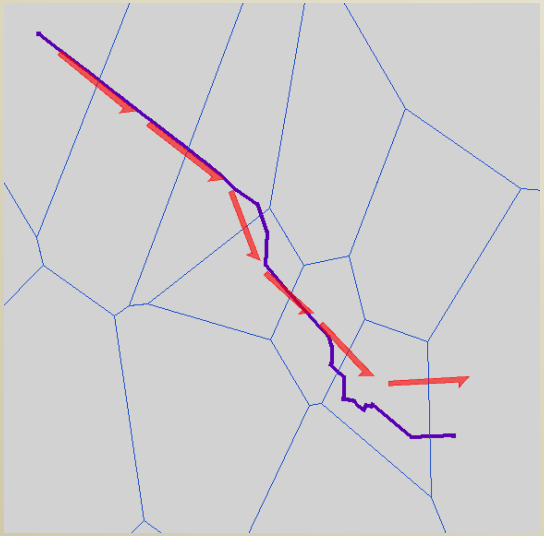


2. Group the extracted points in space so that the groups have a desired spatial extent. The extent is specified by the parameter MaxRadius*, which determines the degree of the generalization.
3. Use the group centres as seeds for Voronoi tessellation of the territory.

* Each group must fit in a circle with the maximal radius MaxRadius.

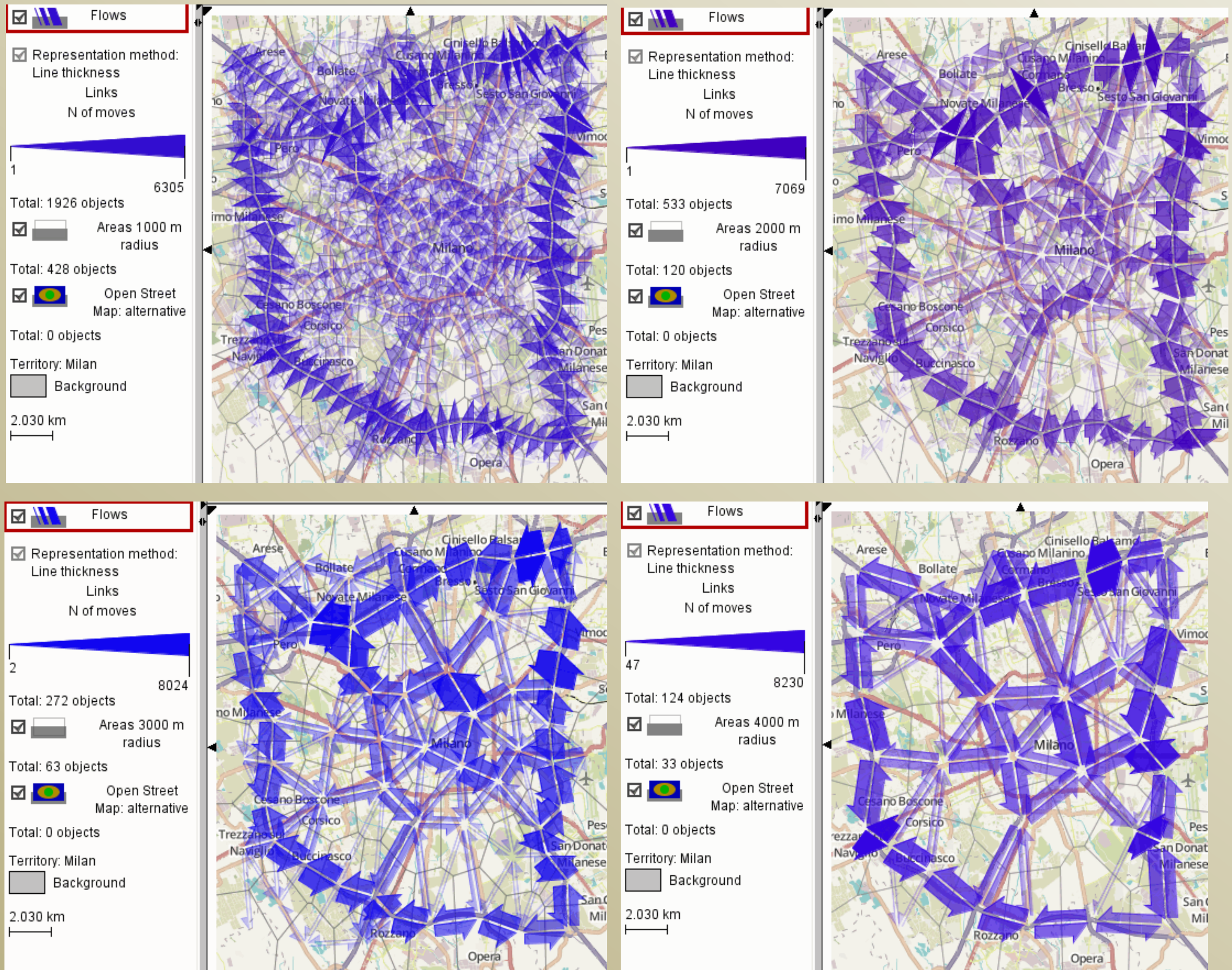


A larger view of point clusters and territory division

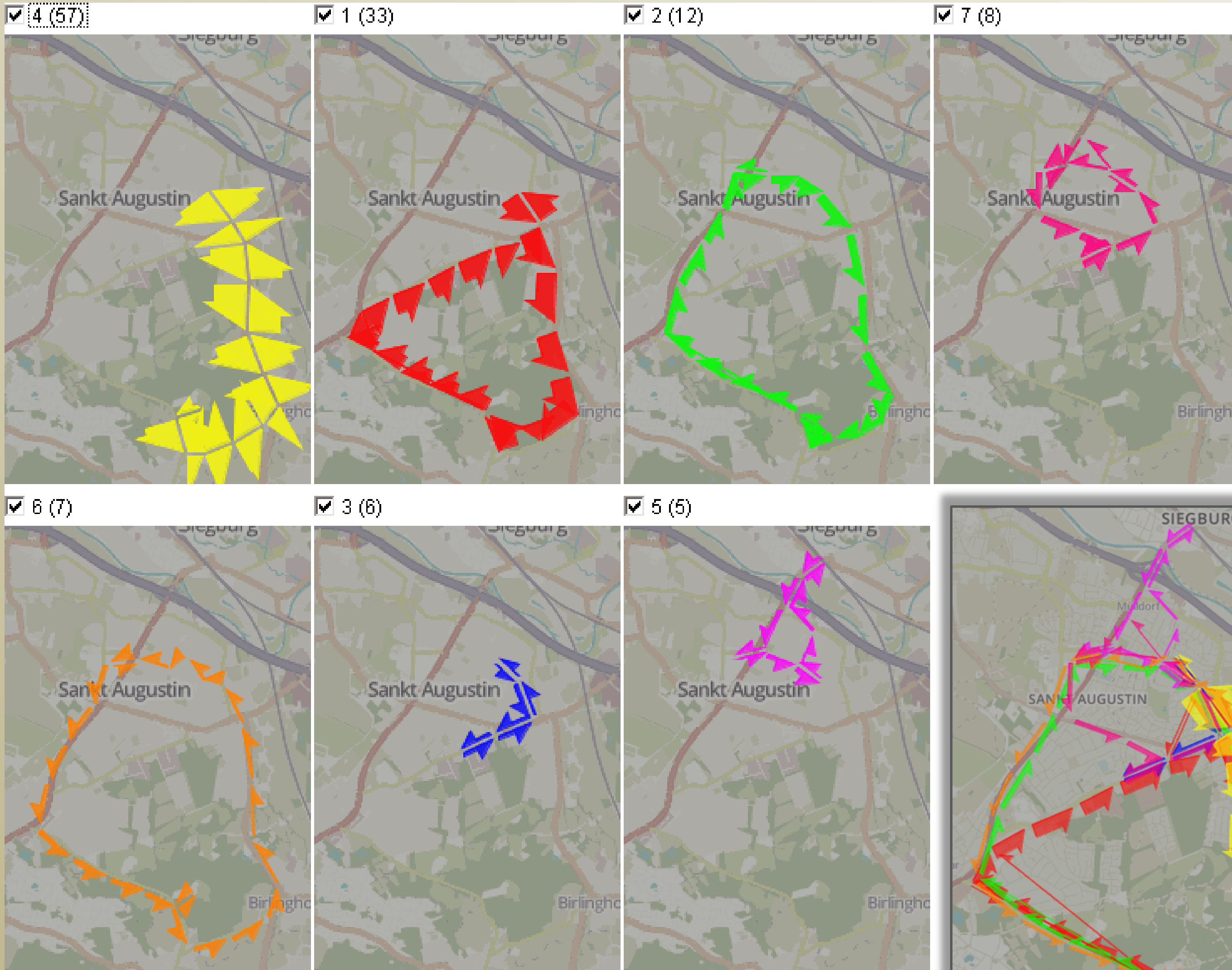


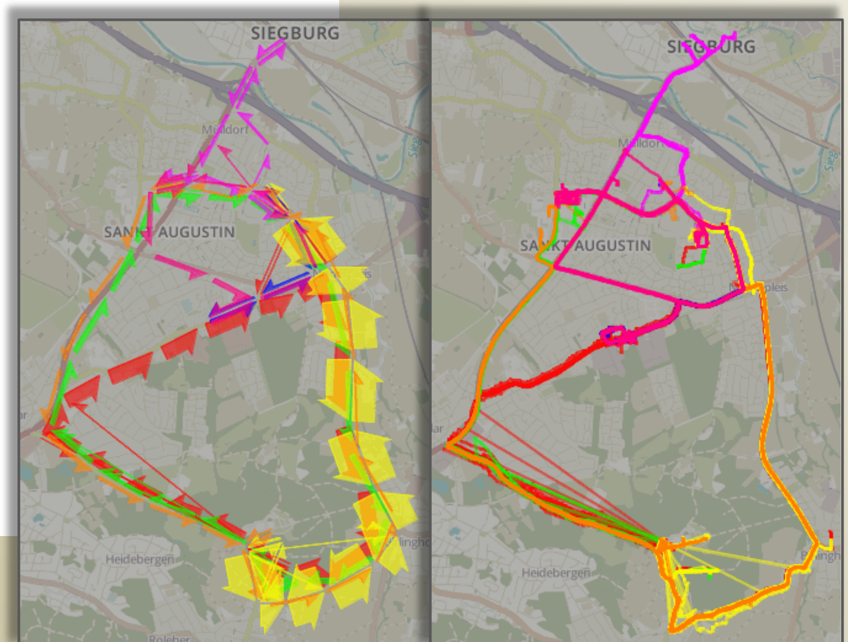
4. Divide each trajectory into segments that link Voronoi cells.
5. For each pair of cells, aggregate the linking segments from all trajectories.
6. Represent the aggregates by flow symbols.

Different levels of spatial abstraction



The parameter MaxRadius allows choosing a suitable level of spatial abstraction, depending on how much detail is needed.





Summarization of groups (clusters) of trajectories:

The approach is applied separately to each group.

Where to read more

Natalia Andrienko, Gennady Andrienko

Spatial Generalization and Aggregation of Massive Movement Data

IEEE Transactions on Visualization and Computer Graphics (TVCG), 2011, v.17 (2), pp.205-219

http://doi.ieeecomputersociety.org/10.1109/TVCG.2010.44

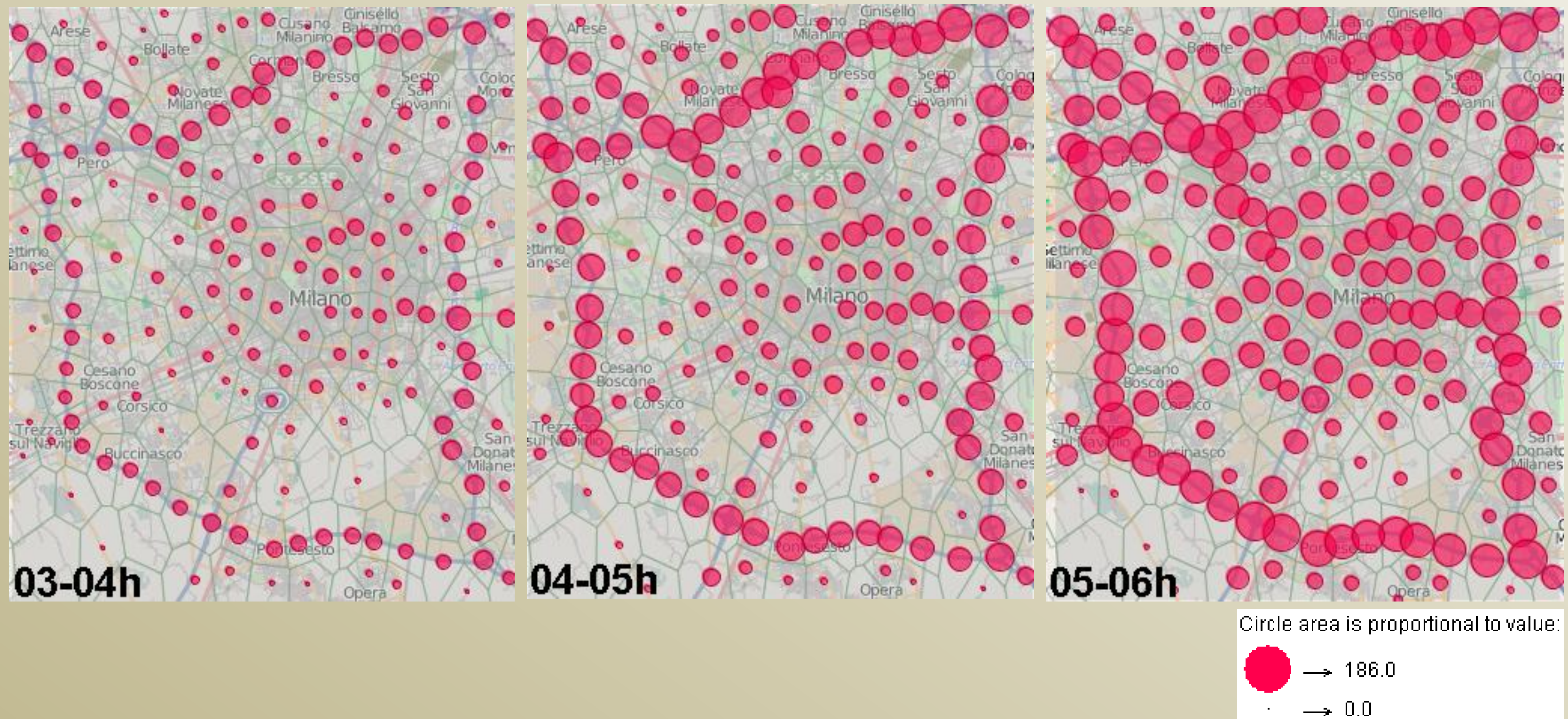
Note 1: The clustering and tessellation method described in the paper is applicable not only to points from trajectories but to any points, e.g., Twitter events, bike docking stations, … - recall the previous lectures and exercises!

Note 2: When the whole set of trajectories does not fit in the RAM of the computer, a random sample of points can be taken from a database and used for creating a tessellation. This tessellation can then be used for aggregating the data in the database.

Note 3: The tessellation can also be used for spatio-temporal aggregation.

Aggregation by space and time: cells

…

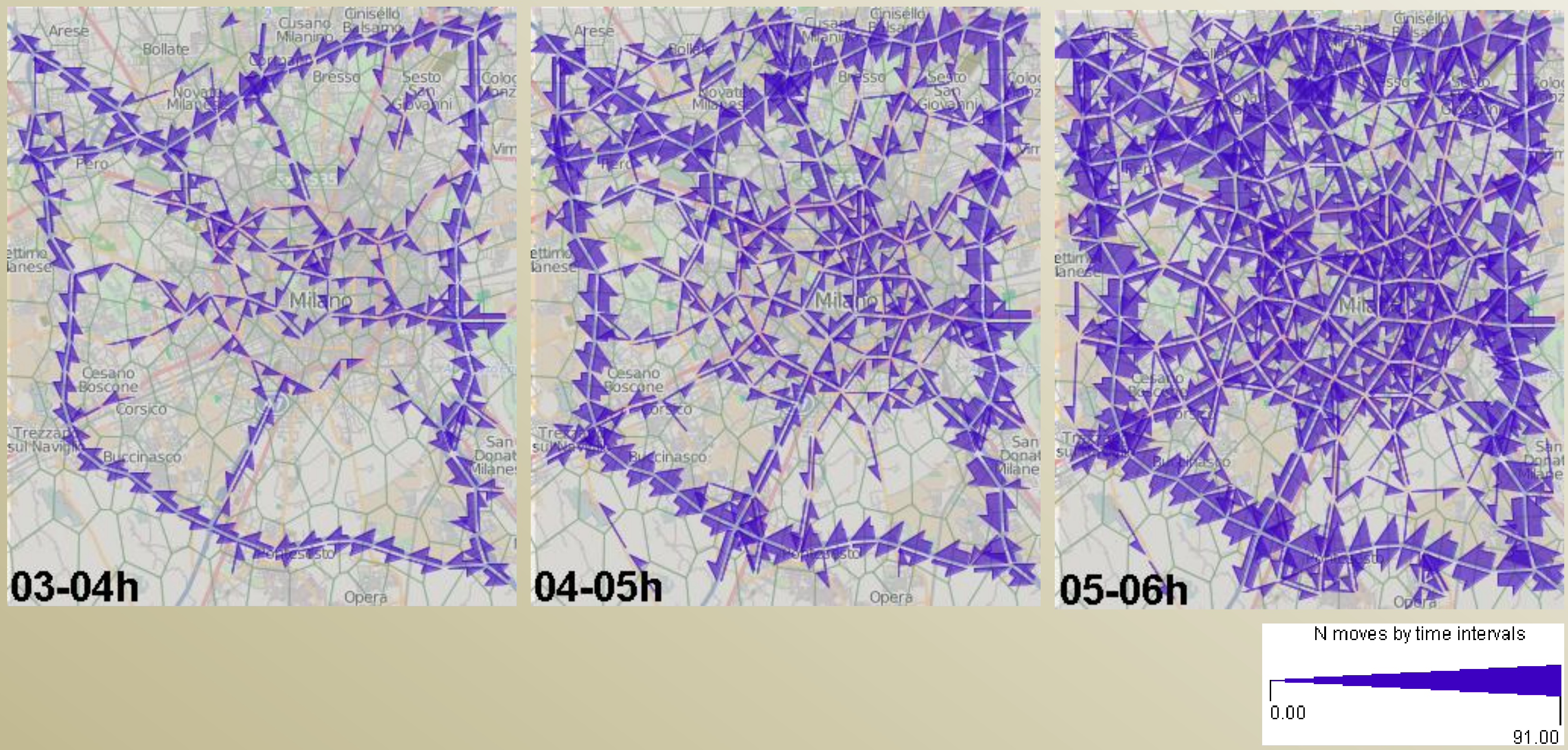


…

Spatial situations: presence

Aggregation by space and time: links

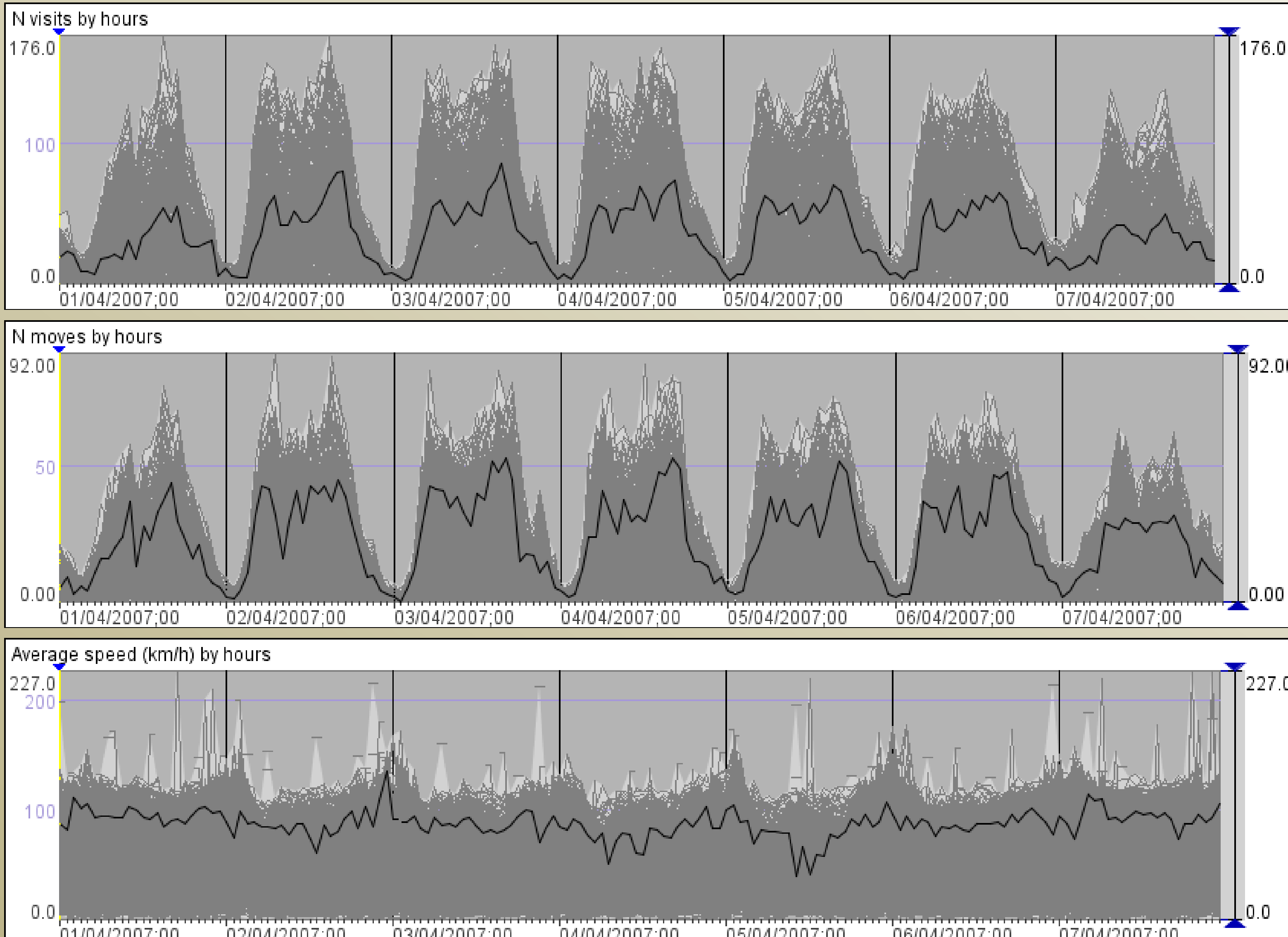
…



…

Spatial situations: flows

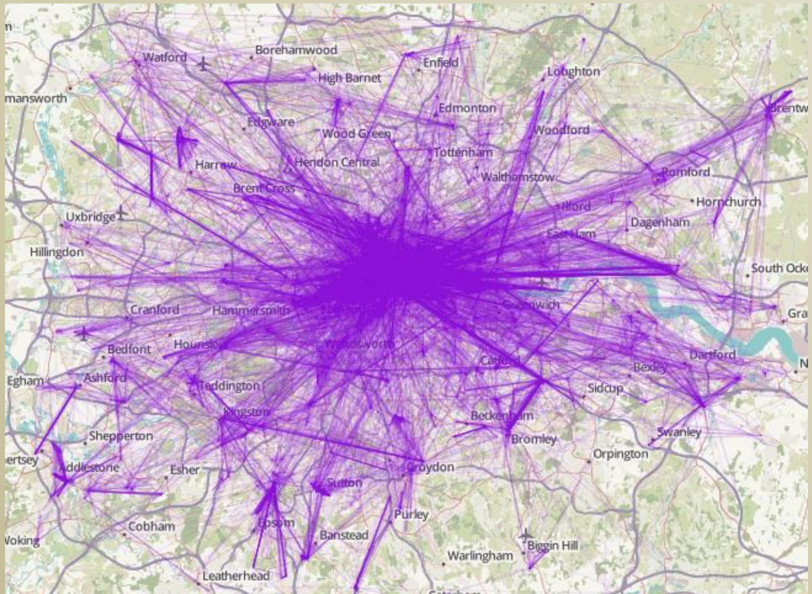
Local time series for the cells and links



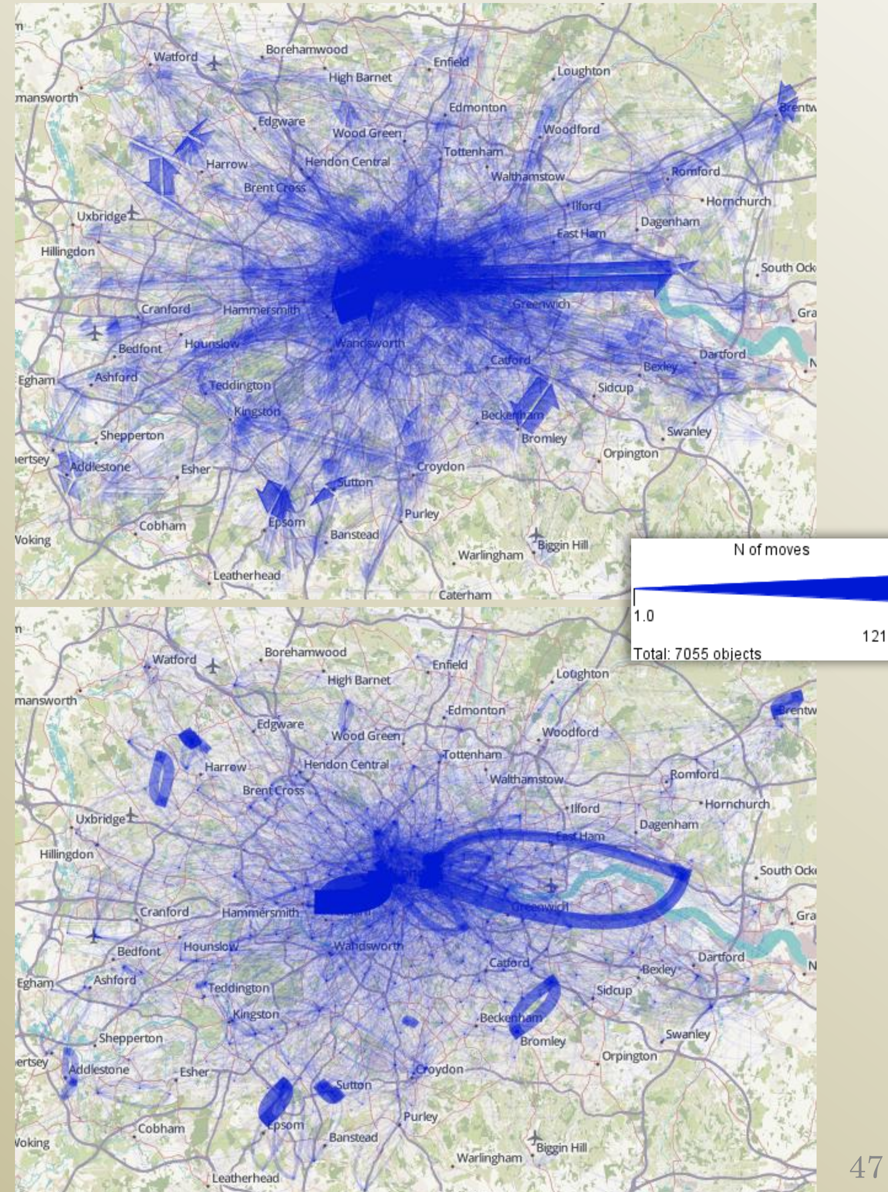
Summarisation of episodic trajectories

- The territory division is done in the same way as for quasi-continuous trajectories.
- When the trajectories are transformed into segments connecting cells, two consecutive points may fall in non-neighbouring cells.
  - Building a path through neighbouring cells by interpolation is invalid!
- The aggregation result will include links going across several (sometimes many) cells.
  - Computation of some aggregates (mean speed, mean transition duration, …) is not meaningful.
- Flow maps are very cluttered due to numerous crossings and overplotting of flow symbols.
  - Analogously to flow maps of aggregated OD moves (recall from the previous lecture).

Example of summarisation of episodic trajectories

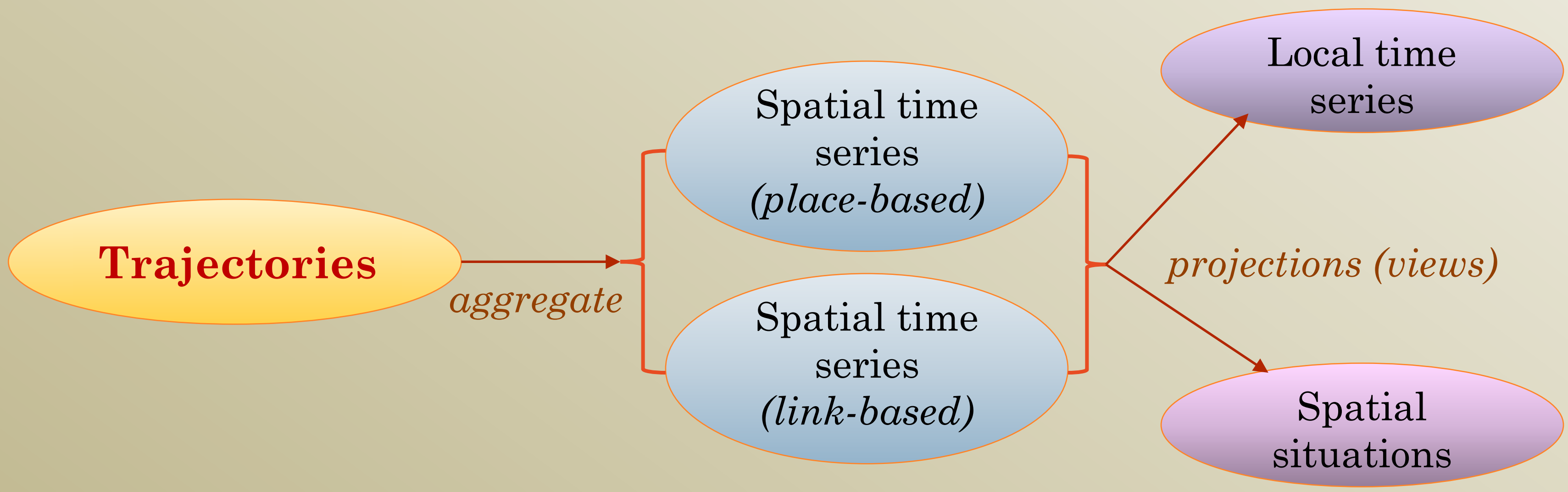


A sample of 214 trajectories of Twitter users (shown with 10% opacity) has been aggregated with MaxRadius = 2km. The flows are shown with varying opacity from 5% to 100%. Curved flow symbols may be better in such flow maps.



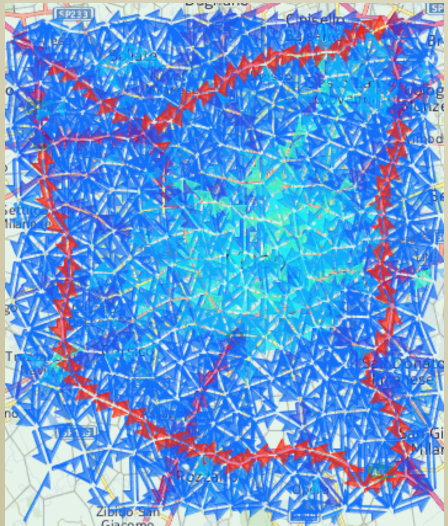
Summarization and aggregation of trajectories

- Can also be done by pre-existing territory division.
- Results are analogous to results of aggregation of OD moves.



☺ You already know how to deal with spatial time series!

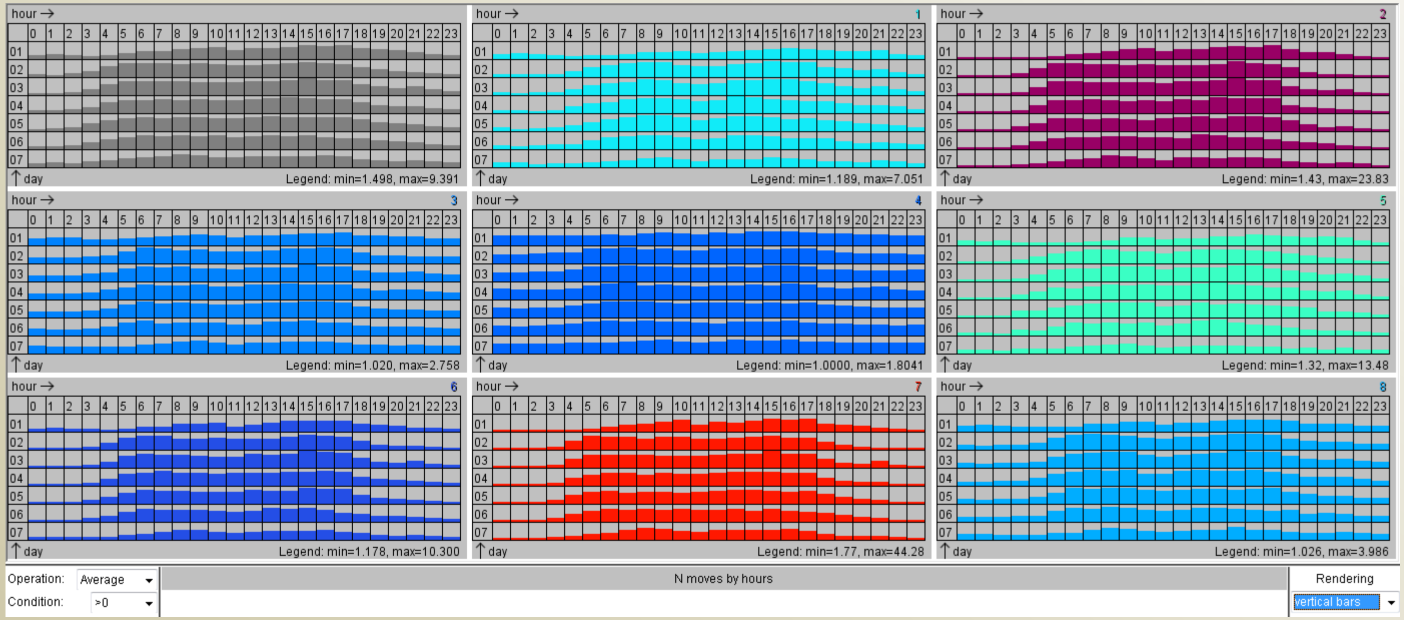
Example: analysis of link-based local time series using PBC

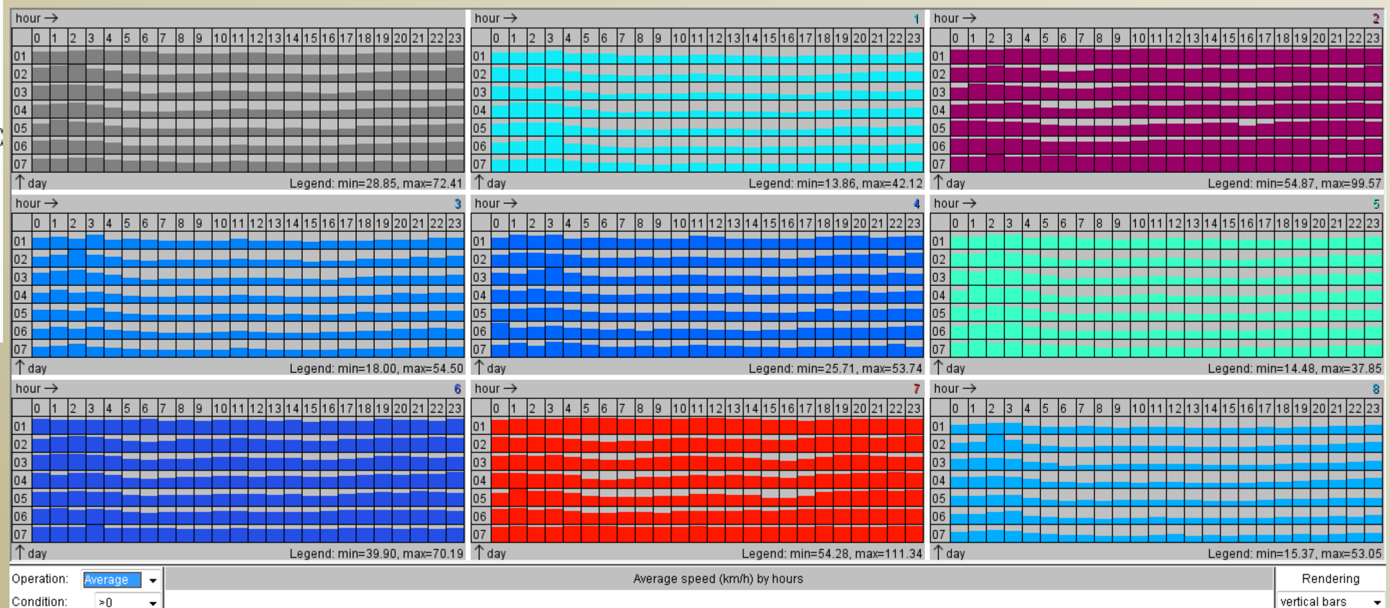


Aggregated moves of cars
Clusters 1

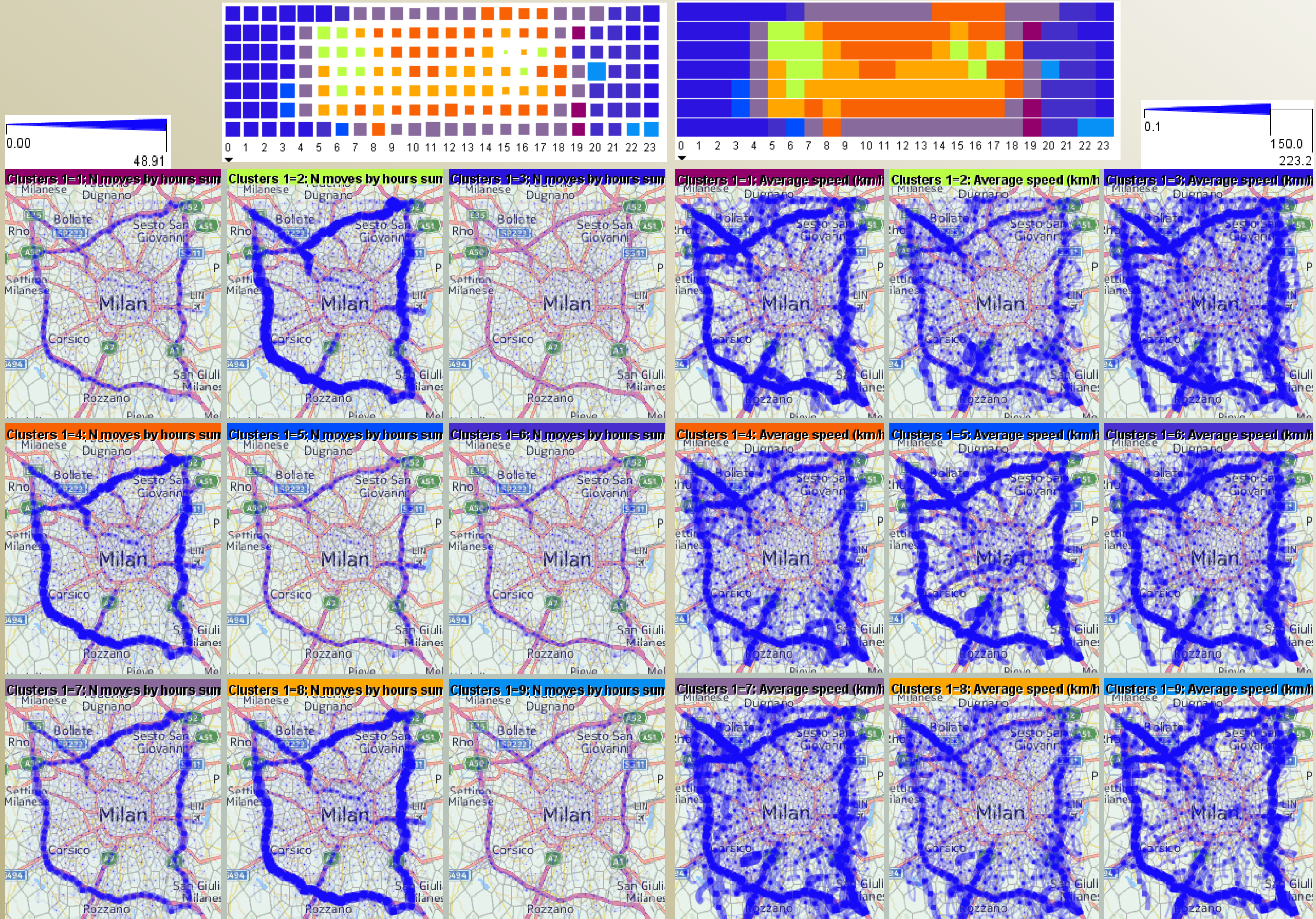
- 1: 138 objects (6.4%)
- 2: 82 objects (3.8%)
- 3: 334 objects (15.5%)
- 4: 1004 objects (46.6%)
- 5: 69 objects (3.2%)
- 6: 160 objects (7.4%)
- 7: 138 objects (6.4%)
- 8: 230 objects (10.7%)

Total: 2155 objects





Example: analysis of spatial situations in terms of the flows using PBC



Questions?

Spatial abstraction and aggregation of trajectories

Extraction of movement events from trajectories

Examples of movement events (m-events)

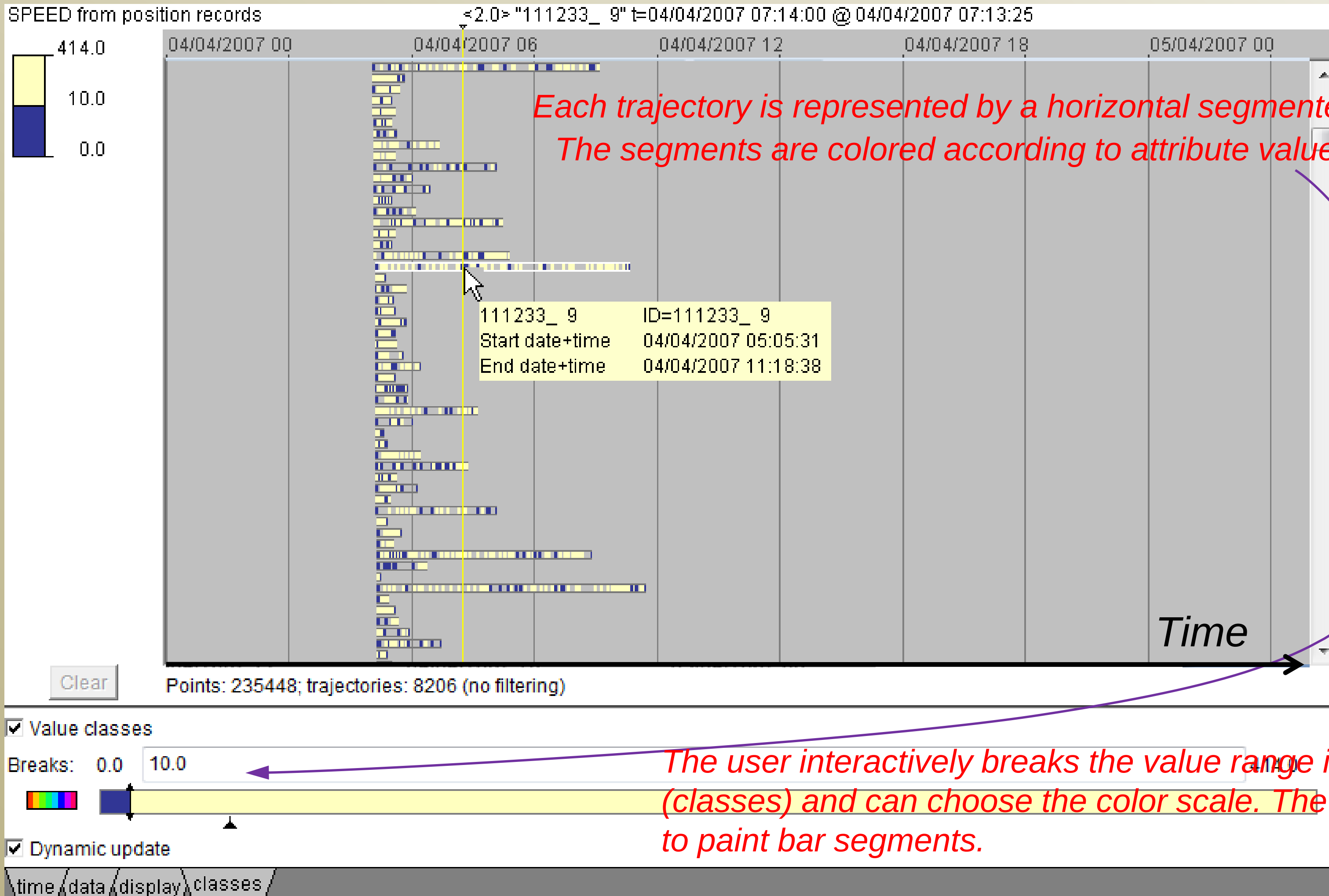
- Stop *(considered earlier)*
- Low-speed driving
- Turn
- High acceleration
- Take-off / landing of an aircraft
- Meeting of two or more moving objects
- Driving late at night
- Stop at a particular place of interest
- Leaving stadium after a football game
- High heart rate {during jogging}

M-events are defined based on values of attributes attached to trajectory positions

- Instant speed, travelled path in time window / from the beginning of the trip
- *Bounding box diagonal*
- *Sinuosity in a time window*
- Heart rate, body temperature…
- Time of day, day of week of trajectory points
- Relationship to places, spatial objects, and events measured as
  - Spatial distance to n^{th} nearest place/object
  - Temporal distance to n^{th} nearest event
  - Neighborhood (counts of objects or events in given S,T,ST windows)

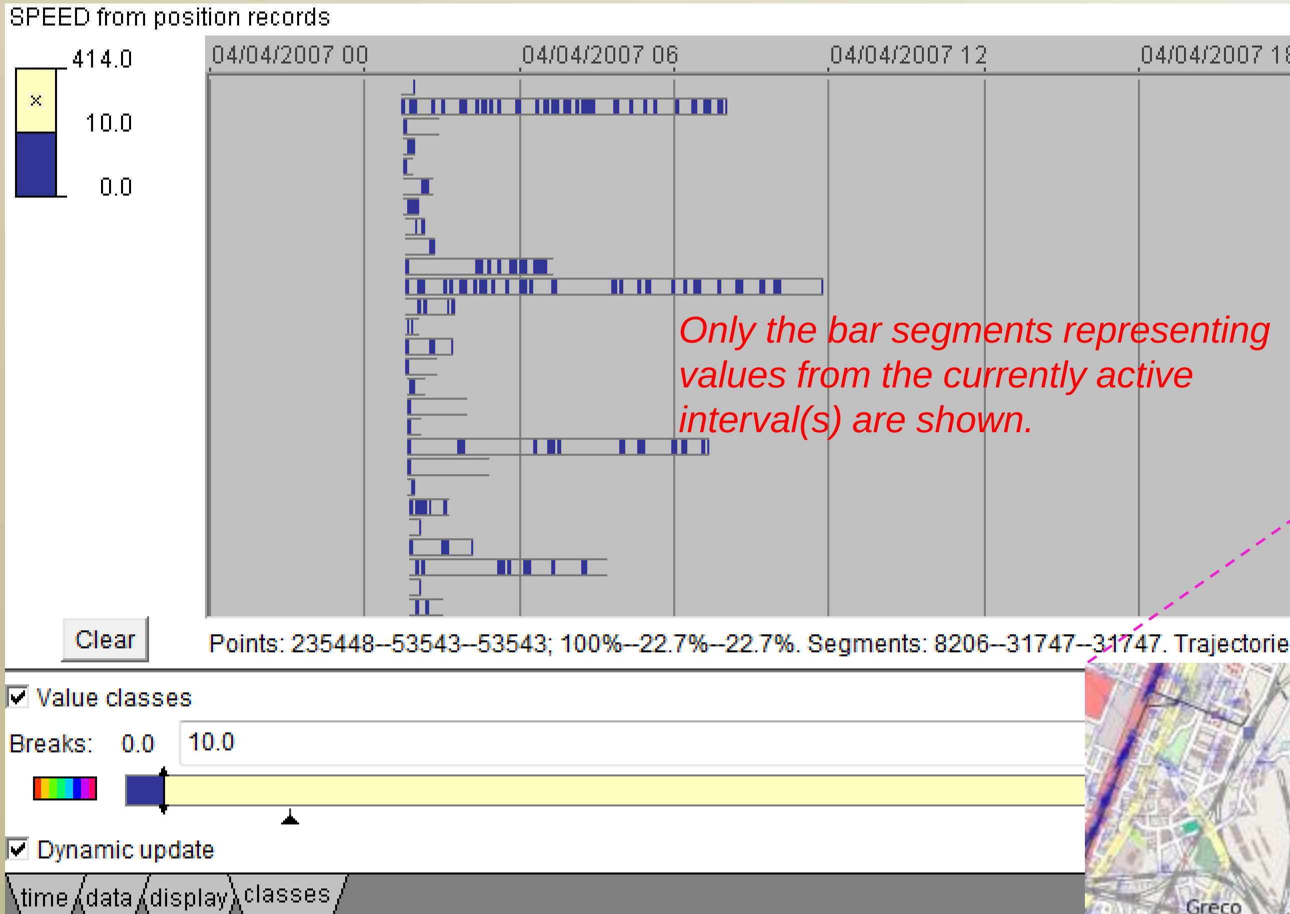
- Most of these attributes can be computed from the position records.

Extracting m-events from trajectories by interactive operations

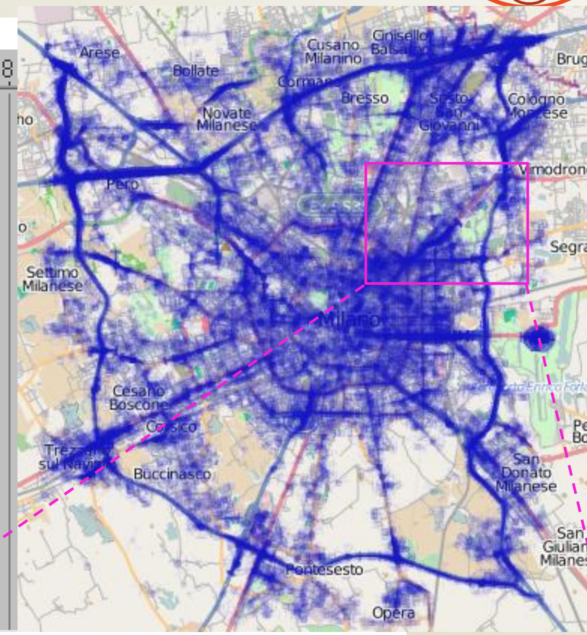


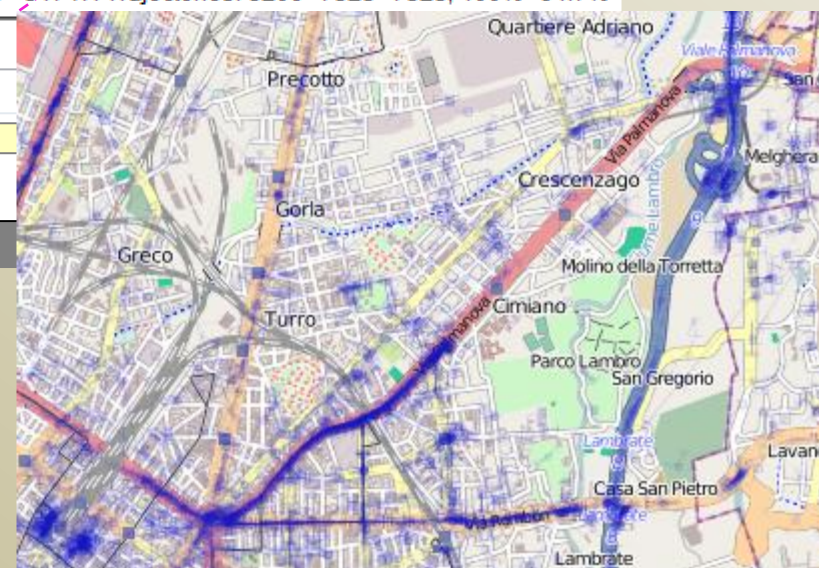
Each trajectory is represented by a horizontal segmented bar. The segments are colored according to attribute values.

The user interactively breaks the value range into intervals (classes) and can choose the color scale. The colors are used to paint bar segments.



Only the bar segments representing values from the currently active interval(s) are shown.

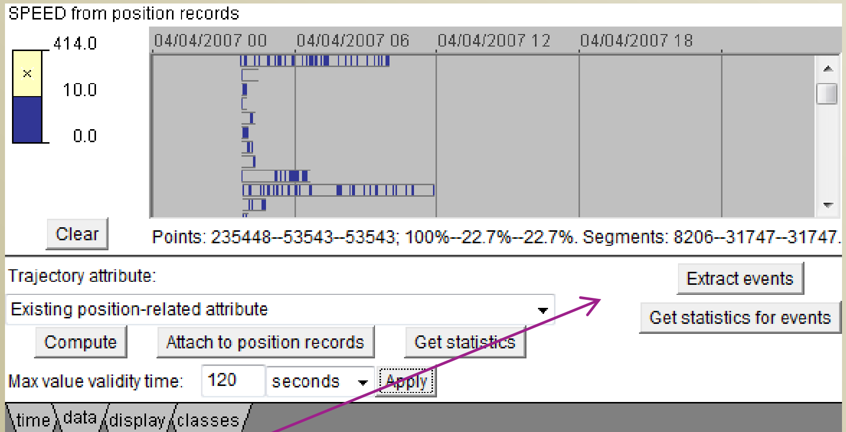




The map shows only the points and segments of the trajectories where the values of the dynamic attribute satisfy the filter.

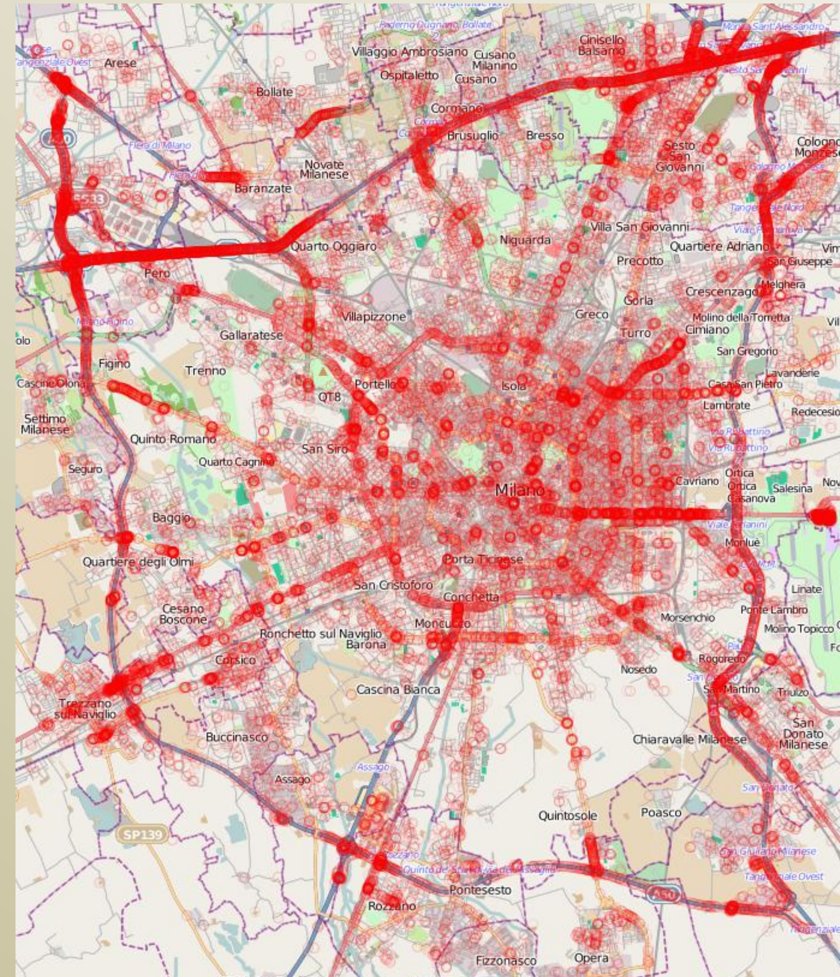
Here we see the points and segments where the speed was not more than 10 km/h.

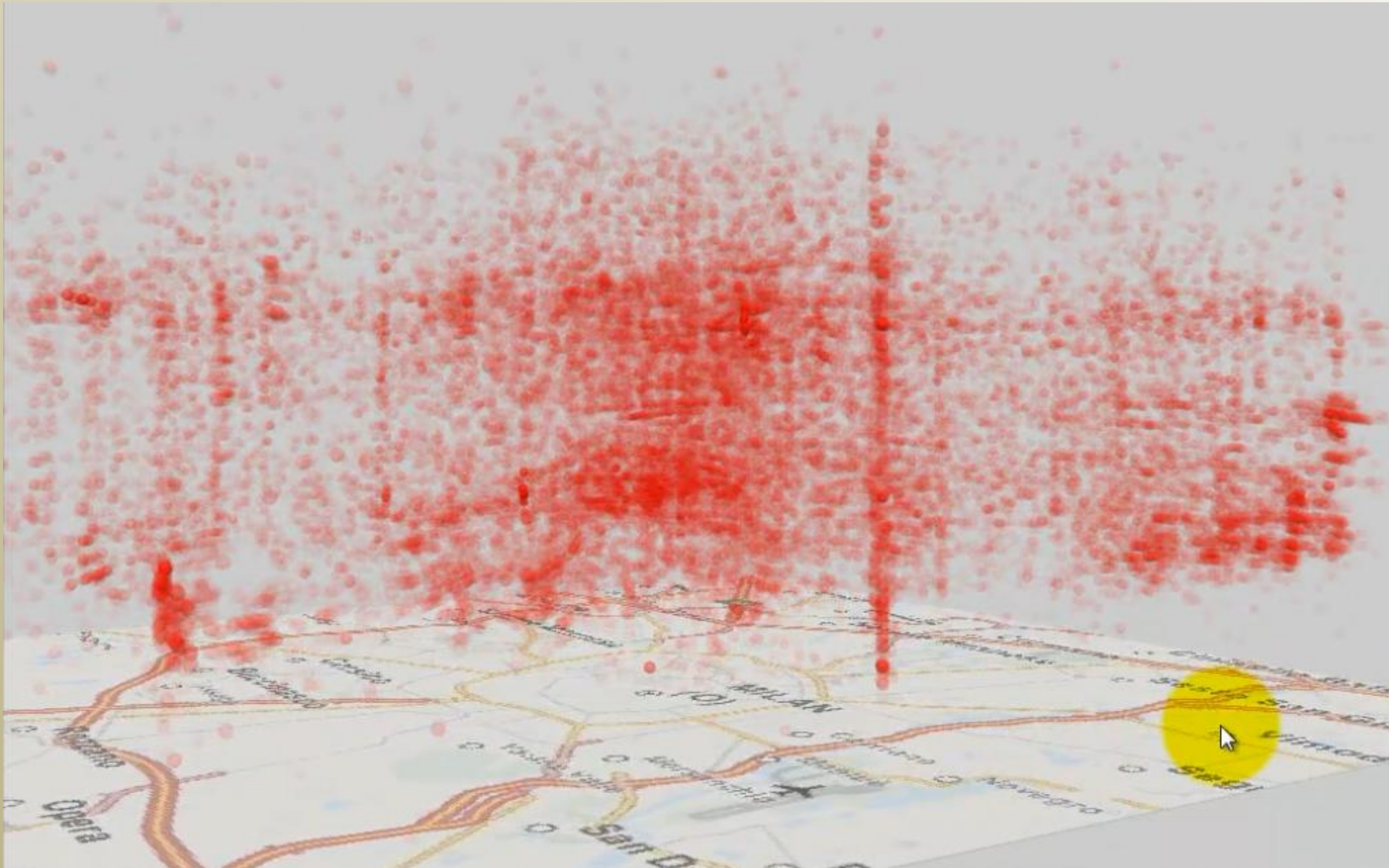
Event extraction based on the segment filter



Pressing a special button extracts m-events from the trajectories according to the current segment filter. The extracted events are organized in a new dataset consisting of points and multi-points with time references and attributes.

The map shows the extracted low speed events as an independent map layer. The m-events are represented by red hollow circles.





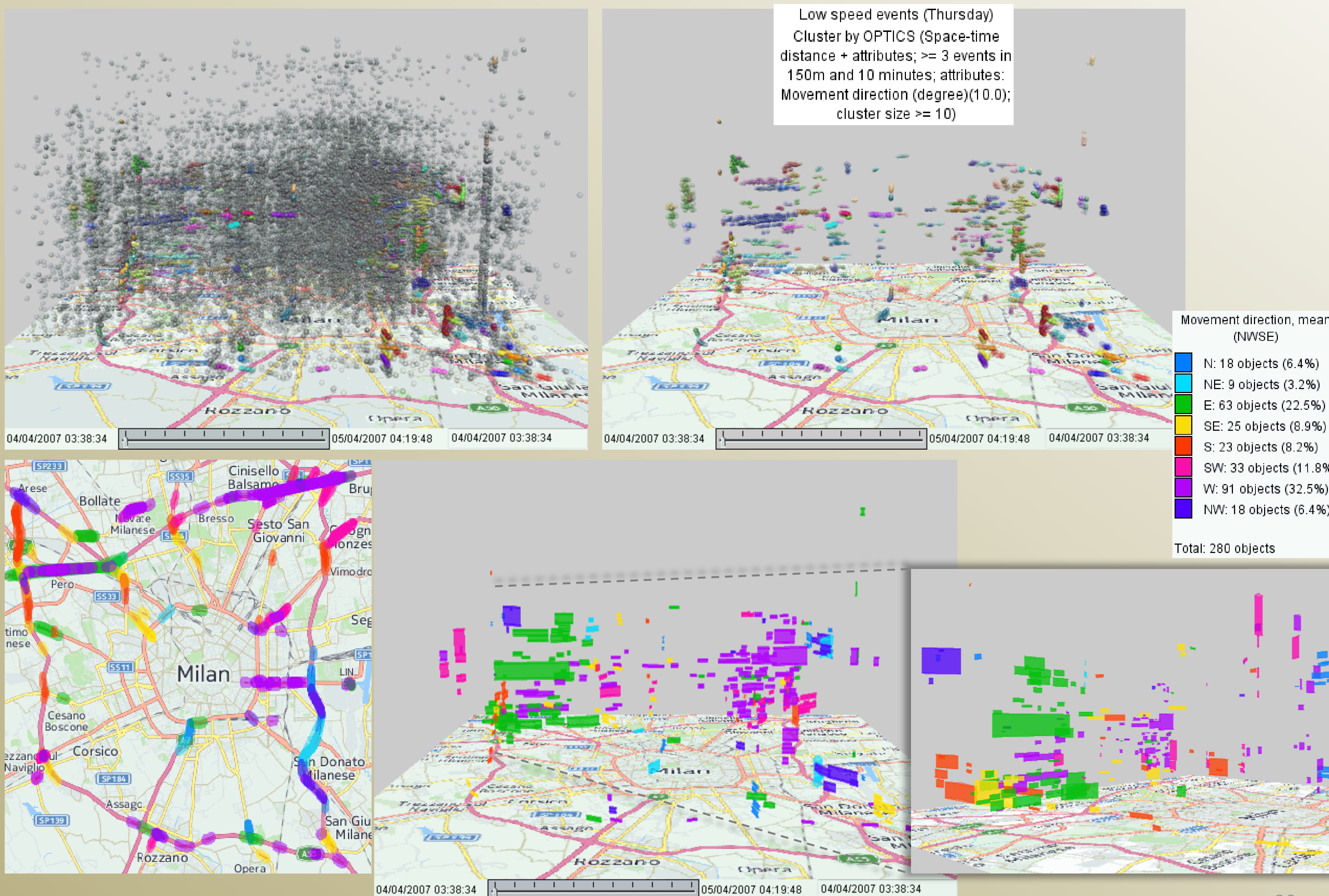
The extracted spatial events are represented in a space-time cube.

Extraction of m-events from trajectories: further notes

- Can be done based on a combination of segment filters, e.g., by the bounding box diagonal and sinuosity *(recall from this lecture)*.
- Can be done not only interactively but also using database queries.
- Analysis of the extracted m-events: use all methods suitable for spatial events.

  ☺ You already know some of them!

Example: analysis of low speed events using DBC (space + time + direction)



Where to read more

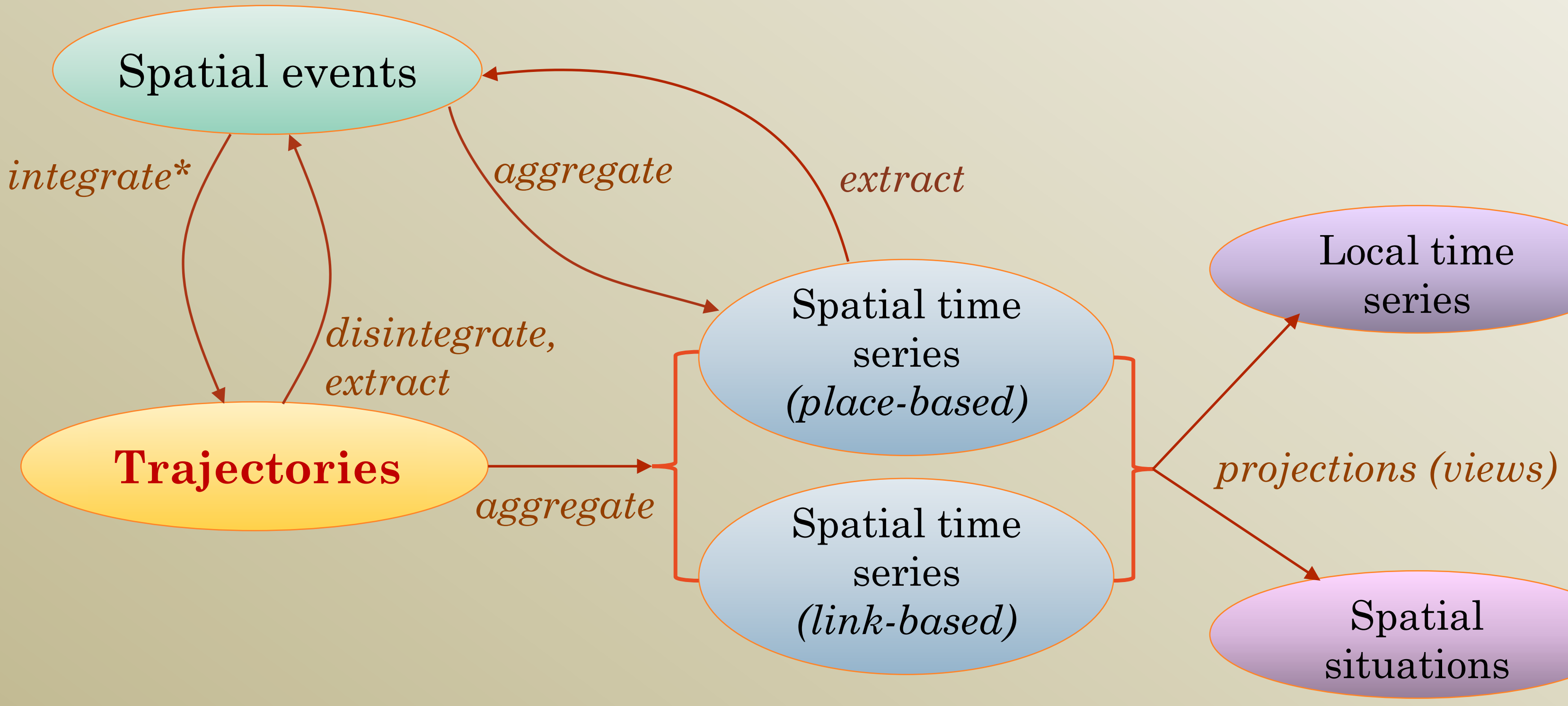
Gennady Andrienko, Natalia Andrienko, Christophe Hurter, Salvatore Rinzivillo, Stefan Wrobel

Scalable Analysis of Movement Data for Extracting and Exploring Significant Places

IEEE Transactions on Visualization and Computer Graphics (TVCG), 2013, v.19 (7), pp. 1078-1094

http://dx.doi.org/10.1109/TVCG.2012.311

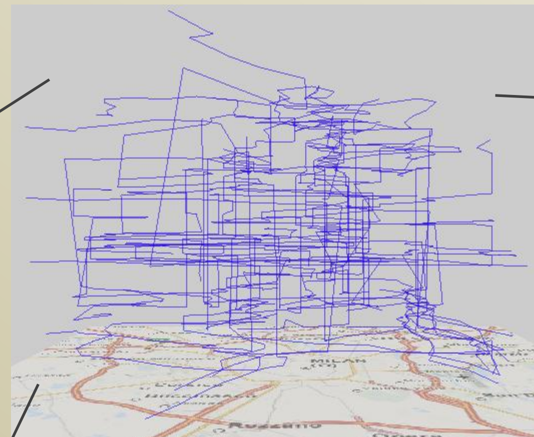
Transformations of trajectories



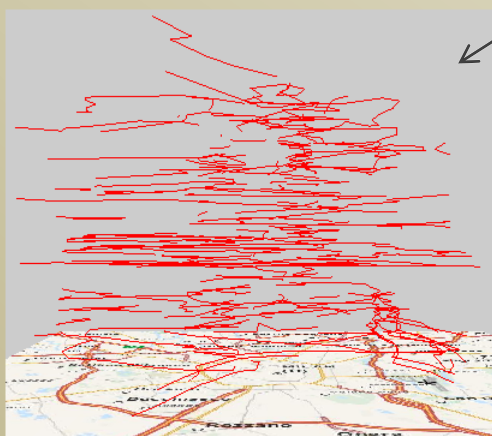
* Any trajectory is composed of spatial events, i.e., each position record represents a spatial event. This is especially clear when trajectories are reconstructed from tweets, phone calls, RFID or Bluetooth readings, etc.

Transformations illustrated

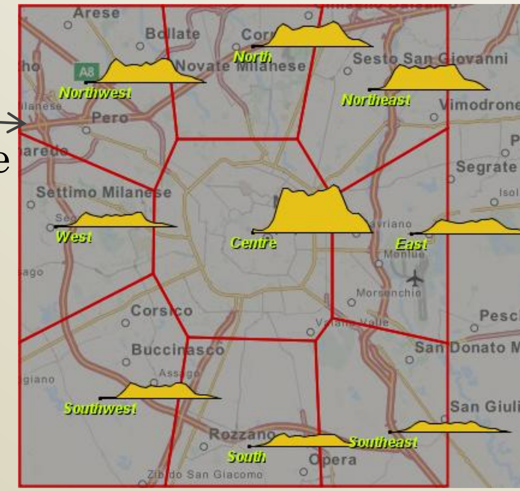
Trajectories



divide into trips

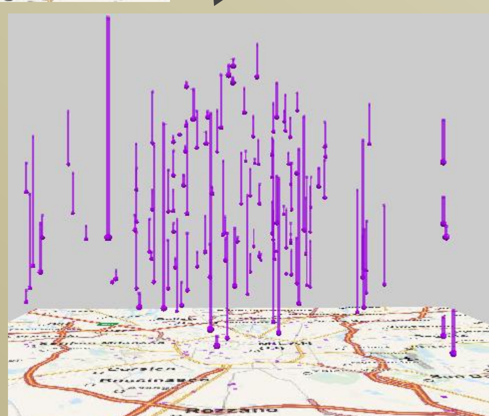


aggregate



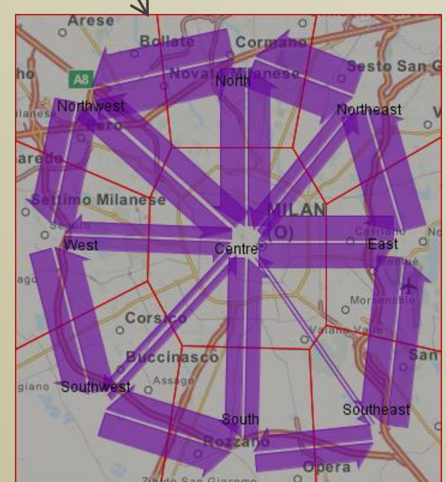
presence of movers in areas by time intervals

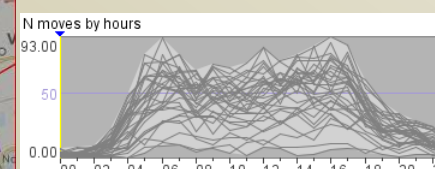
extract events



e.g., stops

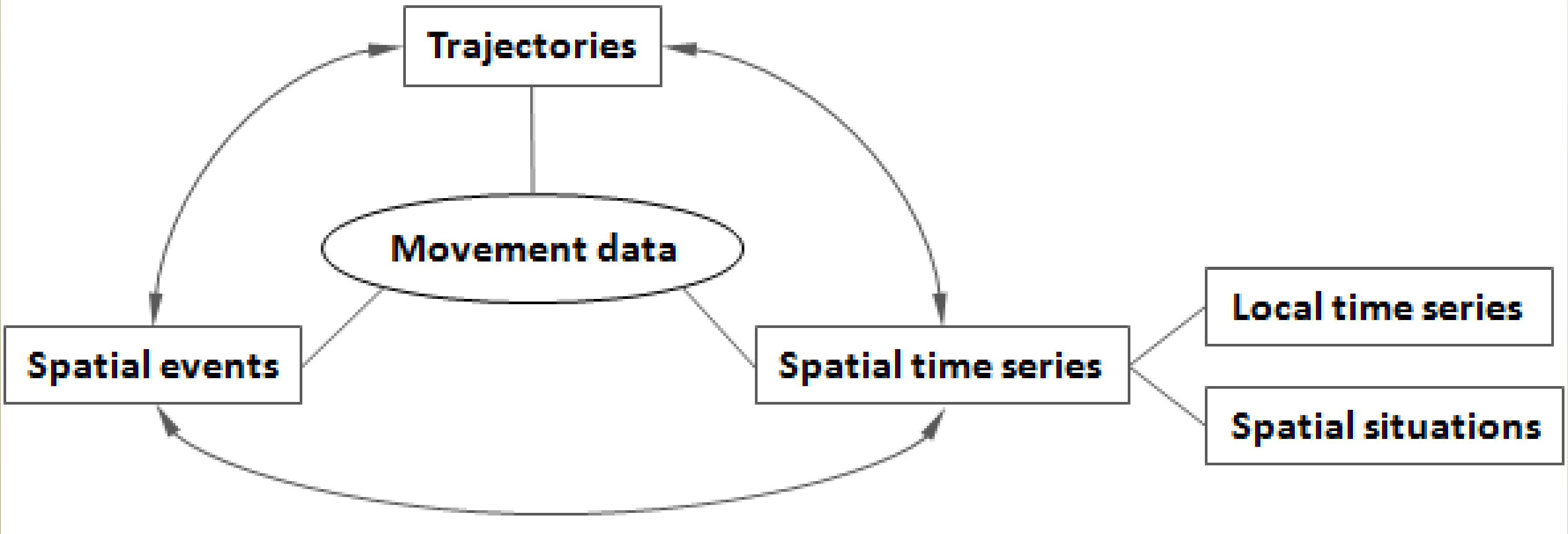
aggregate





flows (aggregate moves) of movers between areas by time intervals

Transformations enable multi-perspective analysis of movement data



Questions?

Extraction of movement events from trajectories

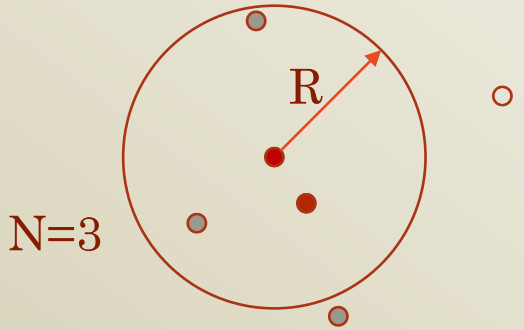
Density-based clustering of trajectories

Distance functions for trajectoris

Density-based clustering *(a reminder)*

Goal: find dense groups of close or similar objects

- For a given object ***o***, the objects whose distances from ***o*** are within a chosen distance threshold (radius) **R** are called <u>neighbours</u> of the object ***o***.
- An object is treated as a <u>core</u> object of a cluster if it has at least **N** neighbours.



- To make a cluster:
  1) some core object with all its neighbours is taken;
  2) for each core object already included in the cluster, all its neighbours are also added to the cluster (if not added yet).
- Some objects may remain out of any cluster (when they have not enough neighbours and do not belong to the neighbourhood of any core object). These objects are treated as "<u>noise</u>".

Density-based clustering

Distance

- For DBC, the user needs to specify the neighbourhood radius (distance threshold) **R**.

  ⇒ The use of DBC requires an understandable definition of **distance** between objects, e.g., spatial distance or spatio-temporal distance.

Distance between trajectories ?

- Trajectories are complex objects
  - consisting of multiple spatio-temporal points, having origins and destinations, particular shapes, lengths, durations, and dynamically changing movement directions and speeds.
- It is hardly possible to define a distance measure that accounts for all these properties.
- Even if such a measure could be defined, it would be hard to understand.

  ⇒It would be quite difficult to choose a meaningful value of R for clustering.

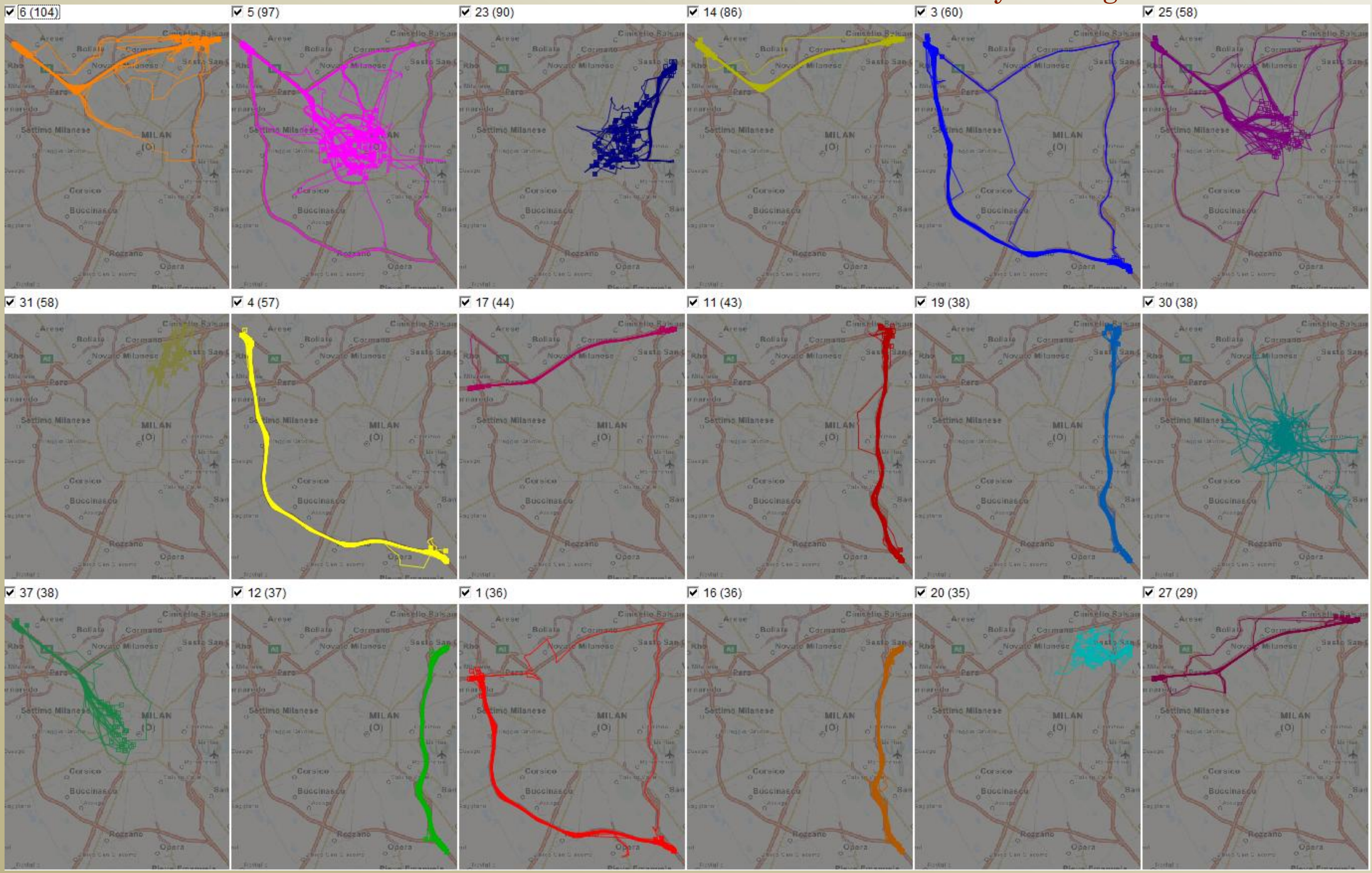
Diverse distance functions for trajectories

- It is more feasible to create a library of simple distance measures (distance functions) addressing different properties, e.g.
  - spatial distance between origins and/or between destinations,
  - average spatial distance between corresponding points along the routes,
  - average spatial distance between points reached at the same times, ...
- Such measures are easy to interpret and computationally efficient
- They support finding answers to different types of questions concerning trajectories.

Example 1

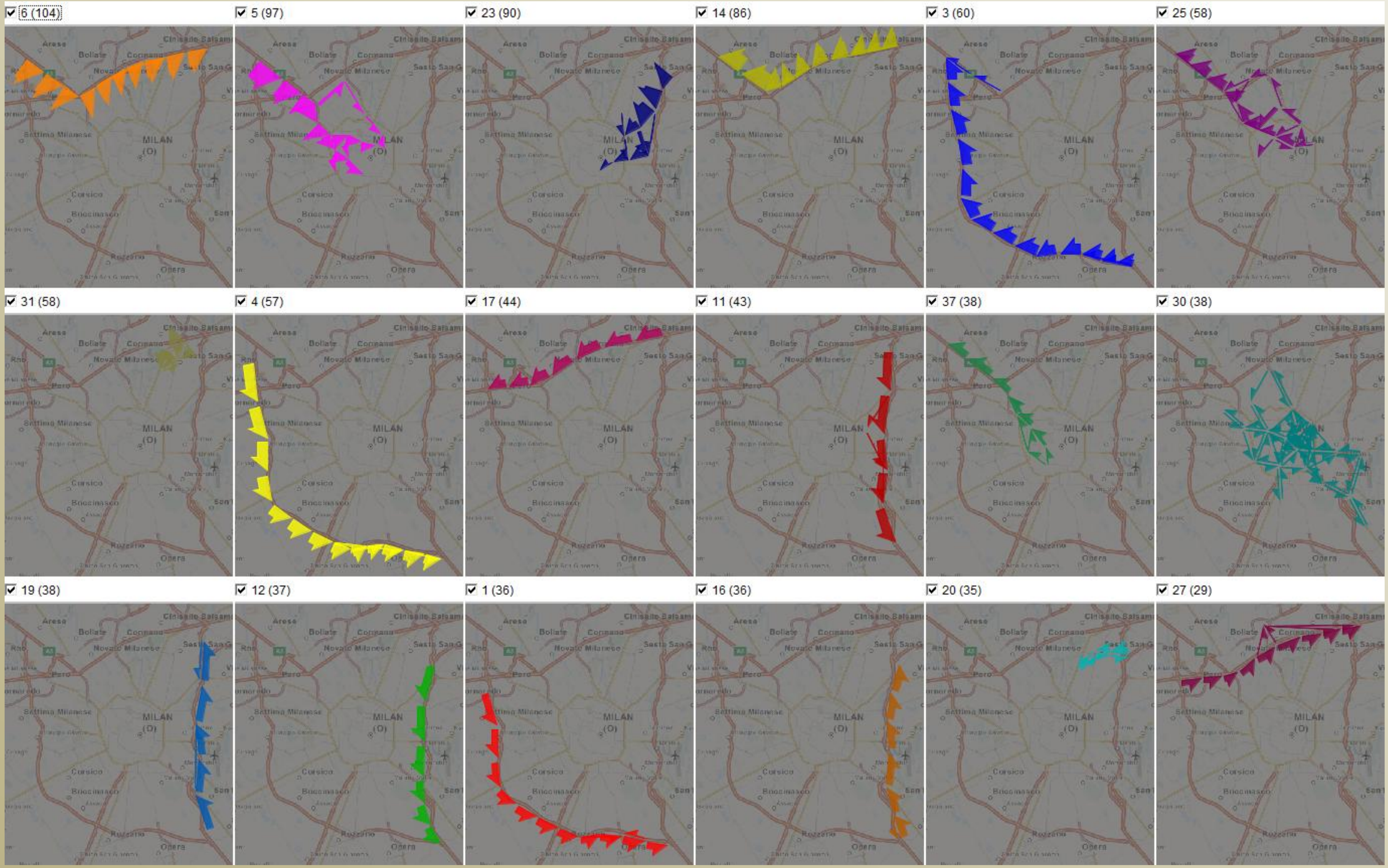
Distance function: the average spatial distance between the origins and between the destinations; R=750m, N=5

Only 18 largest clusters are shown.



The clusters are represented in a summarised form.

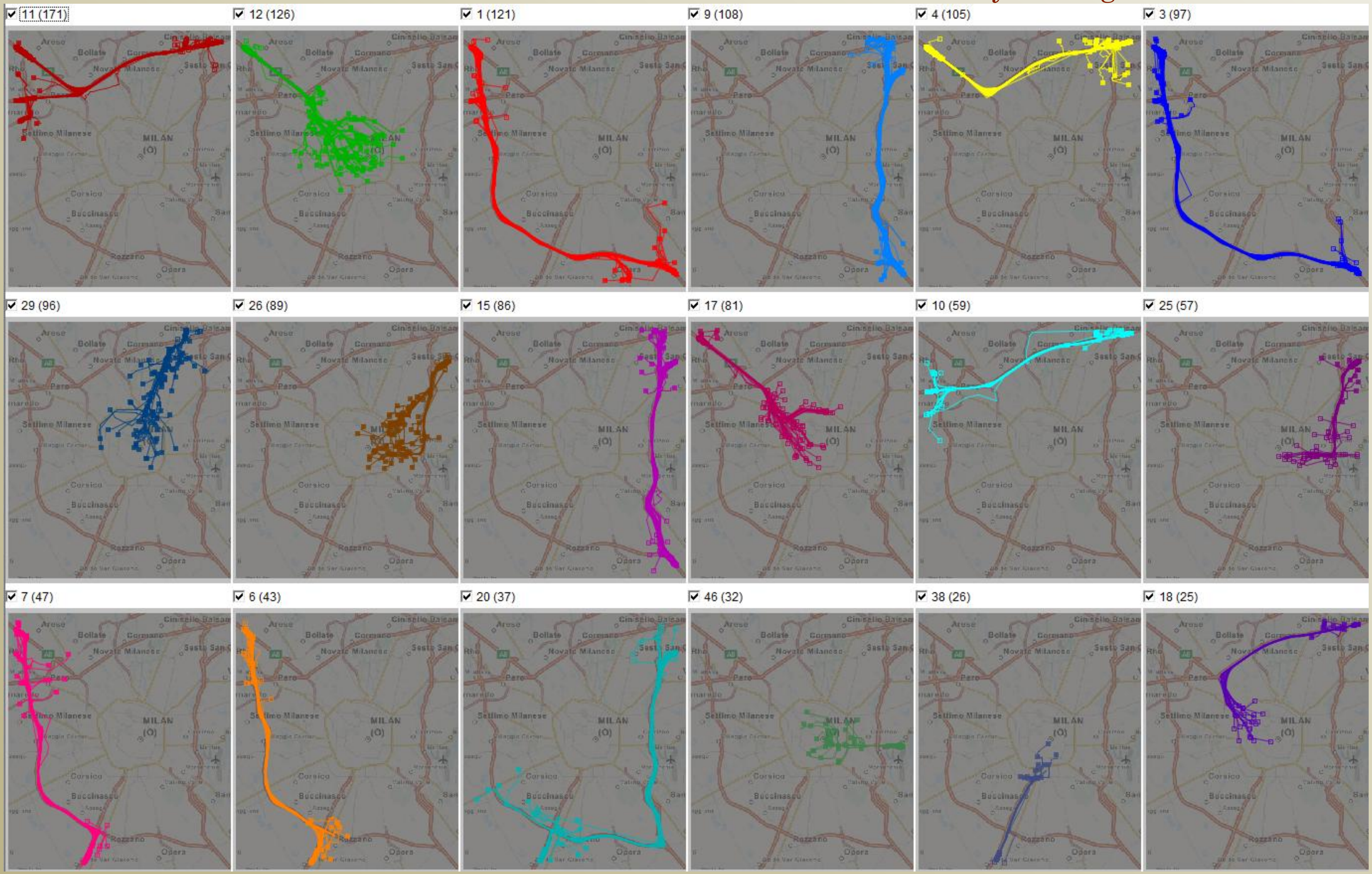
Minor flows are omitted for a clearer view.



Example 2

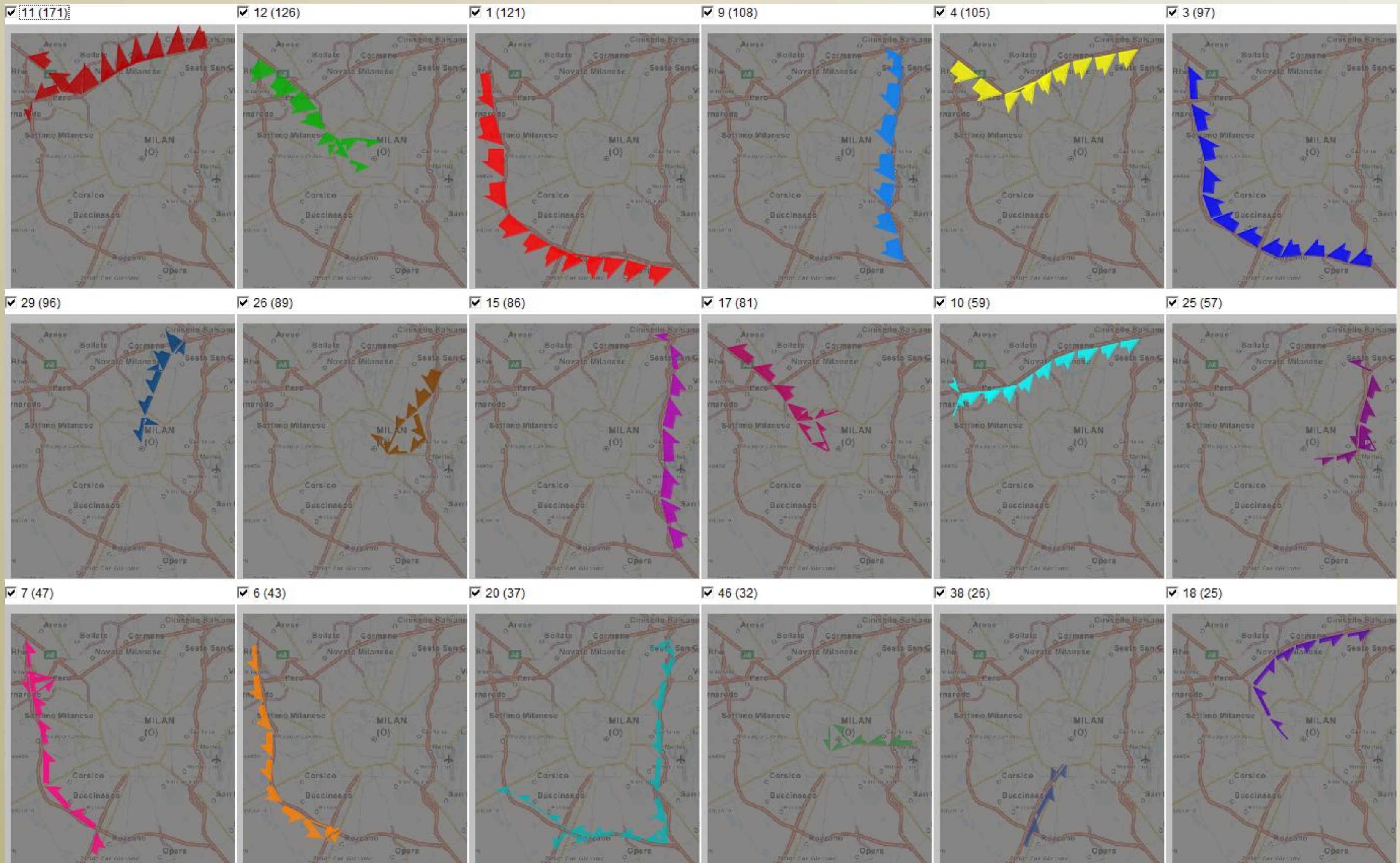
Distance function: "route similarity", i.e., the average spatial distance between the corresponding points along the route; R=750m, N=5

Only 18 largest clusters are shown.



The clusters are represented in a summarised form.

Minor flows are omitted for a clearer view.



Interactive progressive DBC

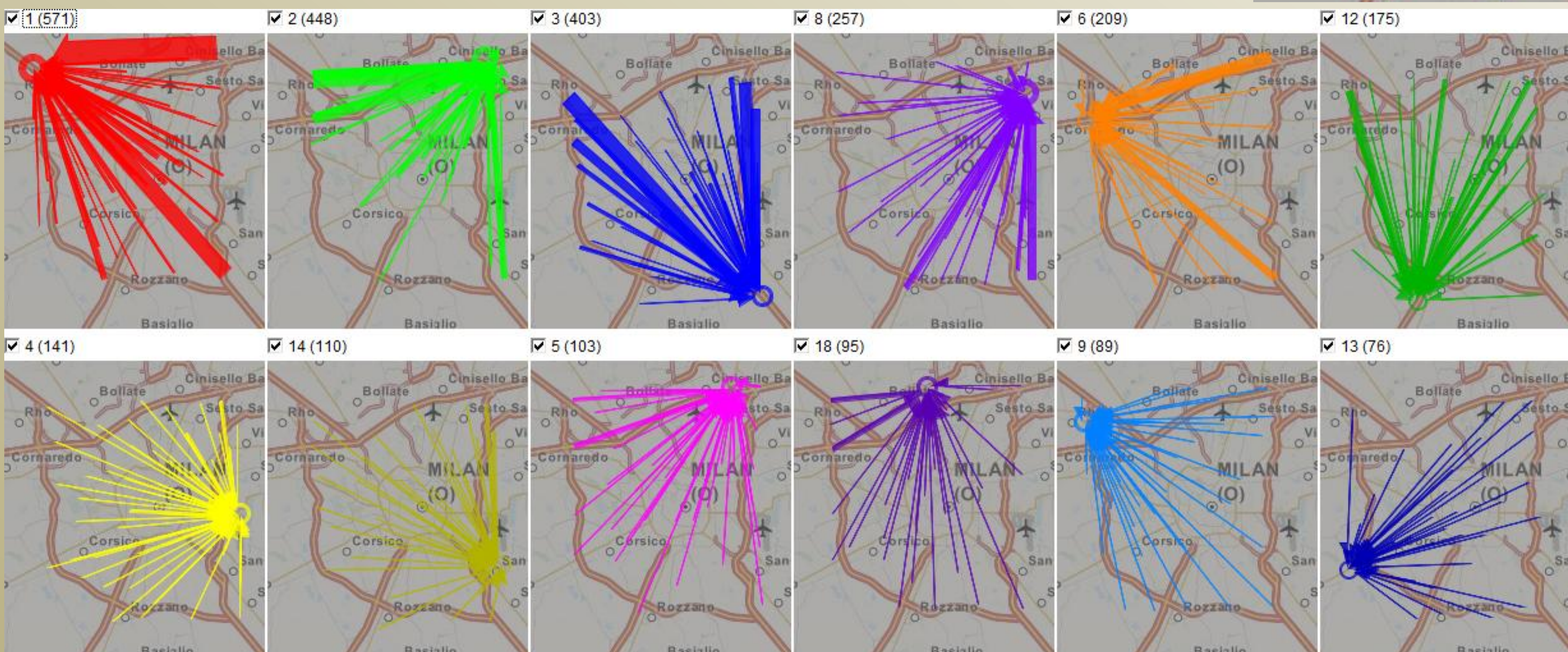
Applying different distance functions (1)

Data: trajectories of cars in Milan

Step 1: clustering according to the spatial proximity of the end points

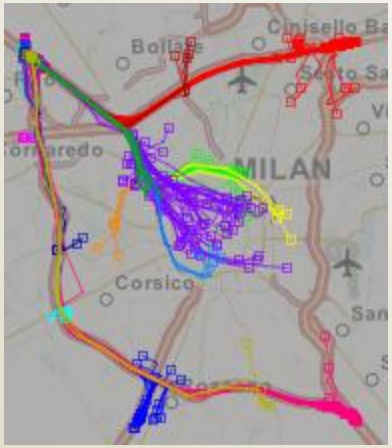
Distance function: "common ends"

<u>Question</u>: what are the most frequent destinations of car trips?



Interactive progressive DBC

Applying different distance functions (2)

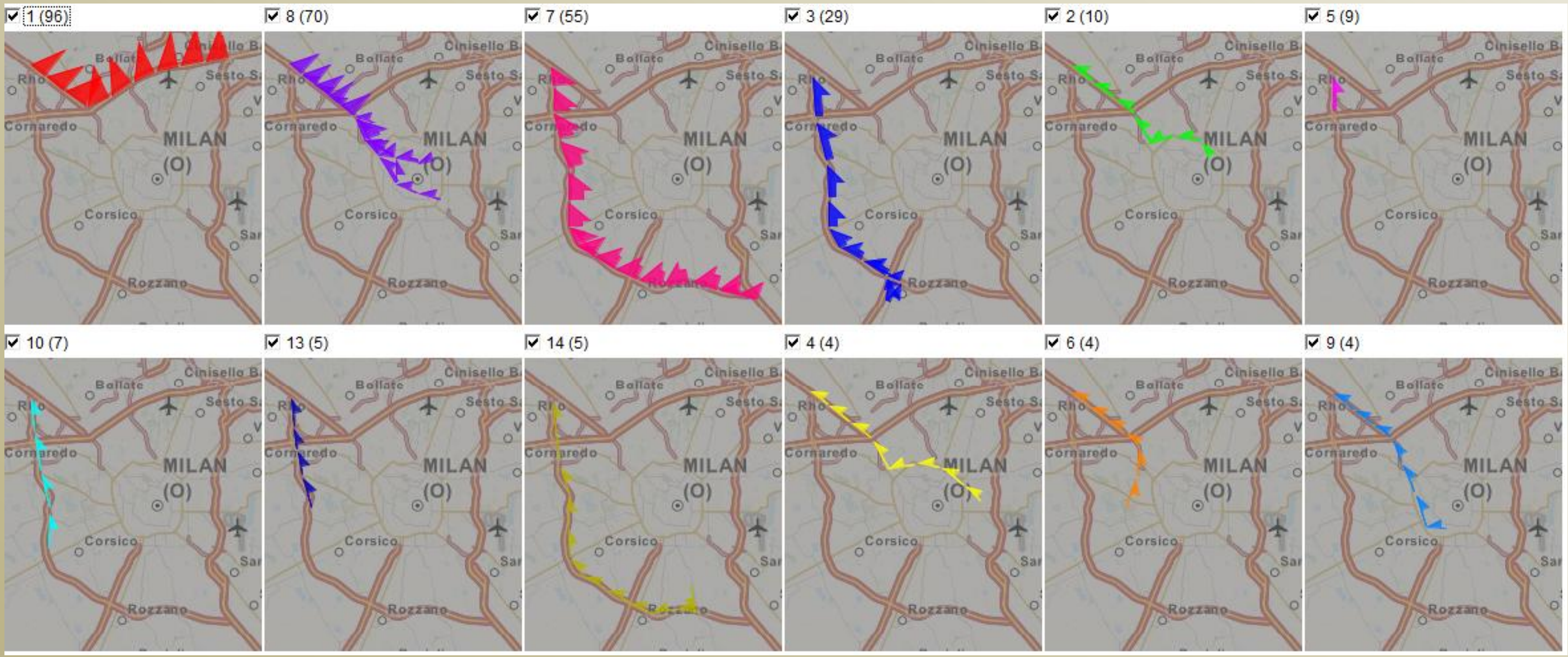


Data: one (or more) selected cluster(s) from the previous step

Step 2: clustering according to the similarity of the routes (shapes)

Distance function: “route similarity”

<u>Question</u>: what routes are usually taken to get to the selected destination?



Interactive progressive clustering

Purposes

- Controlled refinement of previously obtained clusters for
  - reducing internal variation
  - more detailed investigation of data subsets of interest
- Study of a set of complex objects with heterogeneous properties
  - application of diverse distance measures addressing different properties
    - a single distance measure would be hard to implement and results would be hard to interpret
  - incremental construction of multifaceted knowledge by progressively considering different properties

Where to read more

Salvatore Rinzivillo, Dino Pedreschi, Mirco Nanni, Fosca Giannotti, Natalia Andrienko, Gennady Andrienko

Visually–driven analysis of movement data by progressive clustering

Information Visualization, 2008, v.7 (3/4), pp. 225-239

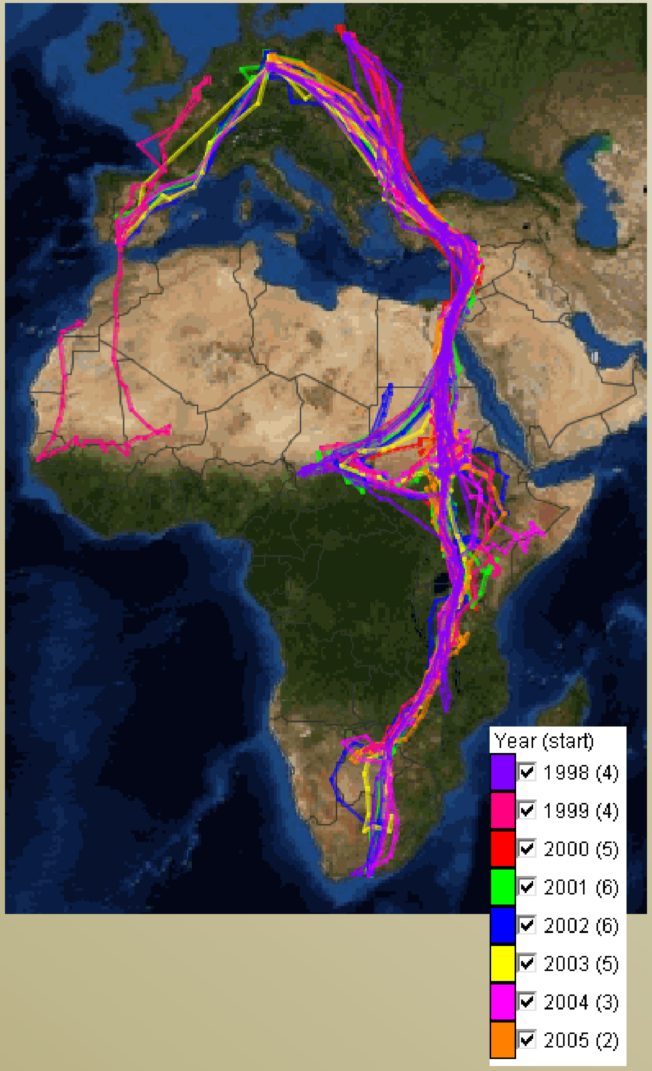
http://dx.doi.org/10.1057/palgrave.ivs.9500183

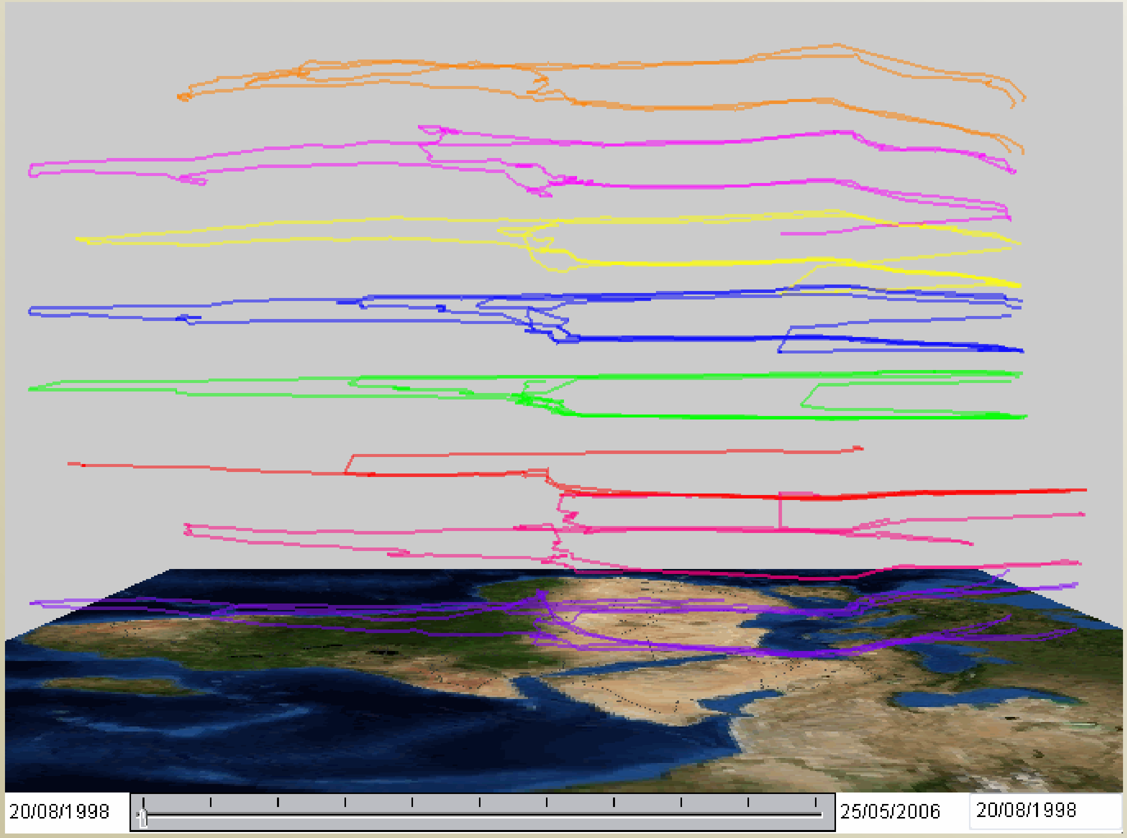
Questions?

Density-based clustering of trajectories,
progressive DBC

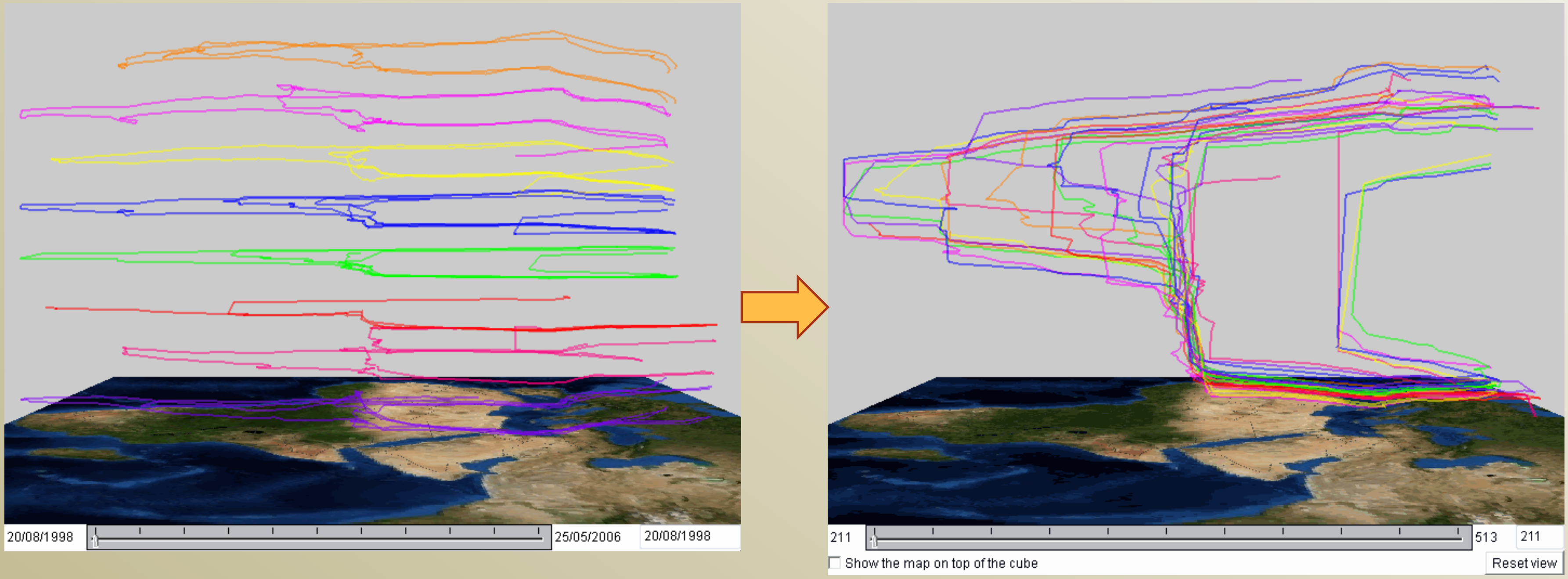
Transformation of time references in trajectories

Example 1: seasonal migration of white storks





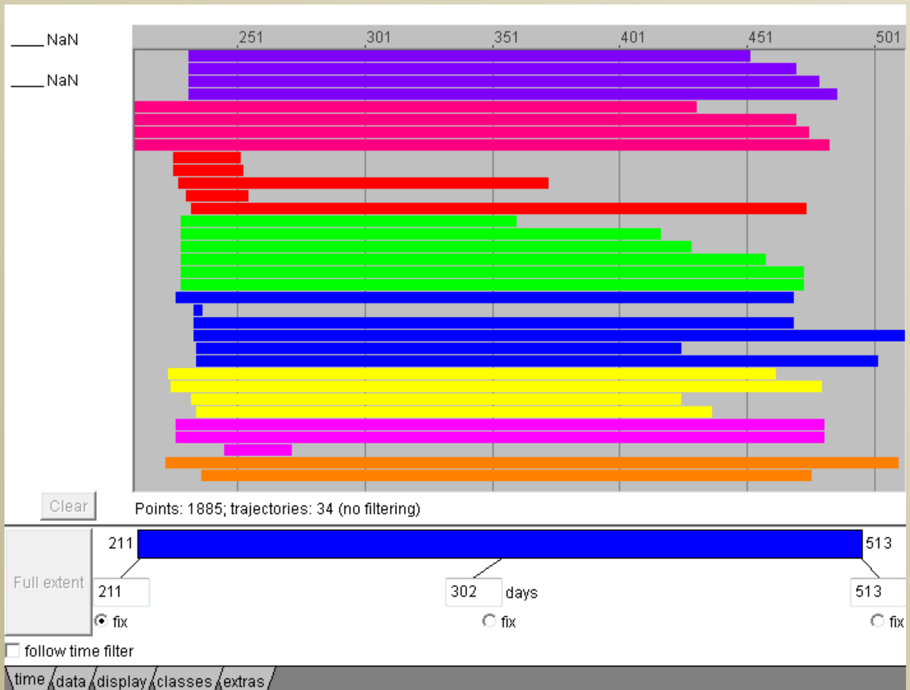
Time transformation to the seasonal cycle

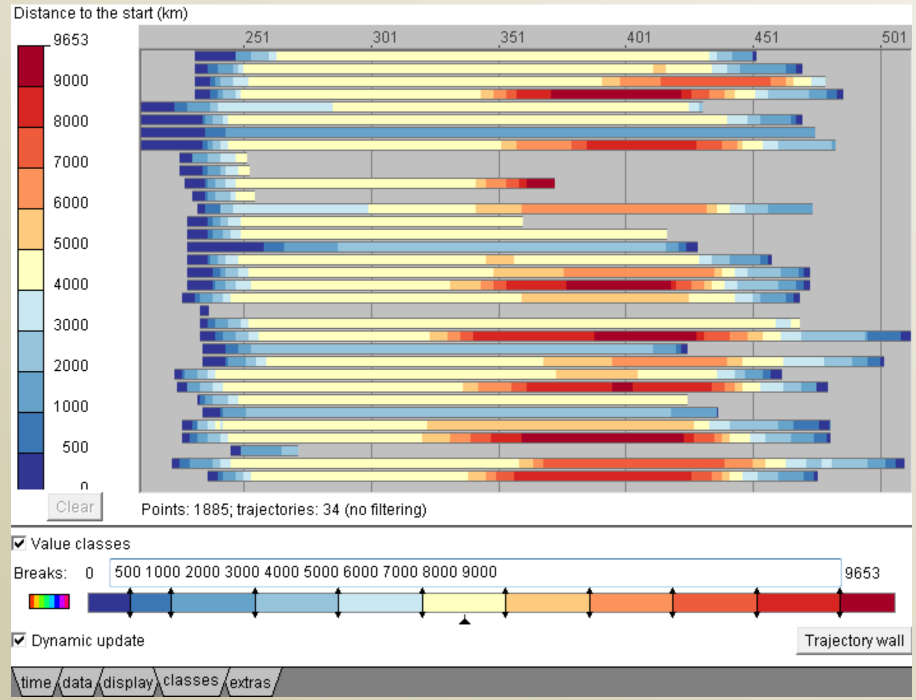


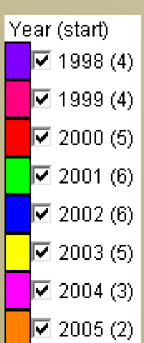
The absolute time references in the position records (calendar dates) are replaced by their relative positions within the yearly time cycle, i.e., each date is replaced by its ordinal N since the beginning of the year.

The transformation allows us to align the trajectories in a space-time cube, which helps us to compare the routes.

Time transformation to the seasonal cycle



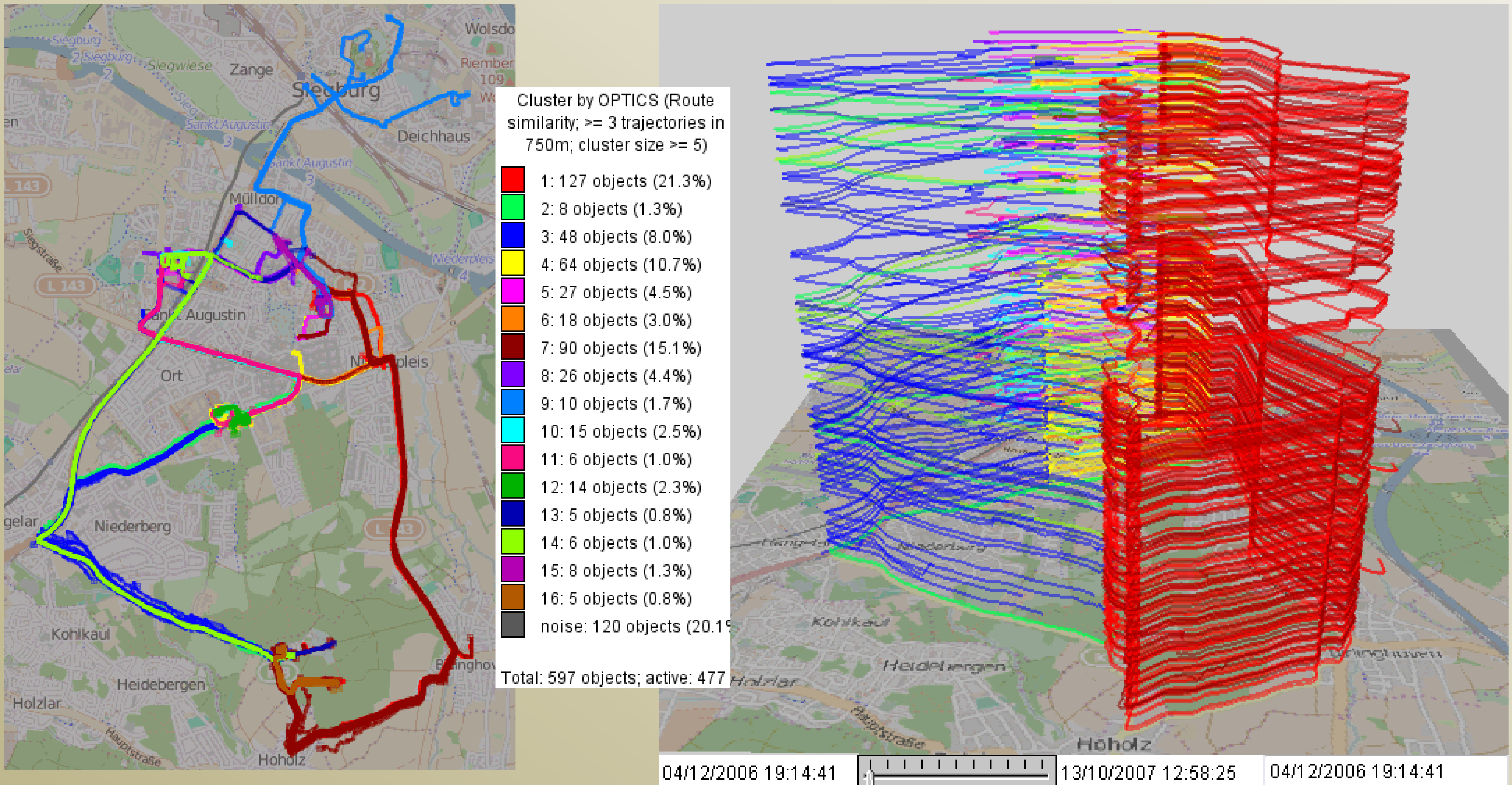




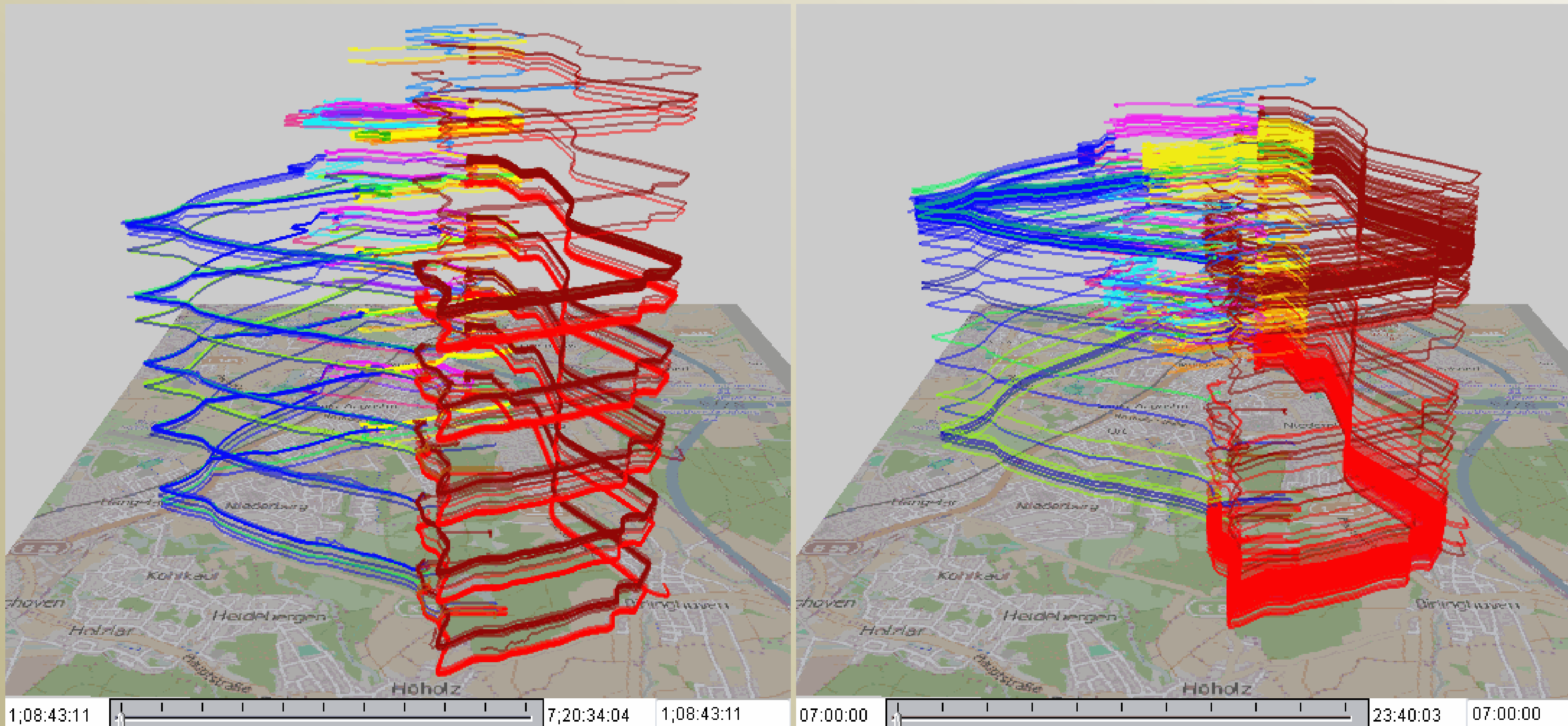
In a timeline view of trajectories, the bars are also aligned, allowing us to compare the times of migration beginnings and ends in different years and the variations of the birds' distances from their home locations over the migration seasons.

Example 2: trips of a personal car

(resulting from track division by 15 minute stops)

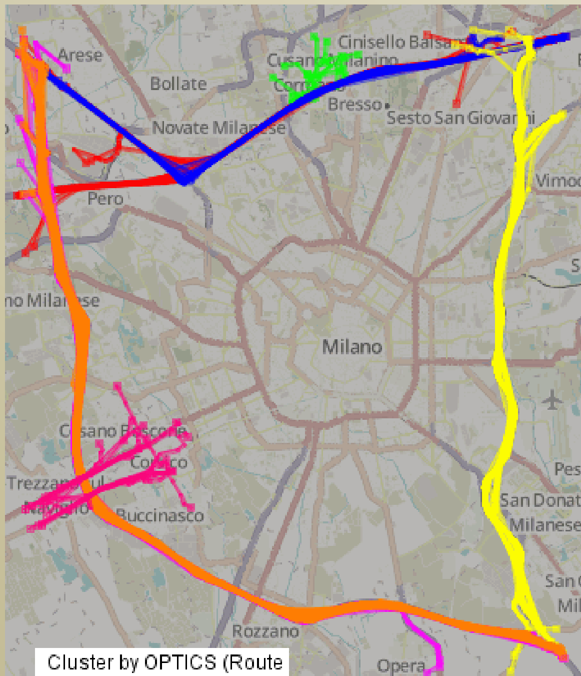


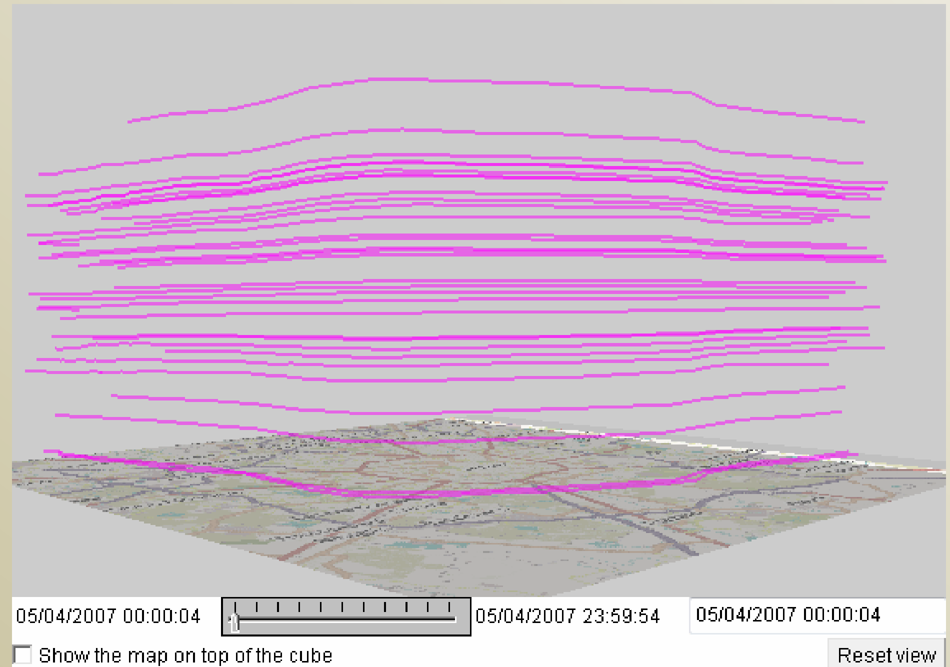
Time transformations to the weekly and daily cycles



The transformations allow us to identify routine weekly and daily patterns of the personal mobility behaviour.

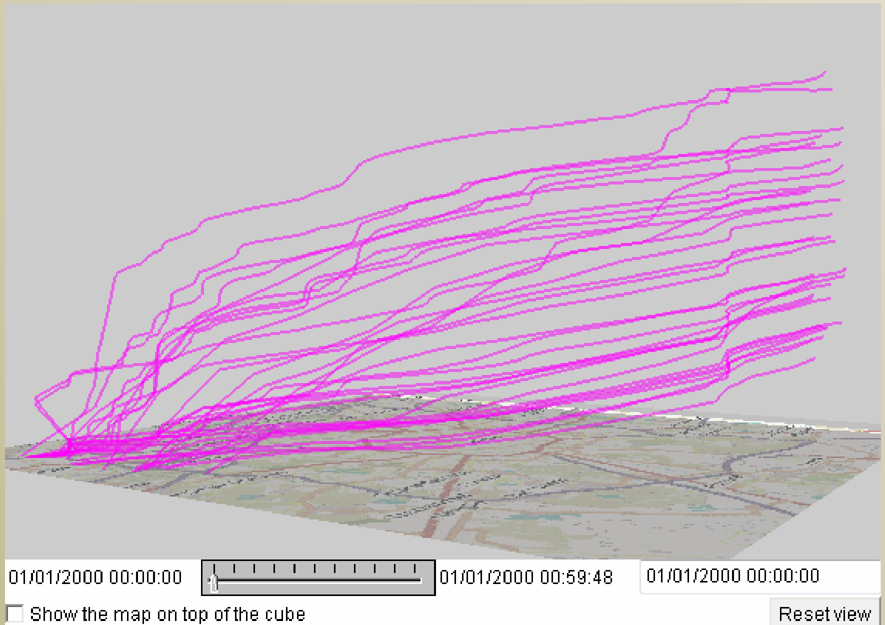
Example 3: a sample of car trips from Milan (division by 15 minute stops)

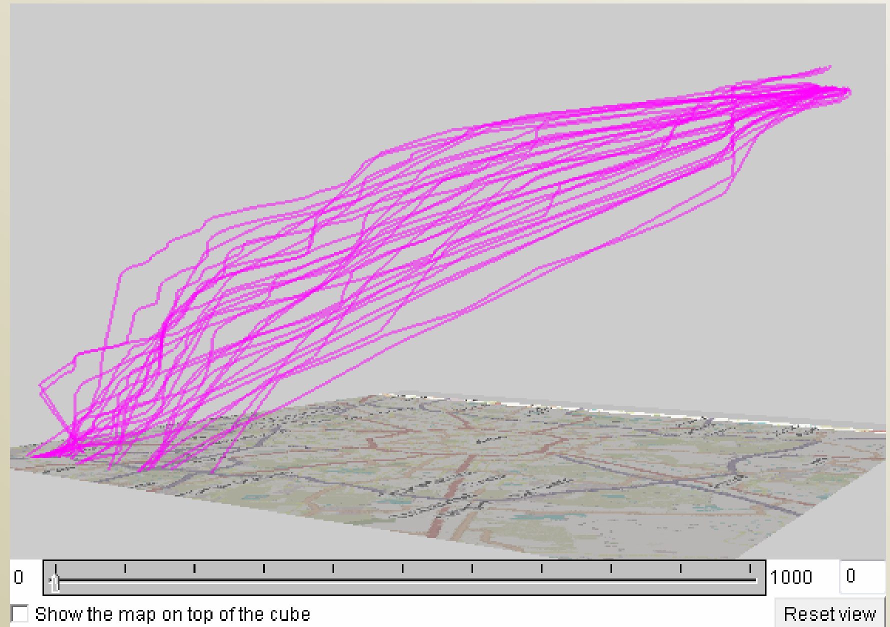




One of density-based clusters of trajectories by route similarity is chosen for a detailed inspection.

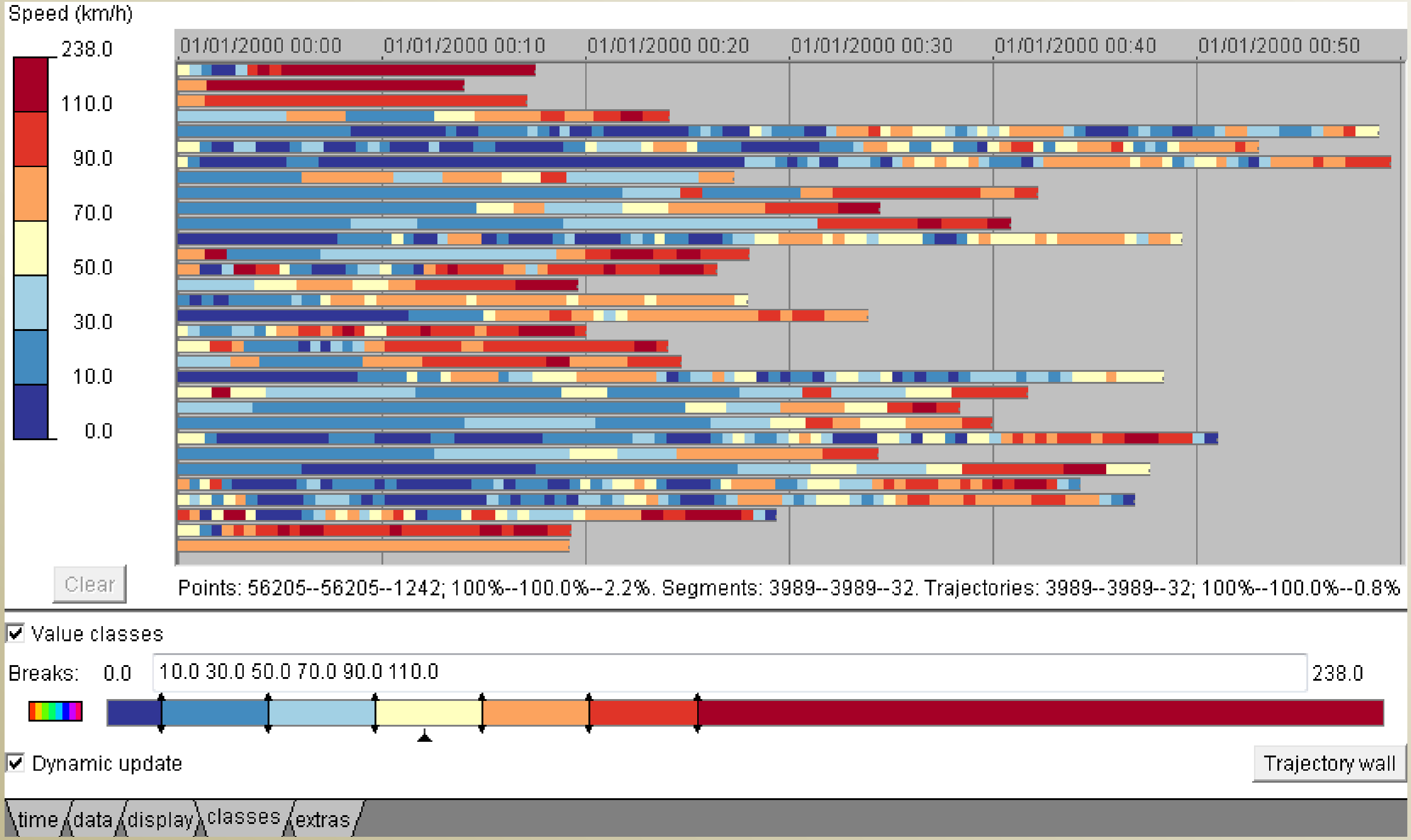
Time transformation to the start or end times of the trips





The absolute time references are replaced by relative (i.e., the time differences) with respect to the times of the trip starts or ends.

The transformation allows us to compare us the internal dynamics of trajectories following the same or similar routes. We can distinguish trajectories and parts of trajectories with fast and slow movement.



The trajectories can also be aligned in a timeline view. Here we also see that many trajectories had low speeds at the beginnings. We can compare the trajectories in terms of the duration of the obstructed movement.

Time transformations in trajectories

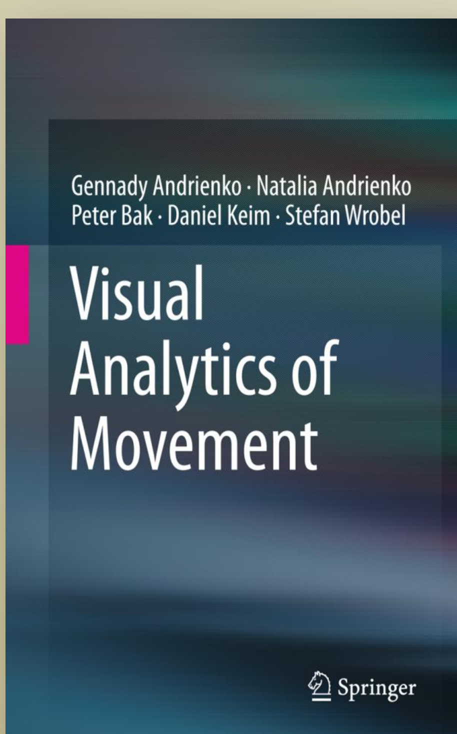
A summary

- Transformation to relative positions within a temporal cycle (seasonal, weekly, daily)
  - Purpose: identify and compare routine movements
- Transformation to trip starts or ends (or both)
  - Purpose: compare the internal dynamics between trajectories following same or similar routes

Questions?

Time transformations in trajectories

Where to read more:



Springer, September 2013, ISBN 978-3-642-37582-8

Ch.1. **Introduction**

Ch.2. **Conceptual framework**

Ch.3. **Transformations of movement data**

Ch.4. **Visual analytics infrastructure**

Ch.5. **Visual analytics focusing on movers**

Ch.6. **Visual analytics focusing on spatial events**

Ch.7. **Visual analytics focusing on space**

Ch.8. **Visual analytics focusing on time**

Ch.9. **Discussion and outlook**