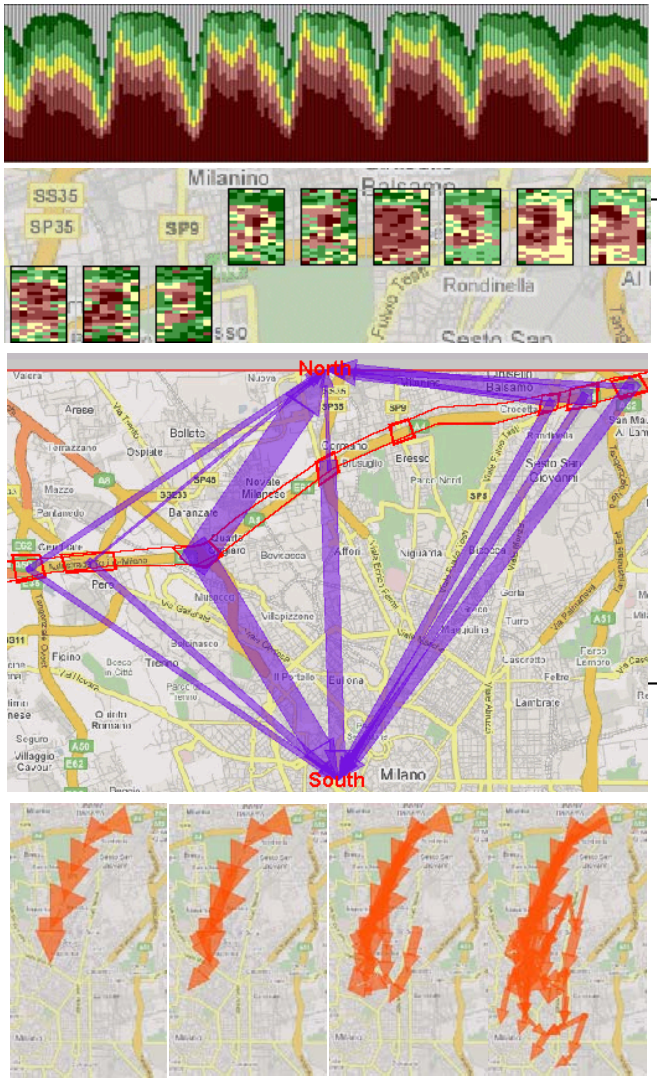
Spatio-temporal Aggregation for Visual Analysis of Movements



Gennady Andrienko & Natalia Andrienko
http://geoanalytics.net

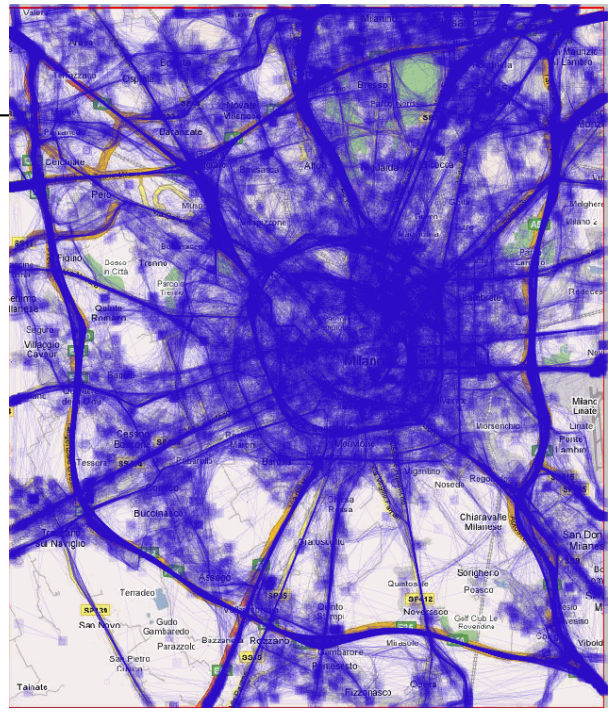
IAIS

Fraunhofer Institut Intelligente Analyse- und Informationssysteme

GeoPKDD
Geographic Privacy-aware Knowledge Discovery and Delivery

Movement data:
typical structure and typical sizes

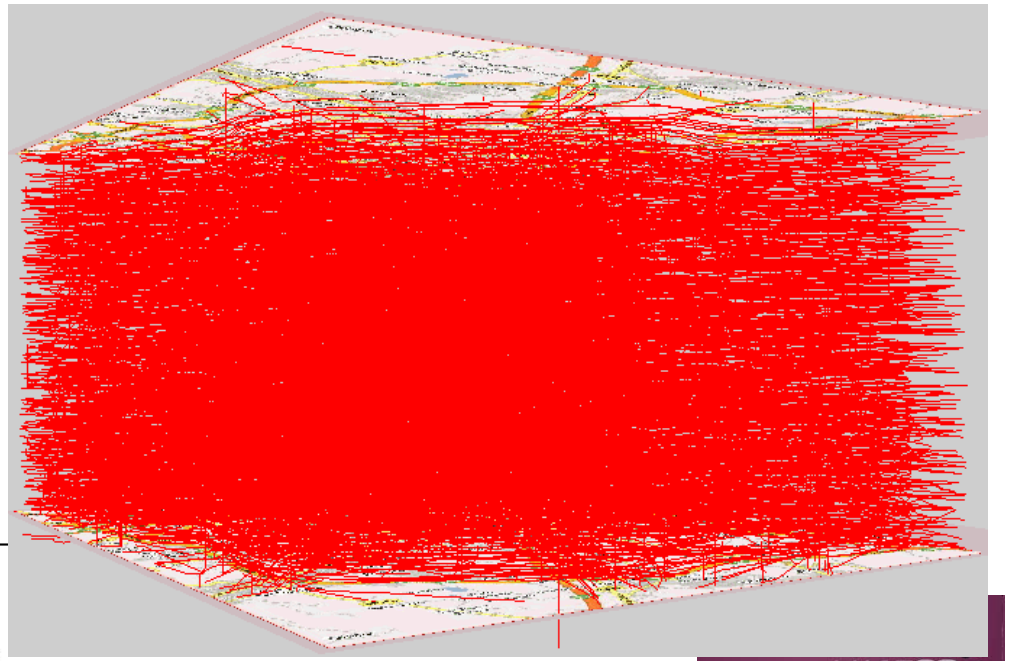
ID	X	Y	time	SPEED	COURSE
600	9.170161	45.43414	02/04/2007 13:47	0	285.3052
600	9.158362	45.43737	02/04/2007 13:52	16	290.5214
600	9.152749	45.43947	02/04/2007 13:57	0	0
600	9.152749	45.43947	02/04/2007 14:00	0	356.4407
600	9.152736	45.43967	02/04/2007 14:01	0	155.9657
600	9.152835	45.43945	02/04/2007 14:07	0	313.
600	9.14287	45.4489	02/04/2007 14:16	24	244.
600	9.129485	45.44251	02/04/2007 14:21	0	
1165	9.266666	45.431	03/04/2007 03:47	88	271.
1165	9.251568	45.43143	03/04/2007 03:49	80	283.
1165	9.232685	45.4361	03/04/2007 03:52	48	296.
1165	9.211592	45.4466	03/04/2007 03:55	42	352.1556
1165	9.210249	45.45635	03/04/2007 04:00	0	0
1330	9.263878	45.55659	02/04/2007 08:32	64	198.3963
1330	9.258997	45.54192	02/04/2007 08:35	68	178.0536
1330	9.259572	45.525	02/04/2007 08:37	74	177.2814
1330	9.260412	45.50731	02/04/2007 08:38	98	206.5272
1330	9.251644	45.48974	02/04/2007 08:40	86	177.9747
1330	9.252272	45.47198	02/04/2007 08:41	62	173.5334
1330	9.254257	45.45447	02/04/2007 08:43	74	175.2762
1330	9.255733	45.43661	02/04/2007 08:44	76	212.353
...	...	...	...	...	...	...



E.g. Milan car movement data:

***2,075,216** position records*

of 17,241 cars during 1 week



Movement data are collected in very large amounts

⇒ traditional visualizations fail

⇒ aggregation and summarization are necessary

Gennady & Natalia Andrienko
http://geoanalytics.net/and

Fraunhofer Institut Intelligente Analyse- und Informationssysteme

Related works

- Temporal, geographical, and categorical aggregation of point events
  - Fredrikson, A., North, C., Plaisant, C., & Shneiderman, B.: Temporal, geographical and categorical aggregations viewed through coordinated displays: a case study with highway incident data. In Proc. Workshop on New Paradigms in information Visualization and Manipulation (Kansas City, Nov. 1999). ACM, NY, 1999, 26-34.
- Aggregation of position records from movement data analogously to point events
  - Dykes, J. A. & Mountain, D. M.: Seeking structure in records of spatio-temporal behavior: visualization issues, efforts and applications, Computational Statistics and Data Analysis, 43 (Data Visualization II Special Edition), 2003, 581-603.
  - … and many others
- Aggregation by origins and destinations of the moves
  - **Flow maps (discrete and continuous):**
    Tobler, W.: Experiments in migration mapping by computer, The American Cartographer, 14 (2), 1987, 155-163
  - **Origin-destination matrices:**
    Guo, D.: "Visual Analytics of Spatial Interaction Patterns for Pandemic Decision Support". International Journal of Geographical Information Science, 21(8), 2007, pp. 859-877
- Geometric summarization of trajectories
  - Buliung, R.N. & Kanaroglou, P.S.: An Exploratory Data Analysis (ESDA) toolkit for the analysis of activity/travel data. Proceedings of ICCSA 2004, LNCS 3044, 1016-1025
  - Schreck, T., Tekusova, T., Fellner, D., & Kohlhammer, J.. Trajectory Based Visual Analysis of Large Financial Time Series Data, SIGKDD Explorations, 9(2), 2007, pp.30-37

Goals of this work

- Collect and systemize the possible approaches to the aggregation and summarization of movement data
- Develop a general framework for the use of aggregation and summarization techniques in visual analysis of movement data

  In particular:

  - Define when the point-oriented techniques are valid and useful and when they are invalid or insufficient

A formal model of movement data

- Movement of multiple entities can be modeled as a function

$$\mu: E \times T \rightarrow S \quad \text{or} \quad \mu(e,t) = s$$

  E: the set of entities *(discrete; unordered)*

  T (time): the set of all time moments *(continuous, linearly ordered, cyclically ordered)*

  S (space): the set of all positions *(continuous; partially ordered)*

- May be viewed in 2 complementary ways:
  - Trajectory-oriented view: $\{ \mu_e: T \rightarrow S \mid e \in E\}$ or $\{ \mu_e(t) \mid e \in E\}$

    $\mu_e(t)$ – the movement of entity e over time (called **trajectory**)

    $\{ \mu_e(t) \mid e \in E\}$ - the set of trajectories of all entities
  - Traffic-oriented view: $\{ \mu_t: E \rightarrow S \mid t \in T\}$ or $\{ \mu_t(e) \mid t \in T\}$

    $\mu_t(e)$ – positions and movement characteristics of the entities at time t (called **traffic situation**, in an abstract sense)

    $\{ \mu_t(e) \mid t \in T\}$ – variation of the traffic situations over time

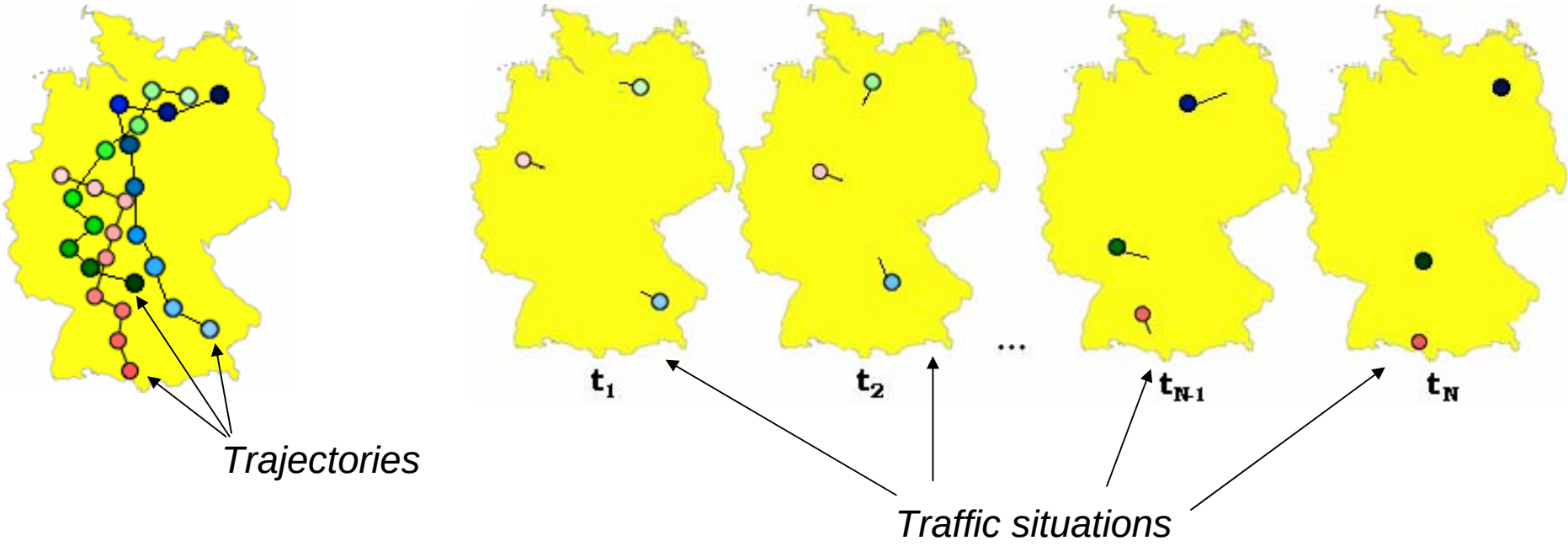
2 views: graphical illustration

Trajectory-oriented view

$\{ \mu_e(t) \mid e \in E \}$

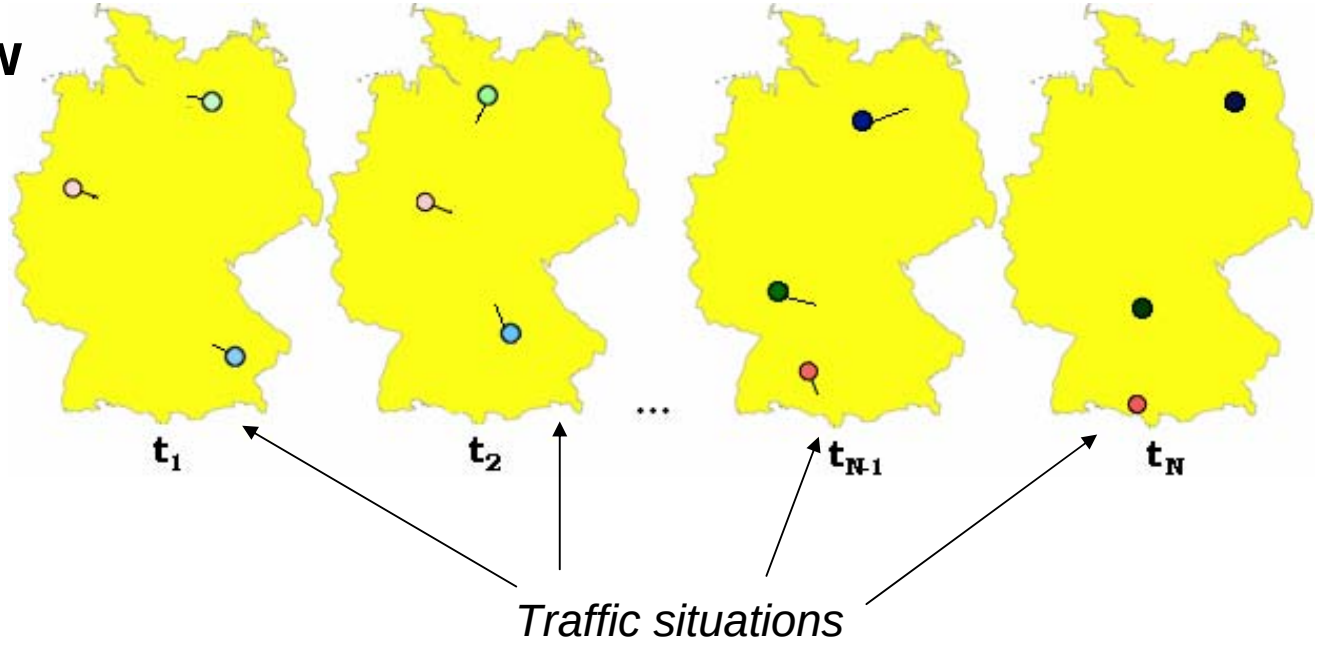
Traffic-oriented view

$\{ \mu_t(e) \mid t \in T \}$



Traffic-oriented view

$\{ \mu_t(e) \mid t \in T \}$



Needed: aggregation and summarization of **traffic situations**

Traffic situation ≈ set of ***point events*** ⇒ point-oriented approaches are applicable

- Spatial aggregation: positions in space → areas (e.g. regular grid)
- Temporal aggregation: moments in time → intervals on time line or in time cycles
- Attributive (categorical) aggregation:
  values of movement attributes → intervals (numeric) or subsets (qualitative)

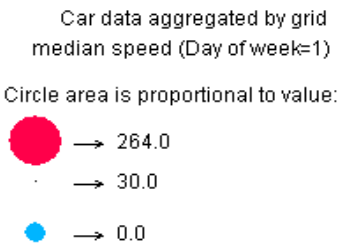
S × T – aggregation (example 1)

- S – aggregation: by cells of a regular grid
- T – aggregation: by hourly intervals (**linear time**)
- Derived characteristics of the aggregates: number of cars, statistics of the speeds (min, max, mean, median, etc.)

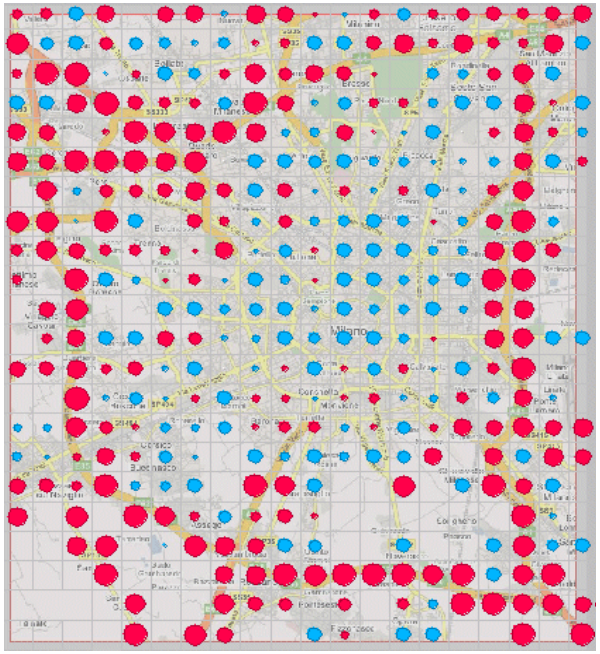
➢ Following slides: examples of visualisations using this aggregation

Animated map with graduated symbols

Attribute: median speed; blue: <30; red: >30



22h



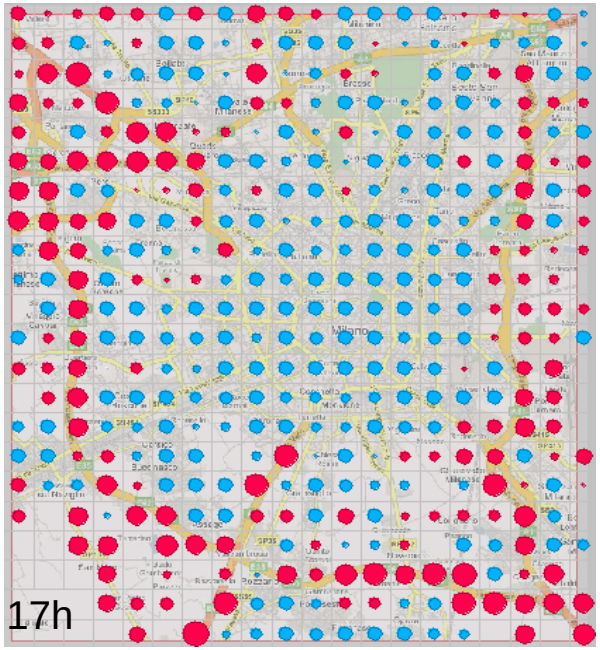
Animated map with graduated symbols

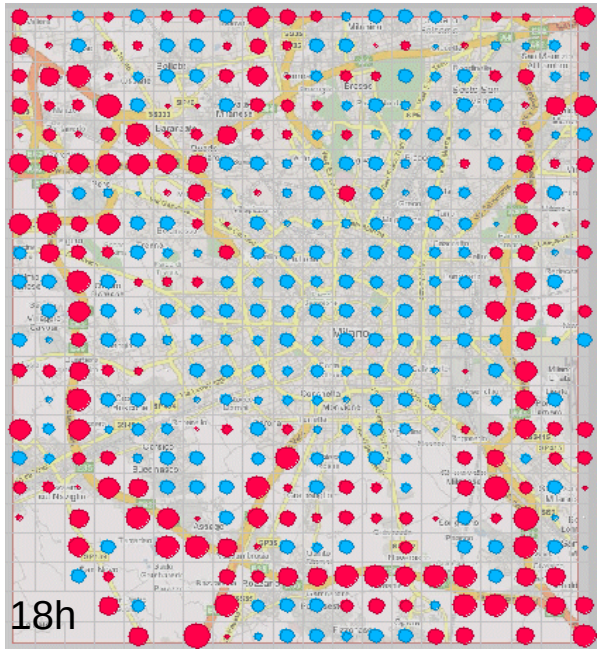
Car data aggregated by grid
median speed (Day of week=1)

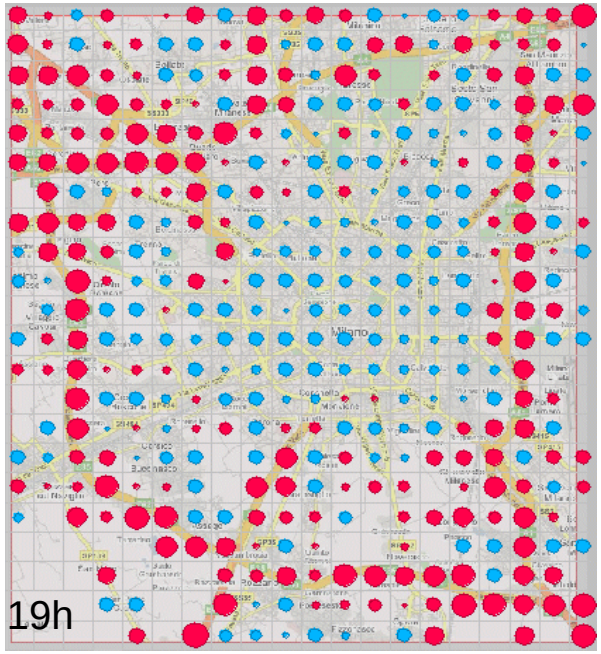
Circle area is proportional to value:

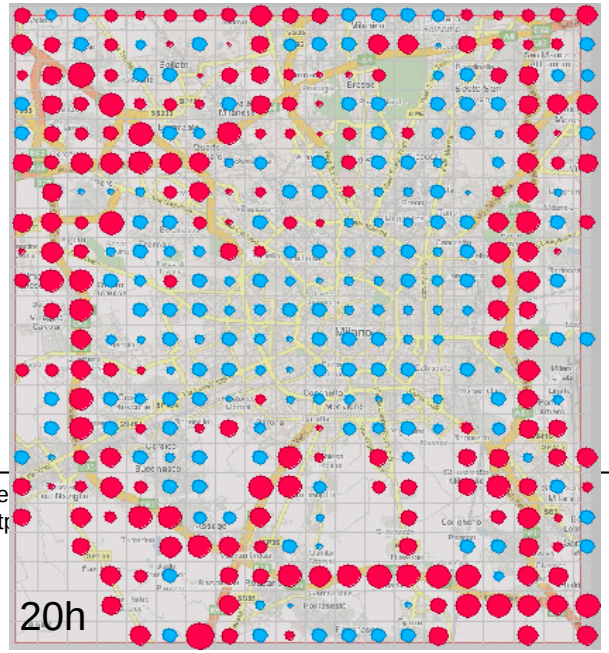
- 264.0
- 30.0
- 0.0

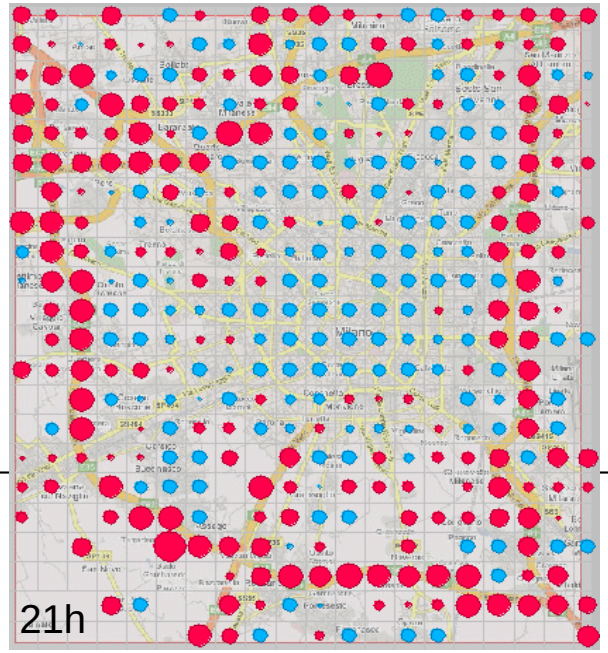
Attribute: median speed; day: Monday; hours: from 17 to 22 (states of the animated map); blue: <30; red: >30

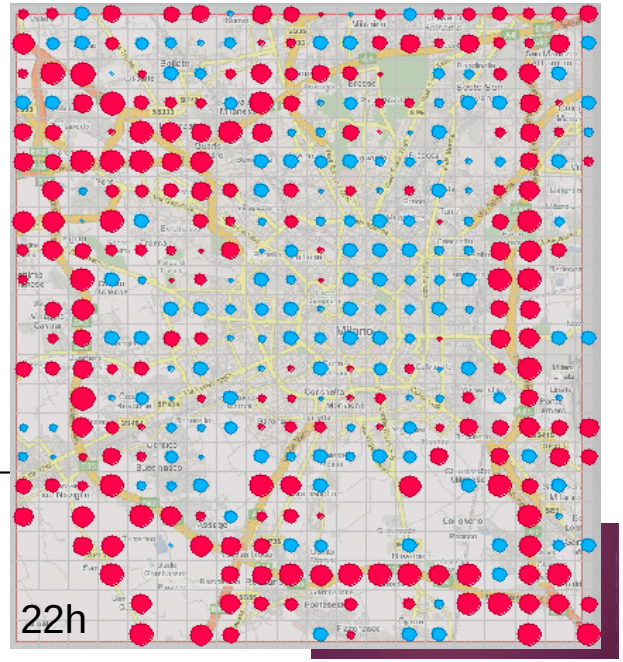










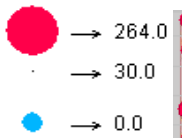


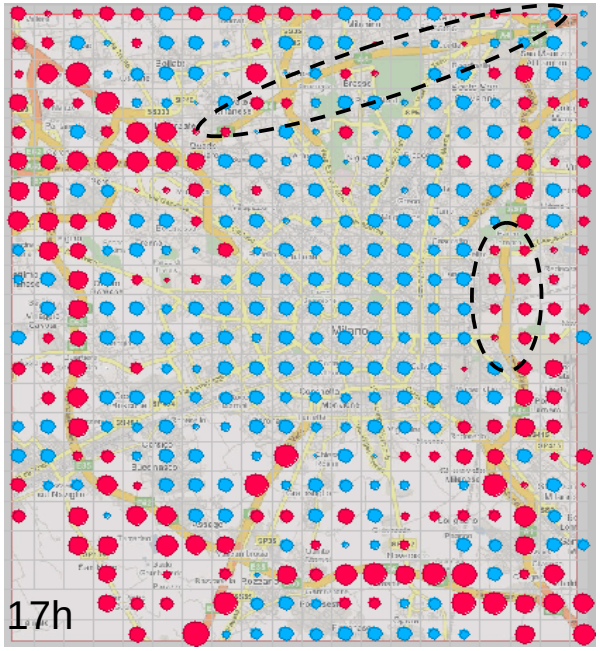
Ge
http

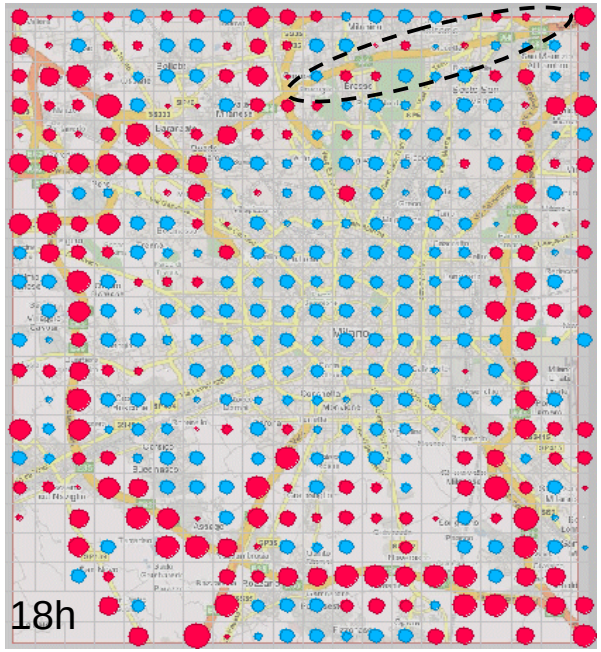
Too low speeds on the major belt roads

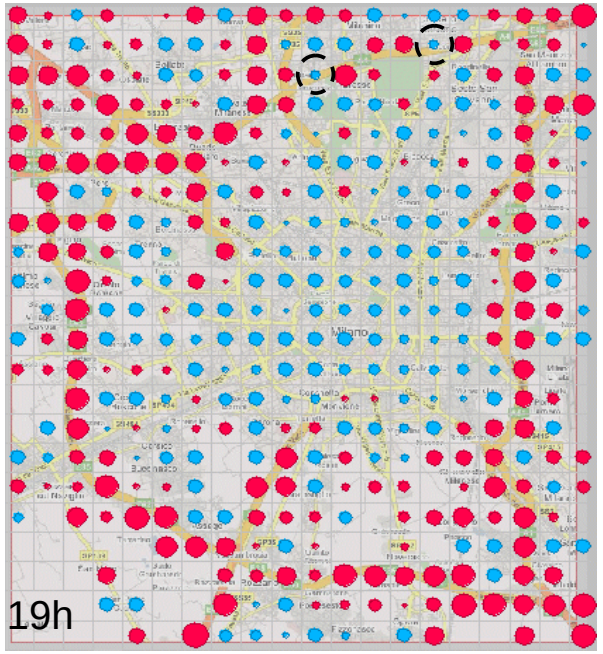
Car data aggregated by grid
median speed (Day of week=1)

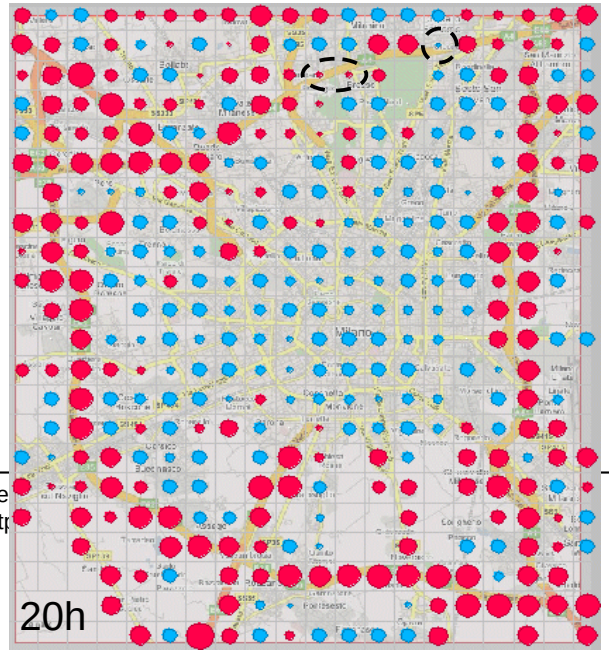
Circle area is proportional to value:

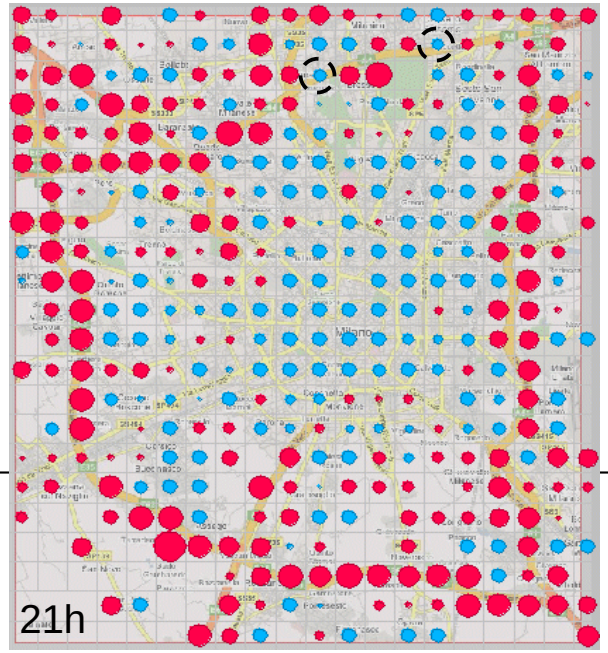


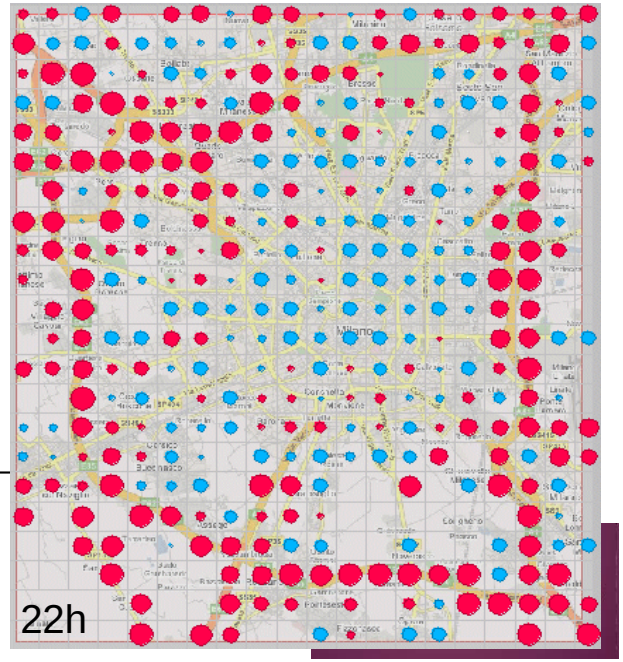










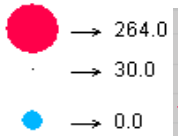


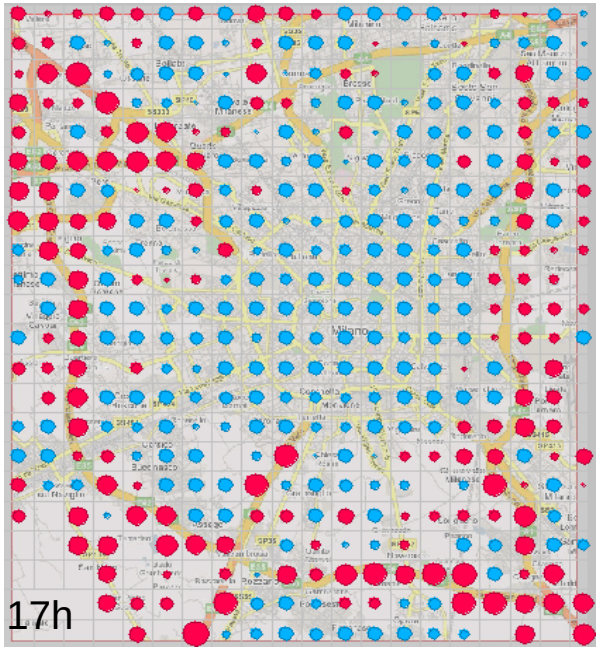
Ge
http

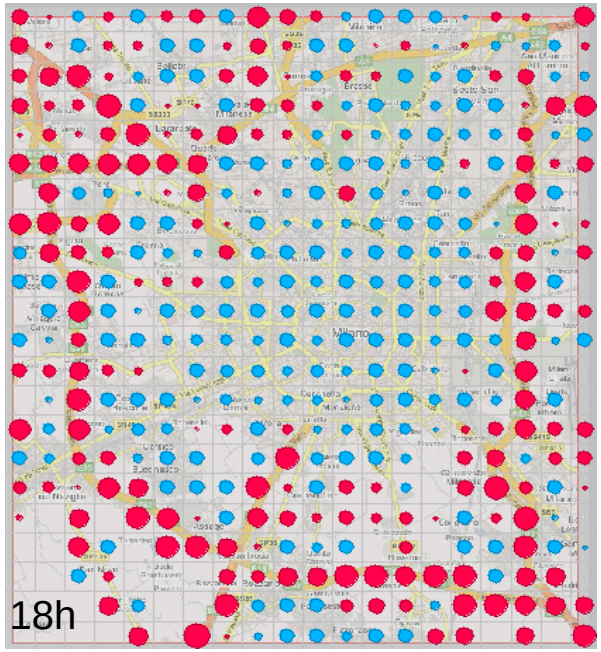
Increase of speeds in late evening

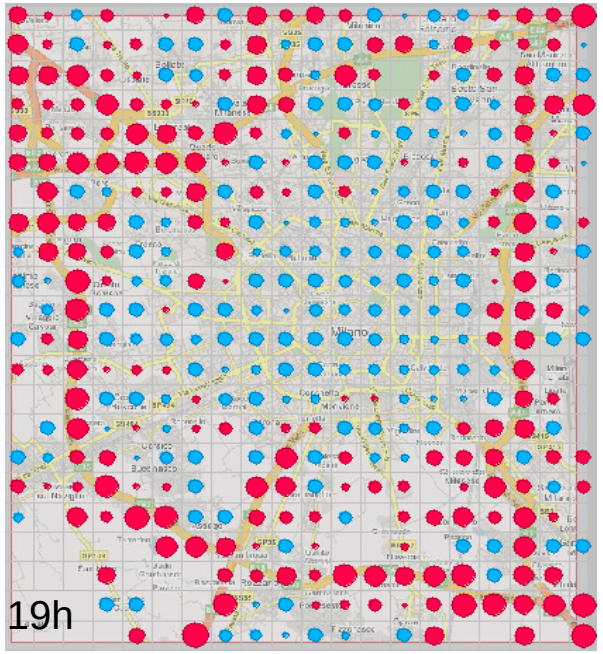
Car data aggregated by grid
median speed (Day of week=1)

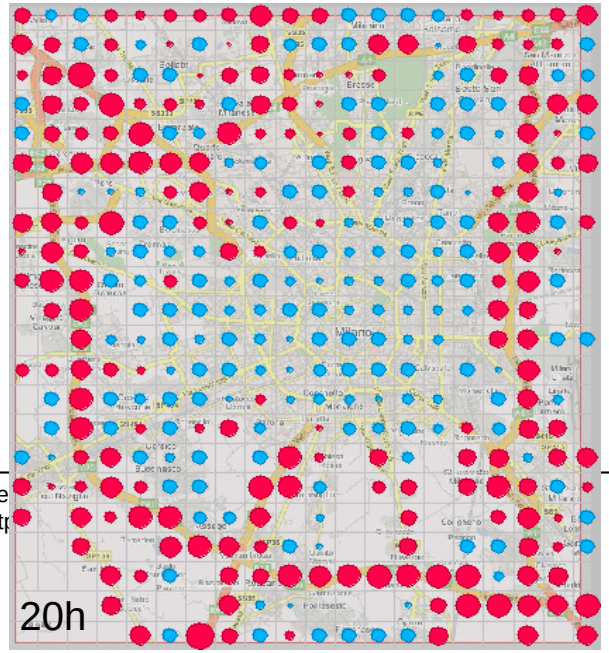
Circle area is proportional to value:

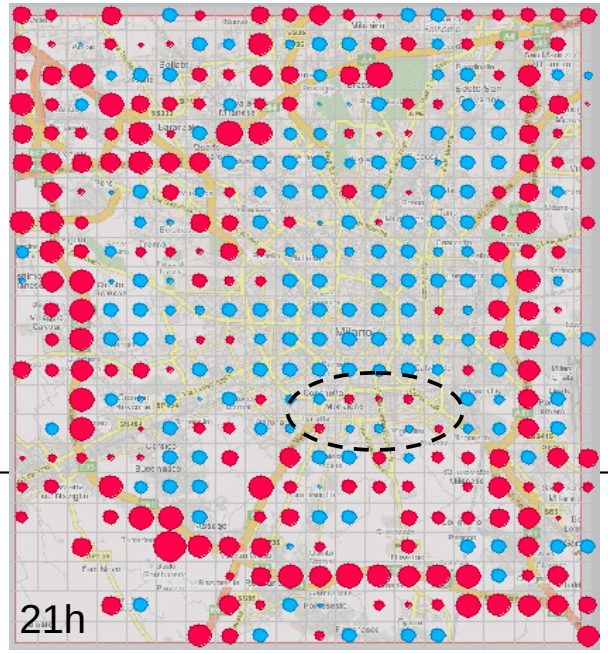


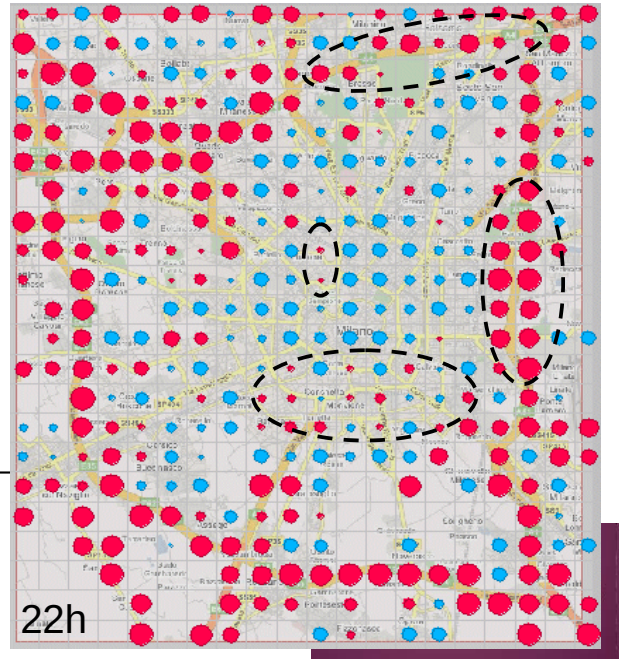








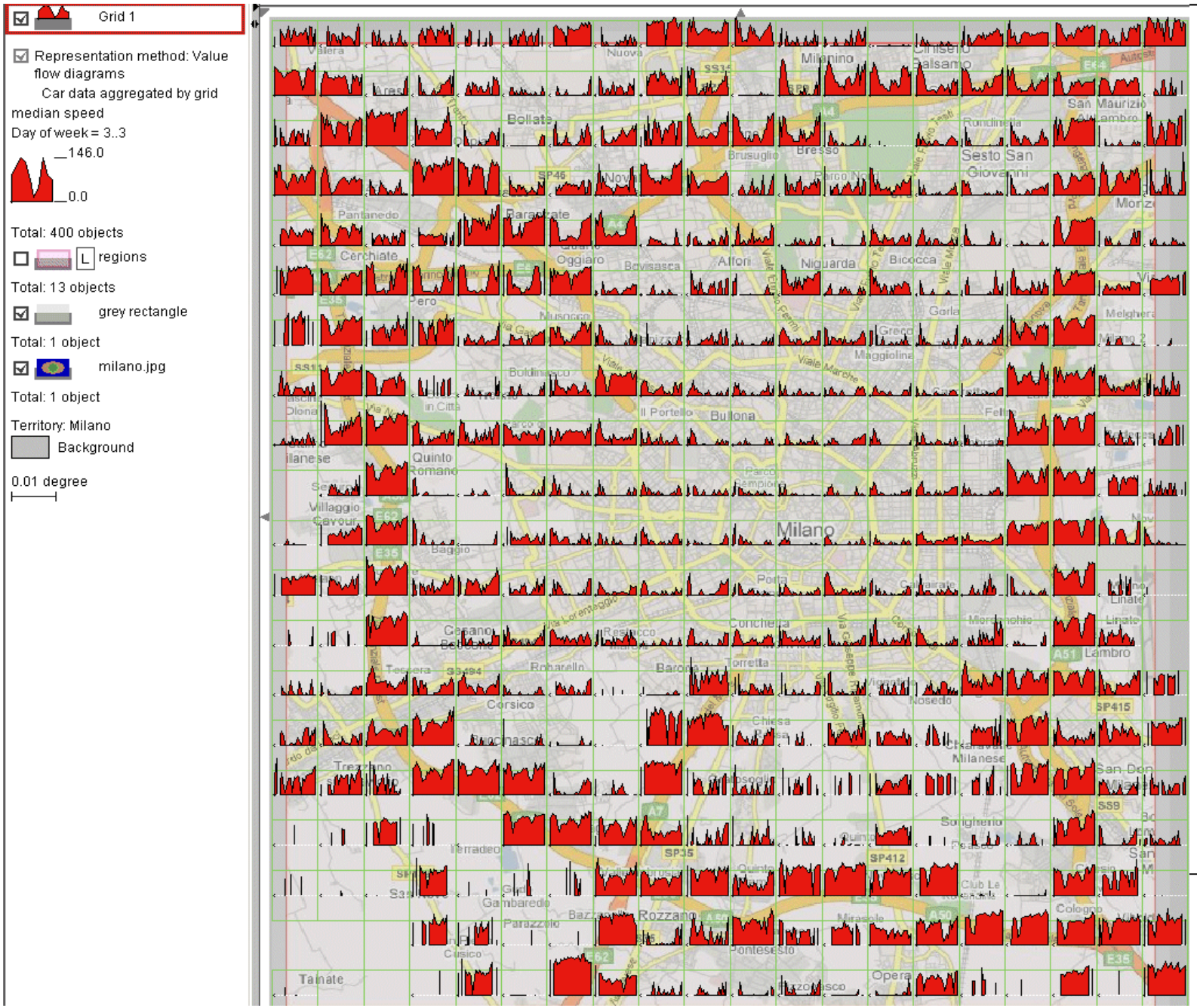




Ge
http

Map with value flow diagrams

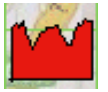
Attribute: median speed; day: Wednesday; hour: from 0 to 23



A frequent temporal pattern: significant drop of the speed in the morning and afternoon rush hours

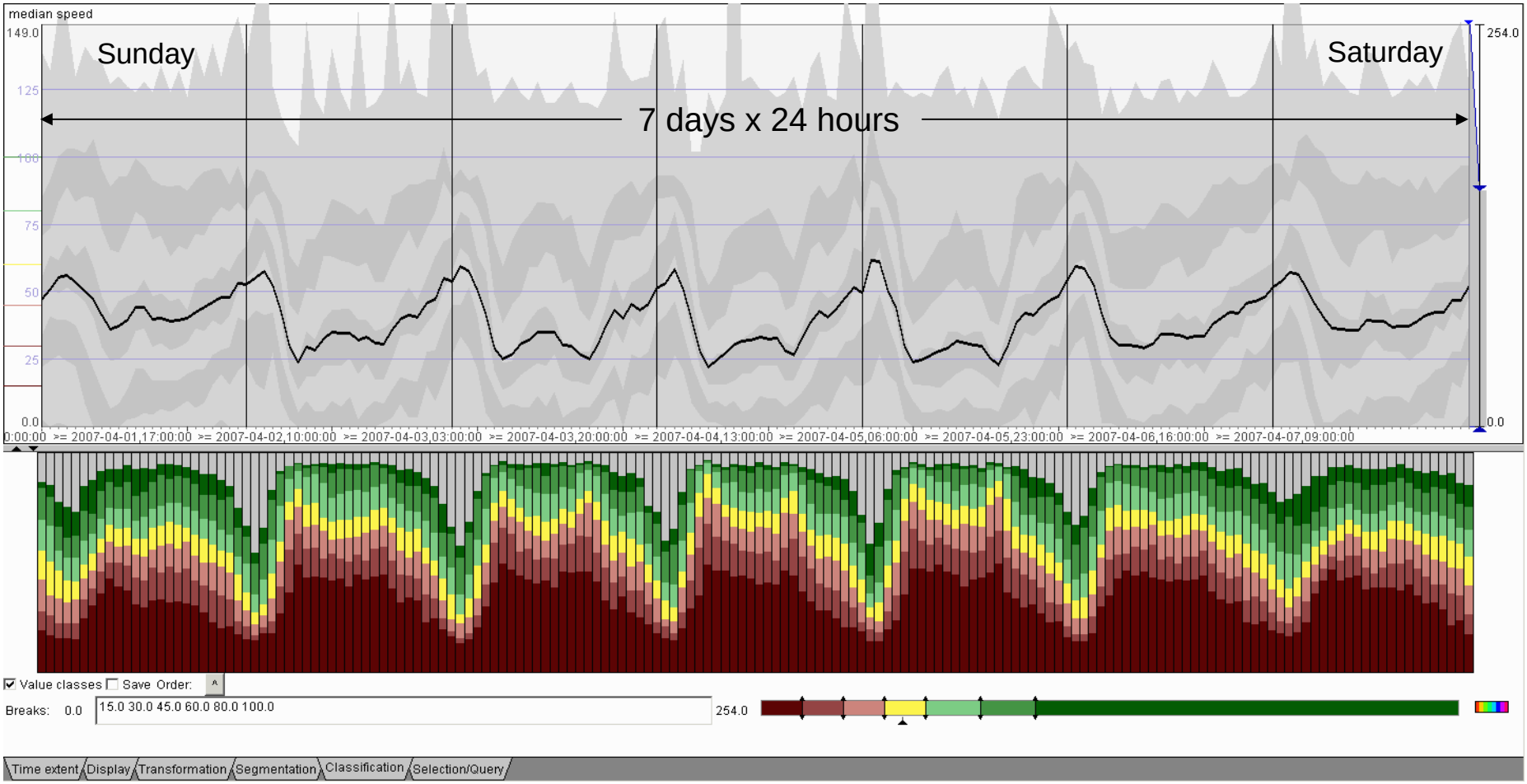
Map with value flow diagrams

Attribute: median speed; day: Saturday; hour: from 0 to 23



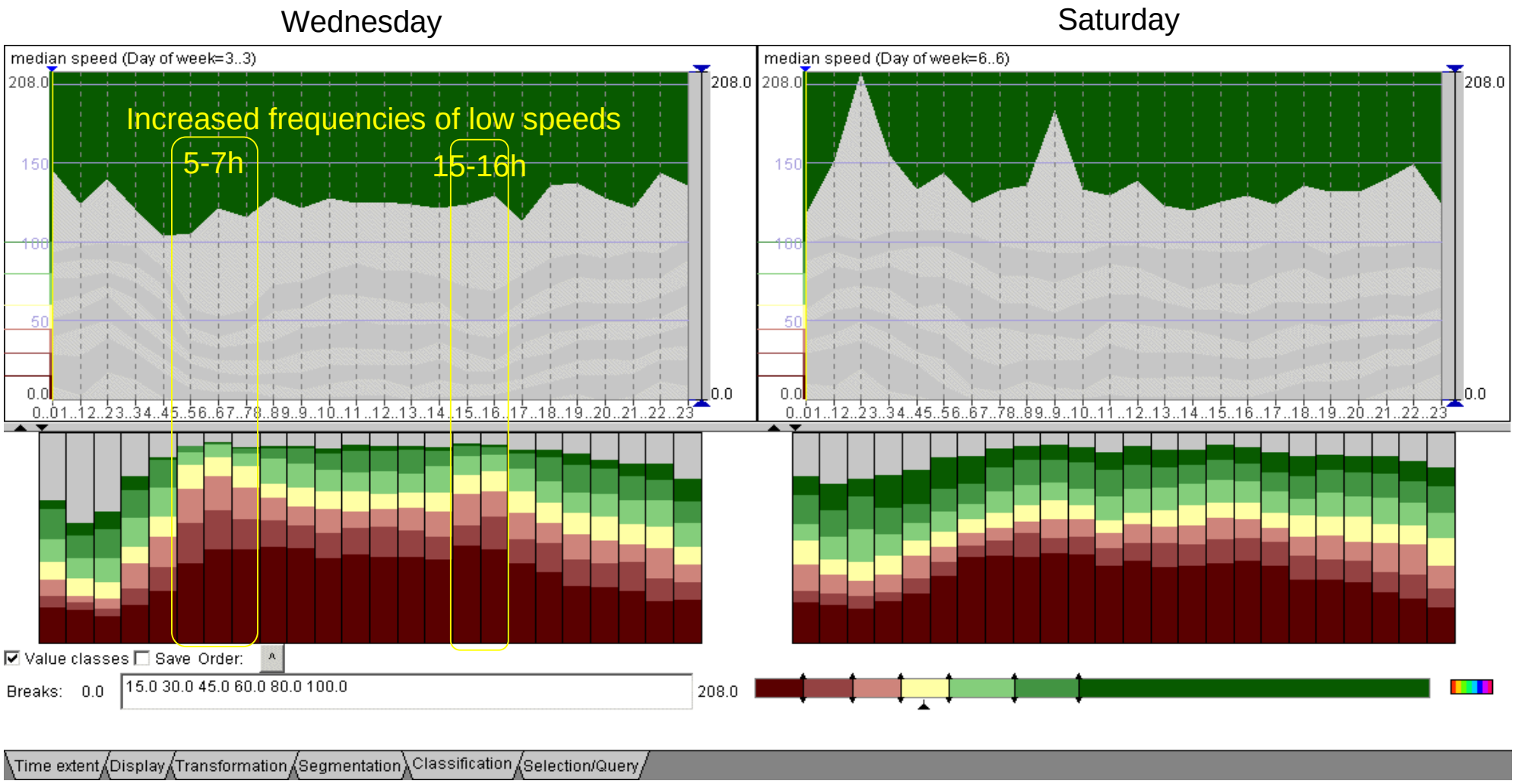
This temporal pattern does not occur on Saturday

Variation of the median speeds in all spatial compartments over the whole time period



Note the similarity of the daily patterns from Monday to Thursday and the difference of the Friday pattern

Variation of the median speeds over time in selected days



S × T – aggregation (example 2)

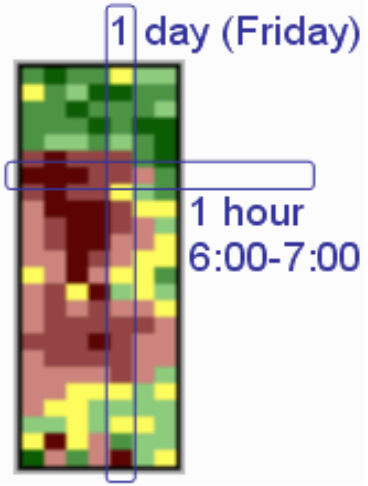
- S – aggregation: by cells of a regular grid
- T – aggregation: by days of the week × by hourly intervals in the day (**cyclic time**)
- Derived characteristics of the aggregates: number of cars, statistics of the speeds (min, max, mean, median, etc.)

Map with mosaic diagrams

Attribute: median speed; columns of the mosaic diagrams: days of the week; rows: hours of the day

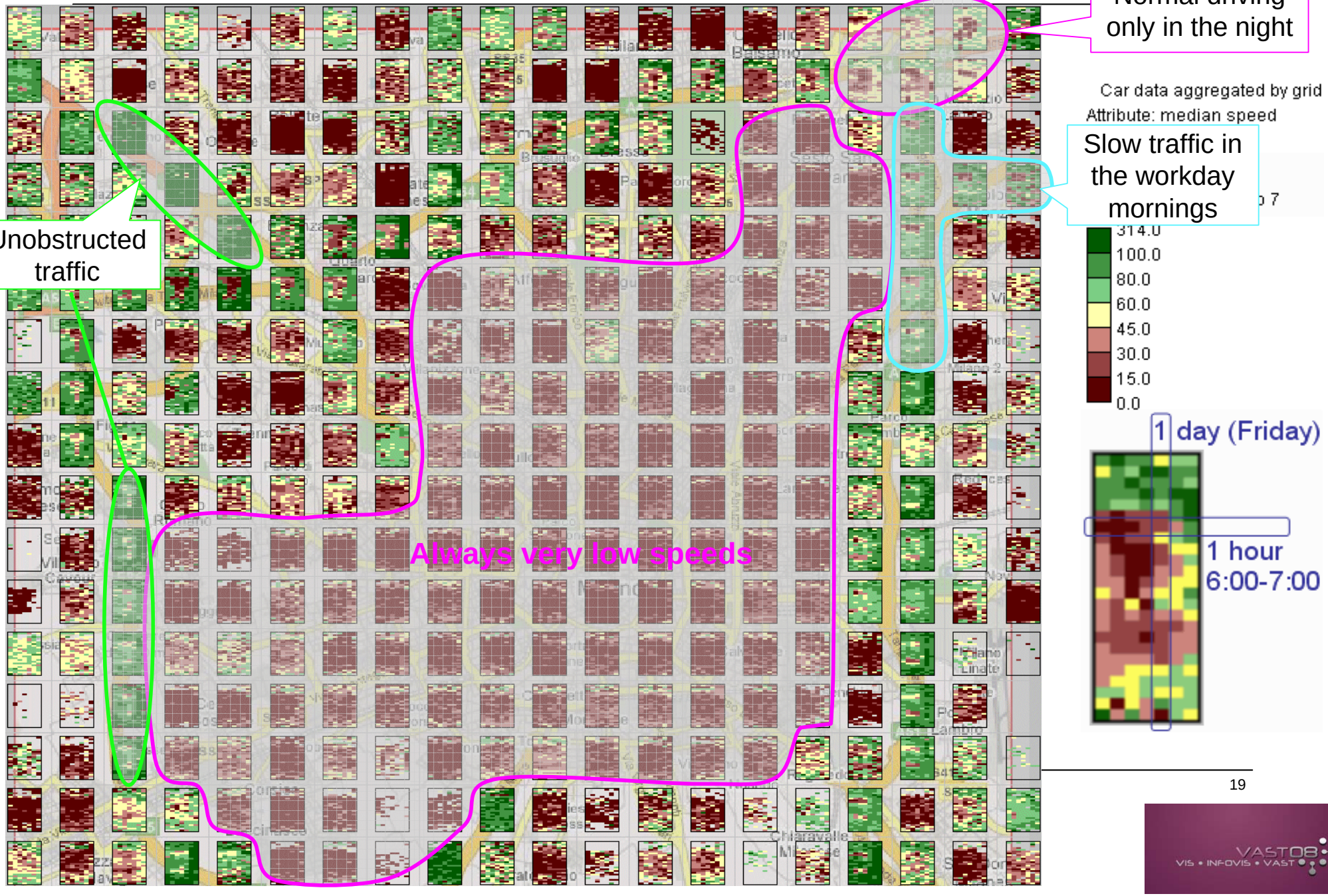
Car data aggregated by grid
Attribute: median speed
Parameters:
Hour: from 0 to 23
Day of week: from 1 to 7

314.0
100.0
80.0
60.0
45.0
30.0
15.0
0.0



Map with mosaic diagrams

Attribute: median speed; columns of the mosaic diagrams: days of the week; rows: hours of the day

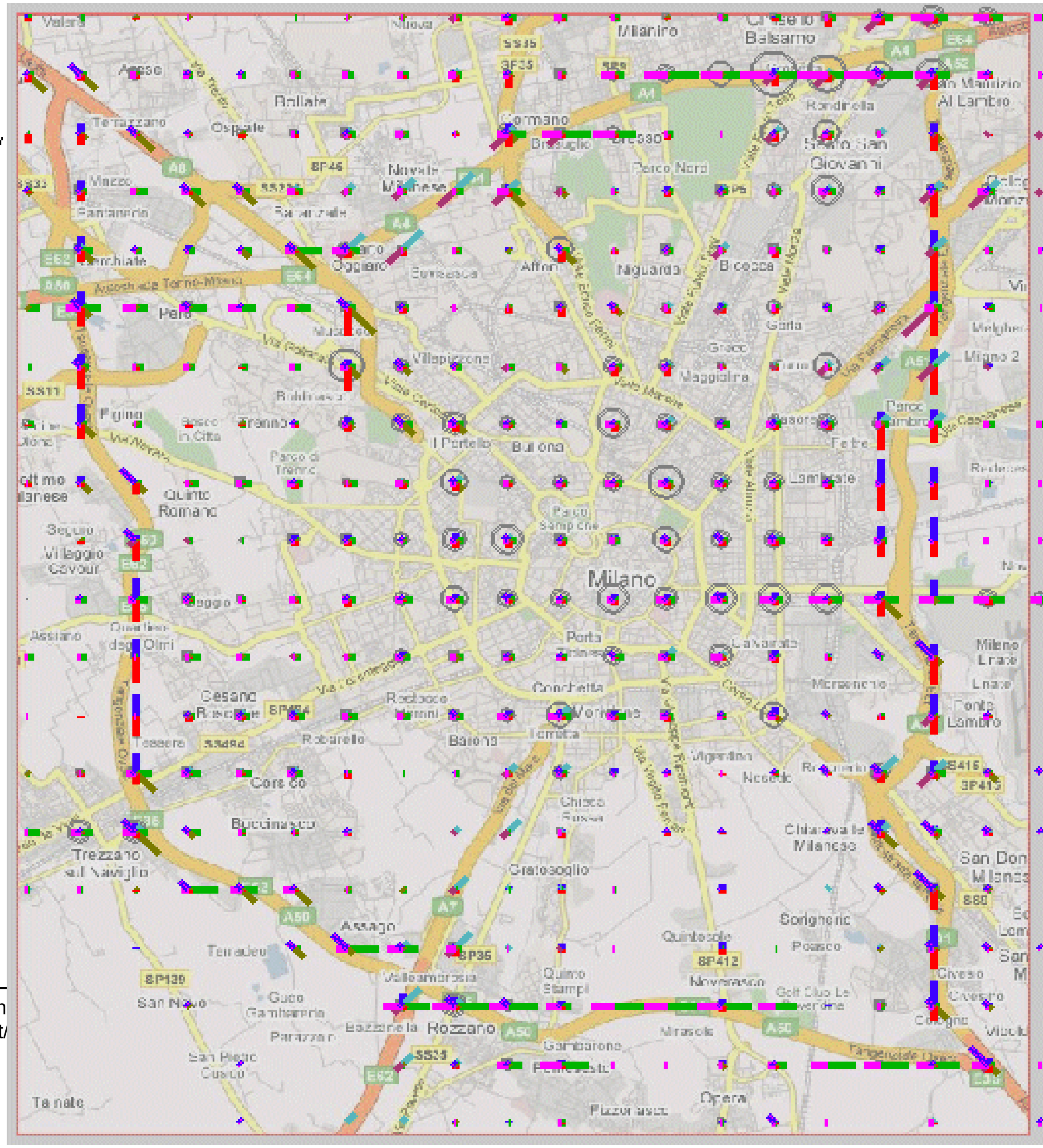


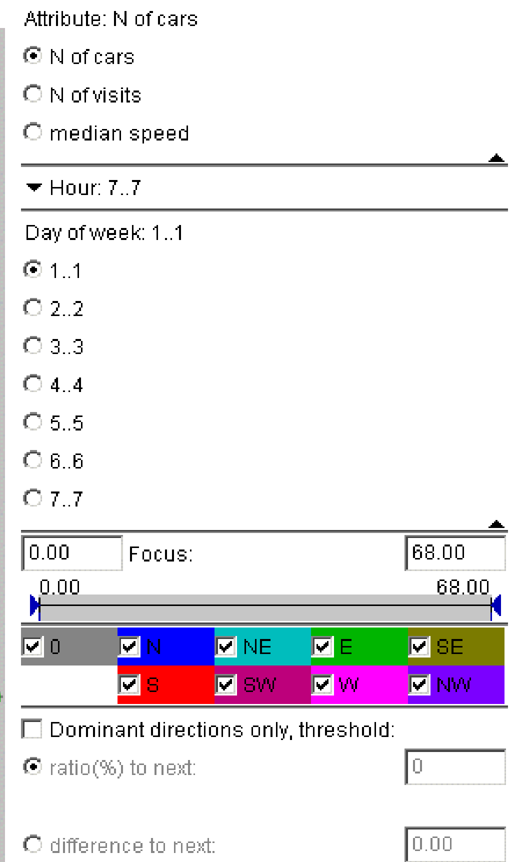
S × T × A - aggregation

- S – aggregation: by cells of a regular grid
- T – aggregation: by days of the week × by hourly intervals in the day (cyclic time)
- A – aggregation: by 8 movement directions (N, NE, E, SE, S, SW, W, NW) + absence of movement (speed below a threshold)
- Derived characteristics of the aggregates: number of cars, statistics of the speeds (min, max, mean, median, etc.)

The bar lengths are proportional to the **numbers of the cars** moving in the respective directions.
The radii of the circles are proportional to the numbers of the cars having the speed below the specified threshold (5km/h).
Selected day: Monday; selected hour: 7h

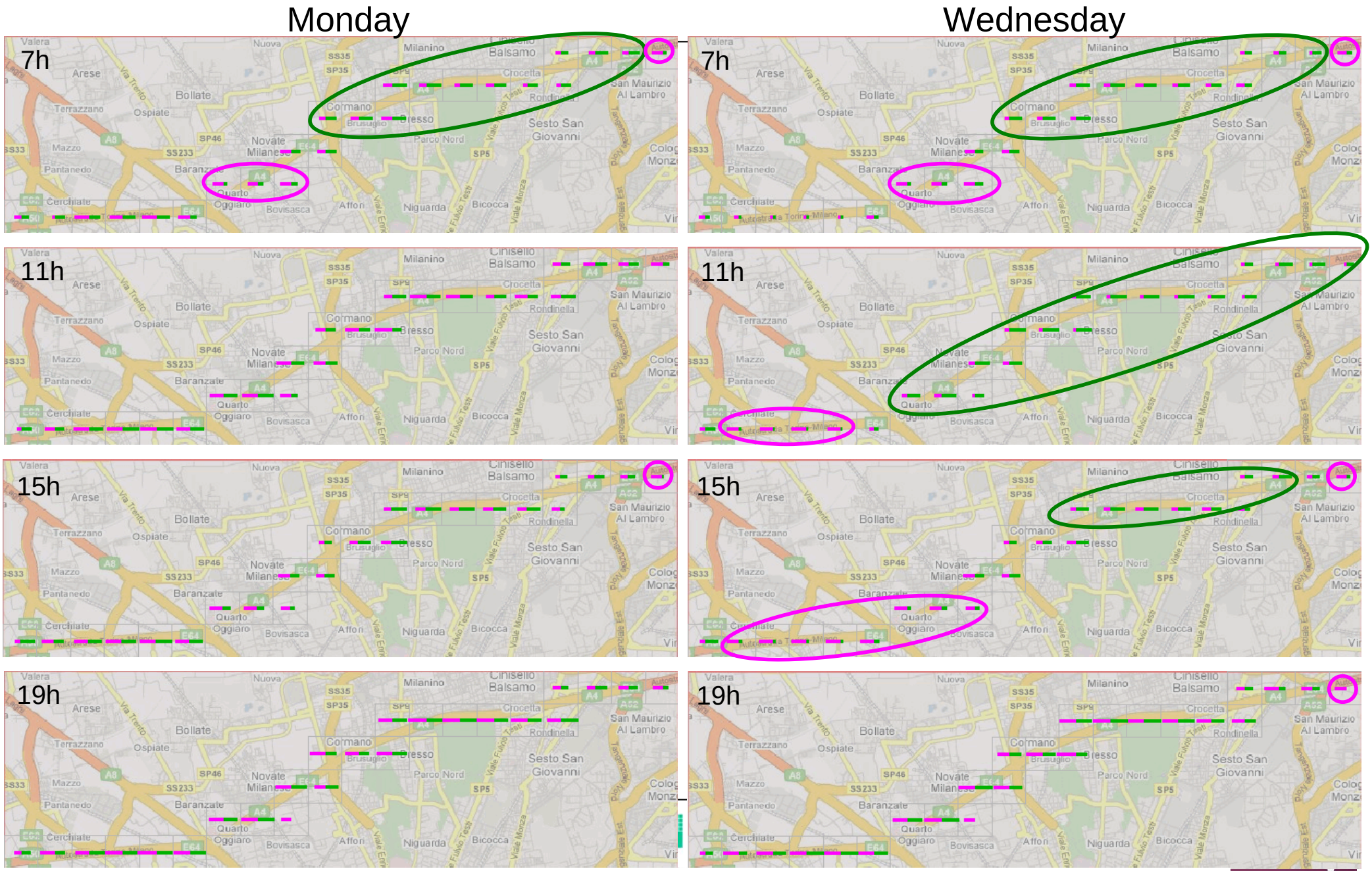
Car data aggregated by grid
N of cars
Hour = 7..7
Day of week = 1..1
No move (speed < 5.0, shown by circles) + 9 directions; all are shown
__68.00
__0





Display manipulator:
- Select aggregate attribute
- Select day of the week
- Select hour of the day
- Select directions
- Focus on a value interval
- Switch to showing only dominant directions and specify parameters

The bar lengths are proportional to the **median speeds** of the cars moving in the respective directions.



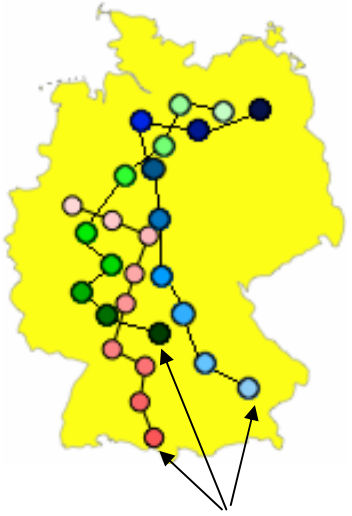
Note the asymmetry of some diagrams signifying different load of the street in two directions

Trajectory-oriented view

$\{ \mu_e(t) \mid e \in E \}$

Needed: aggregation and summarization of trajectories,
i.e. individual movement behaviors of different entities

Trajectory ≈ **line** in S × T - continuum
⇒ point-oriented approaches are not applicable



Trajectories

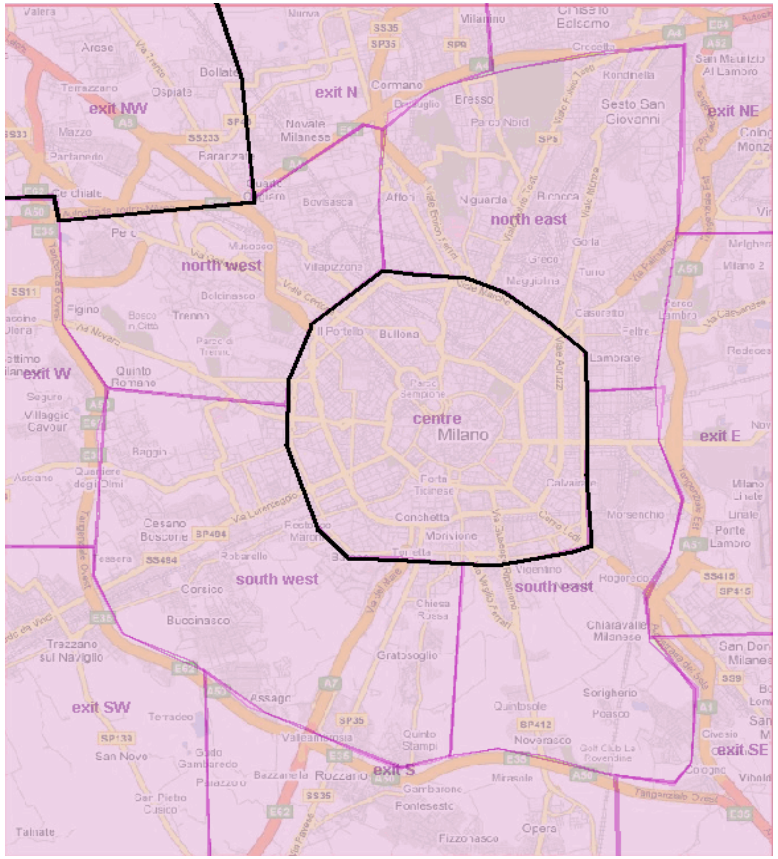
- Attributive (categorical) aggregation: by general characteristics of trajectories, e.g. duration, traveled distance, average speed, …
- Temporal aggregation: trajectories or fragments made during time intervals
- Spatial aggregation: based on spatial characteristics of trajectories
  - Start position
  - End position
  - Route

S × S × T × T - aggregation

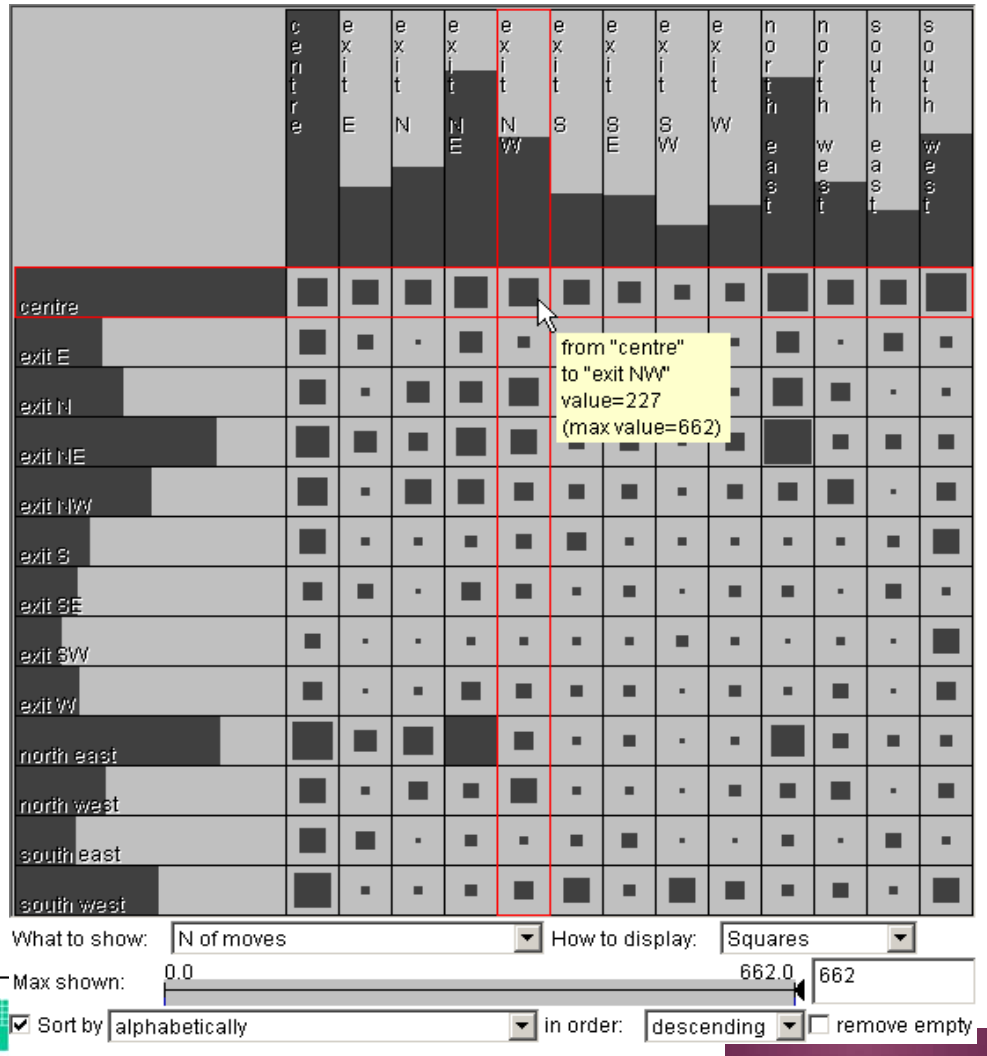
- S × S – aggregation:
  - The territory is divided into areas
  - The trajectories are grouped according to their start and end positions: start_area (tr1) = start_area (tr2) & end_area (tr1) = end_area (tr2)
  - ➢ Result: set of *aggregate moves* <$area_i$, $area_j$>
- T × T – aggregation:
  - The time period is divided into intervals
  - {The trajectories are divided into the fragments corresponding to these intervals}
  - The trajectories/fragments are grouped according to their start and end times: start_interval (tr1) = start_interval (tr2) & end_interval (tr1) = end_interval (tr2)
  - ➢ Result: set of *aggregate moves* <$area_i$, $area_j$, [t_k, t_{k+1}]>
- Derived characteristics of the aggregate moves: number of trajectories, statistics of the trajectory attributes (duration, distance, speed, …)

S×S-aggregation: trajectories are grouped by the origins and destinations

1. The start and end points of the trajectories are referred to areas in space

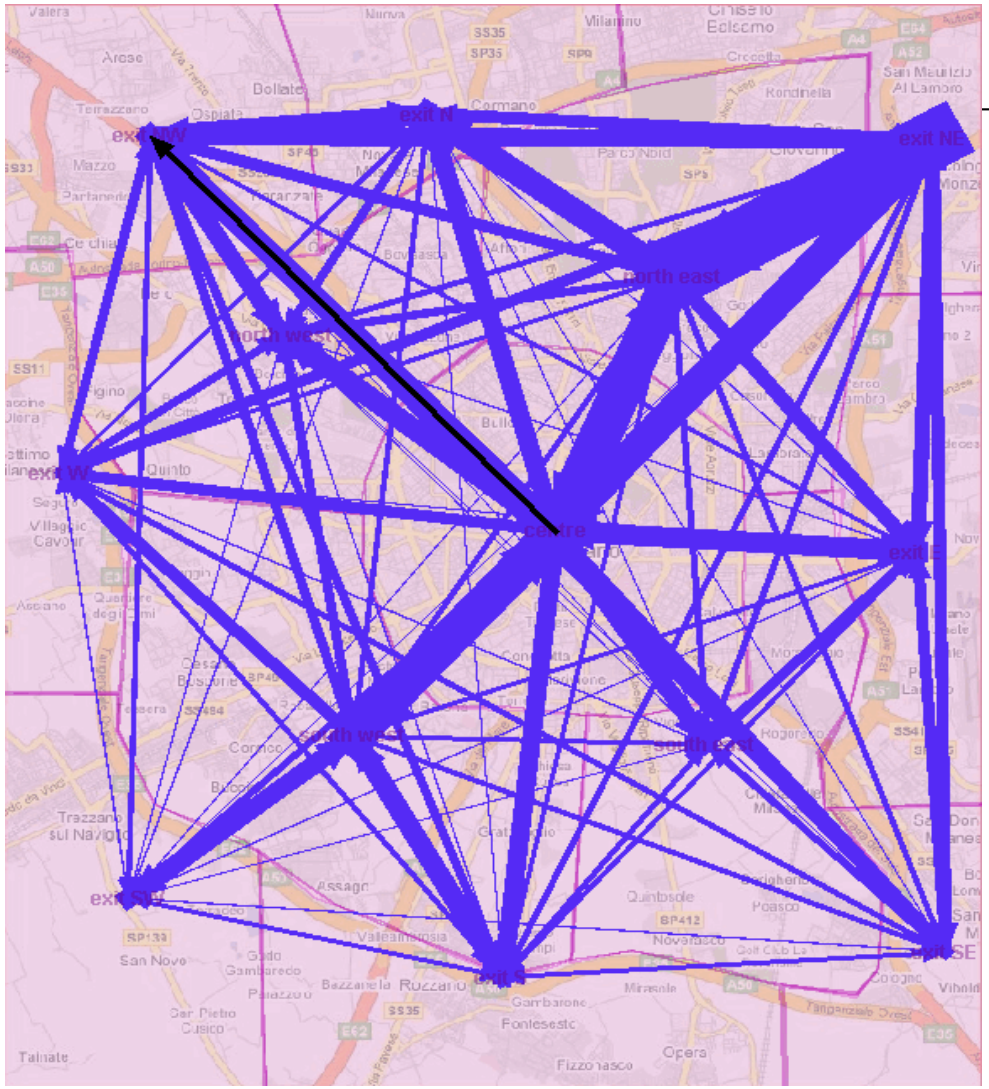


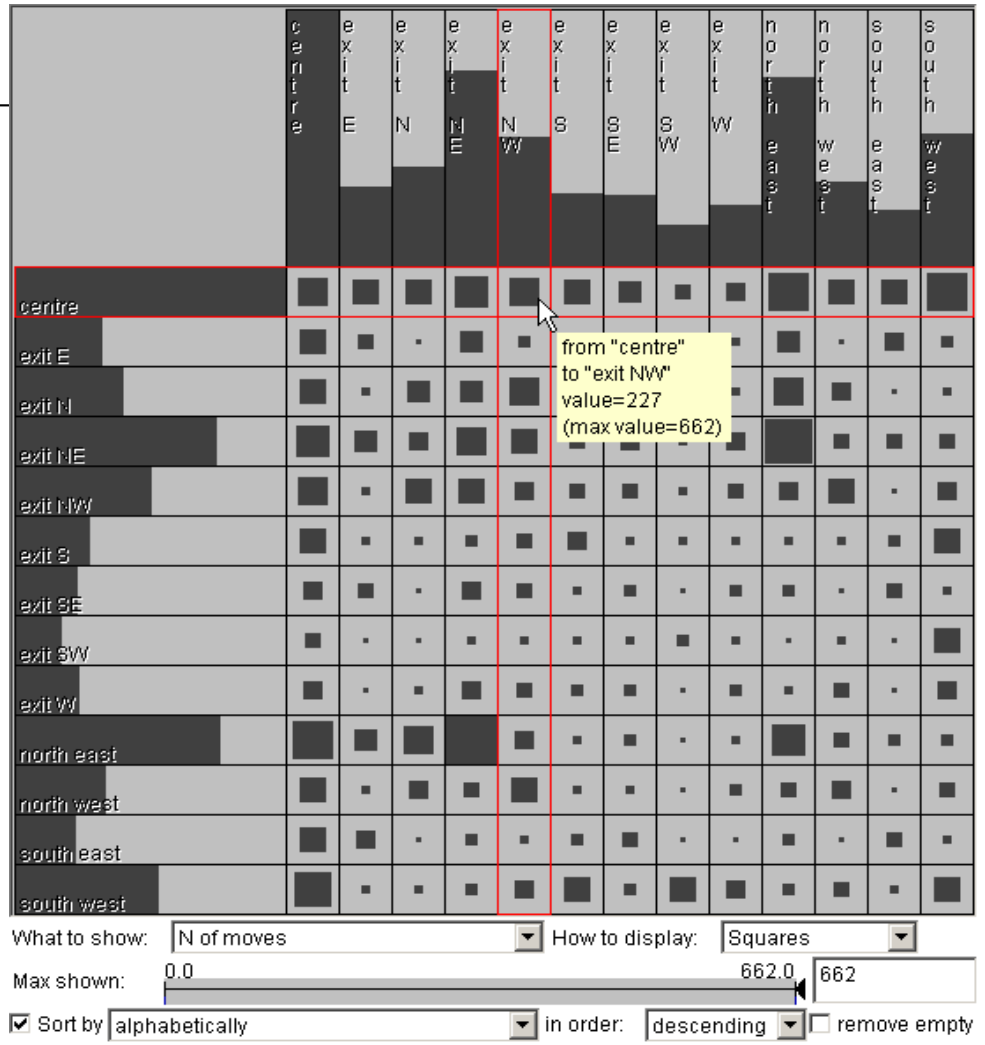
2. For each pair of areas, the number of trajectories and statistics of their numeric properties (duration, distance) are computed



Gennady & Natalia Andrienko
http://geoanalytics.net/and

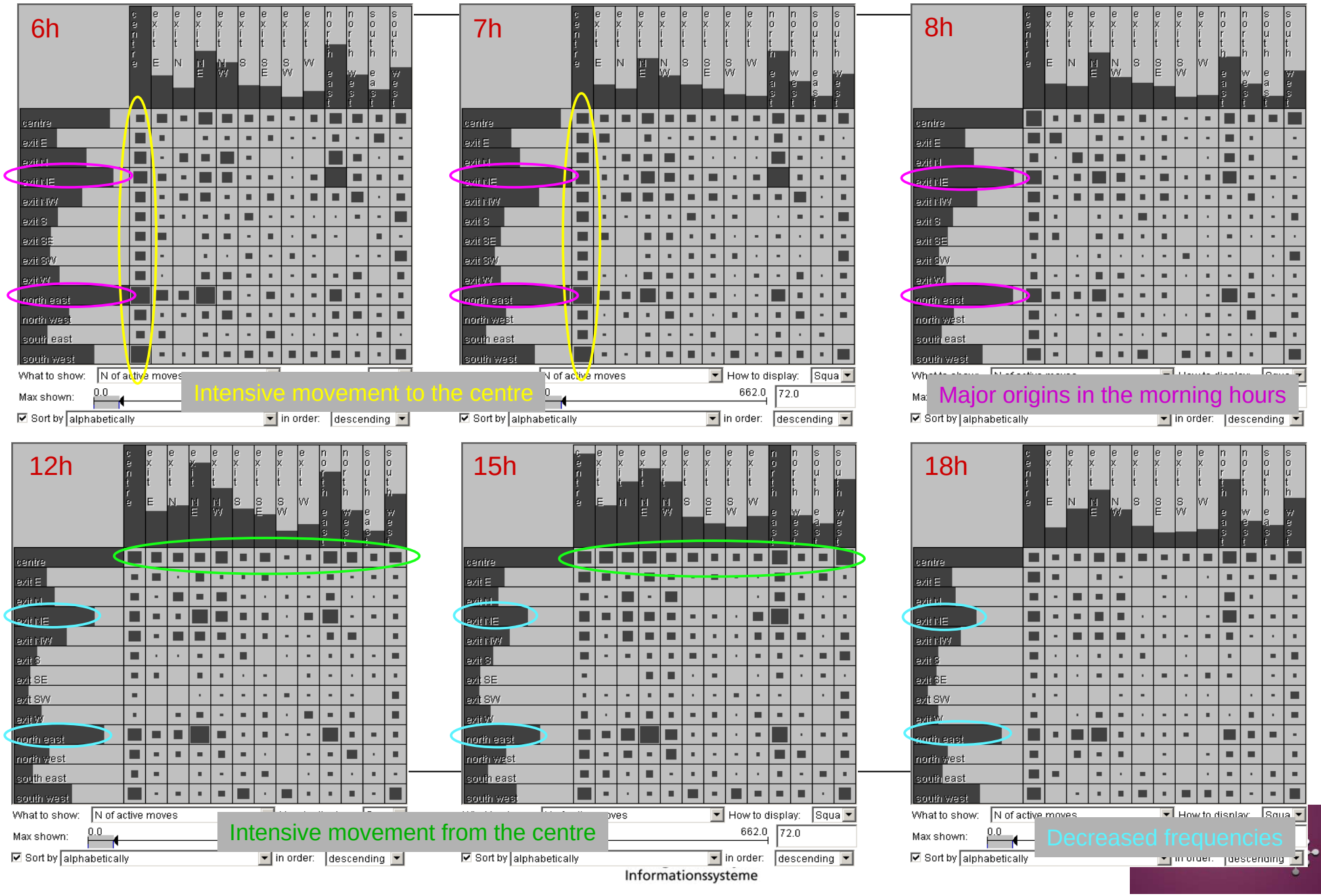
Fraunhofer Institut Intelligente Analyse- und Informationssysteme



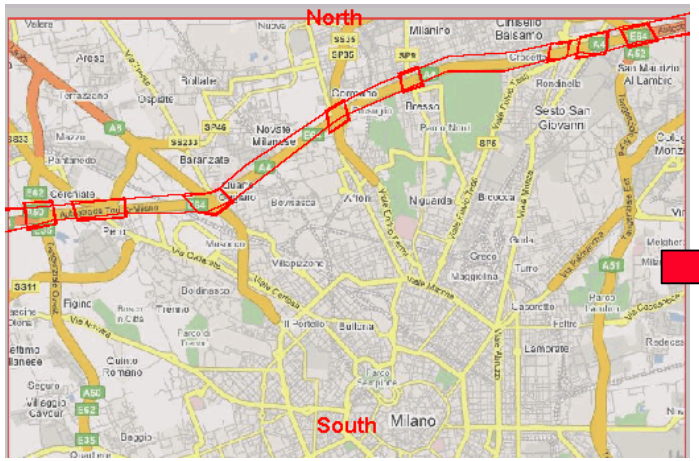


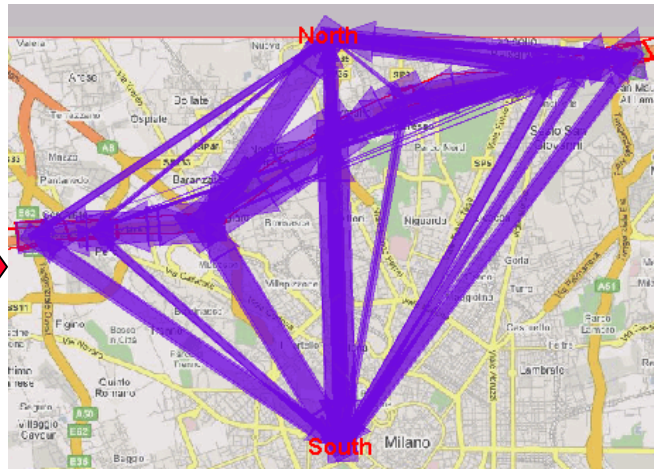
The aggregated data can also be represented on a map by directed lines (vectors) with the thicknesses proportional to the numbers of the trajectories or to the values of another aggregate attribute. However, the origin-destination matrix may be preferable.

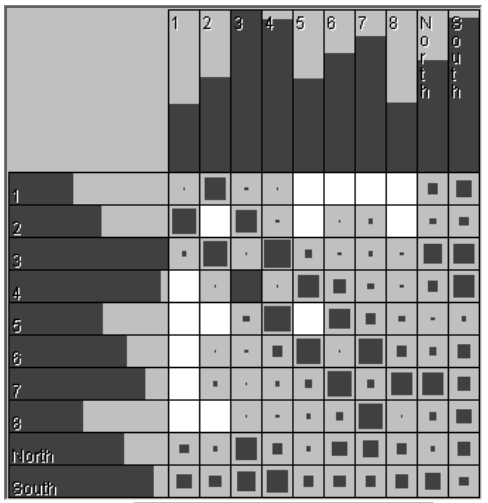
S×S×T×T-aggregation: time intervals of the length of 1 hour

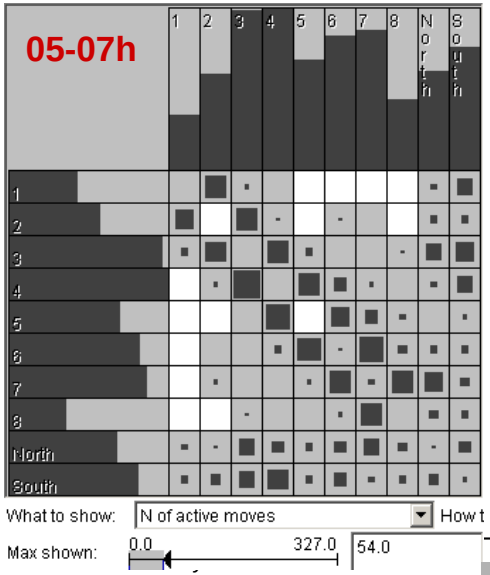


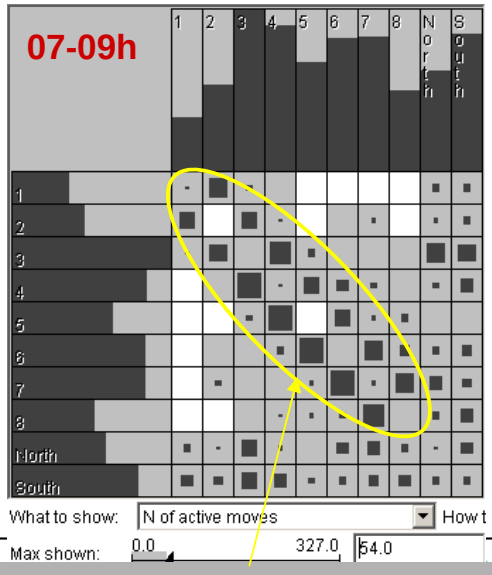
Investigation of the movement through a particular road

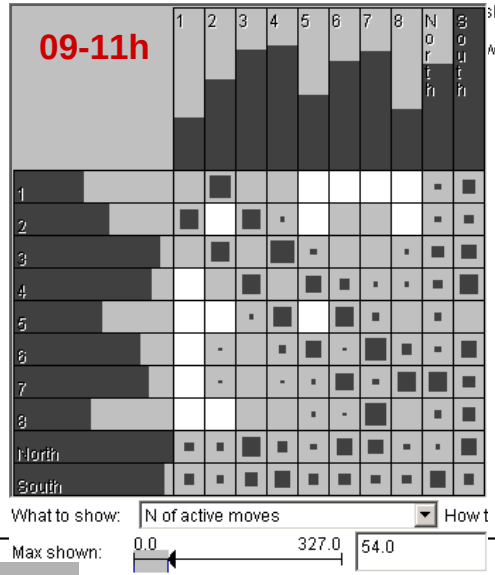












. . .

http://geoanalytics.net/and

Asymmetric movement in two directions

Fraunhofer Institut Intelligente Analyse- und Informationssysteme

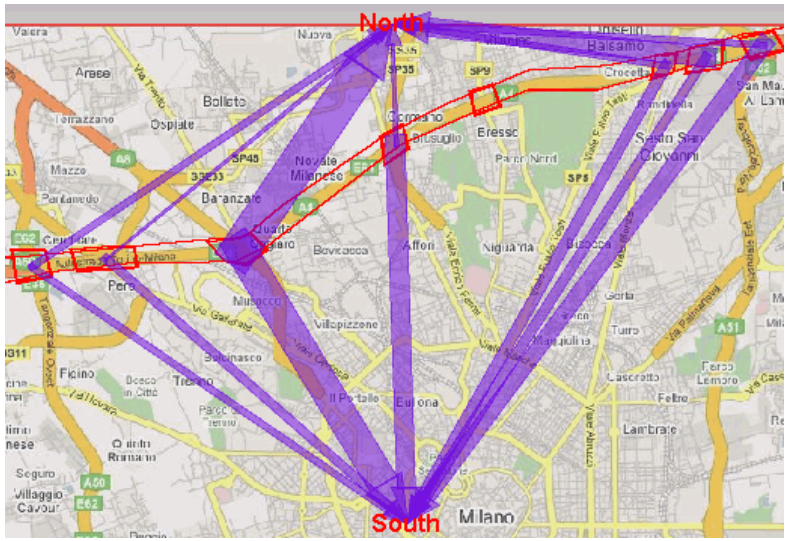
VAST08 VIS • INFOVIS • VAST

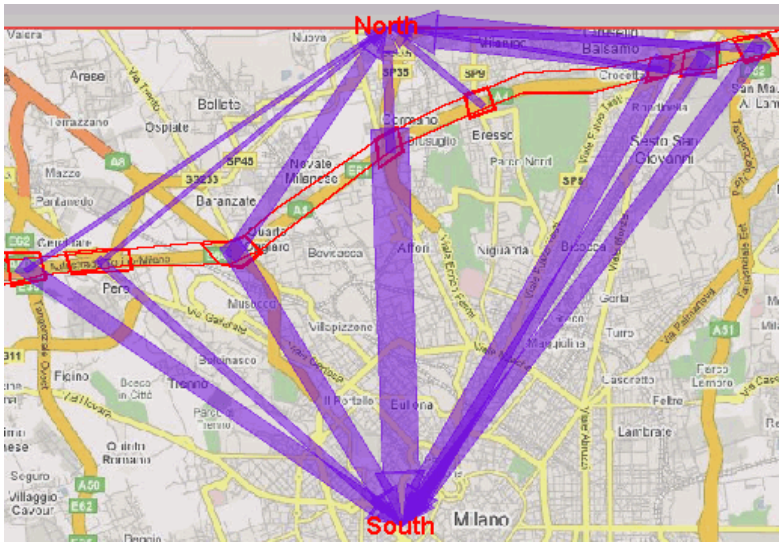
Focusing on the cars leaving the road or entering the road

07-09h

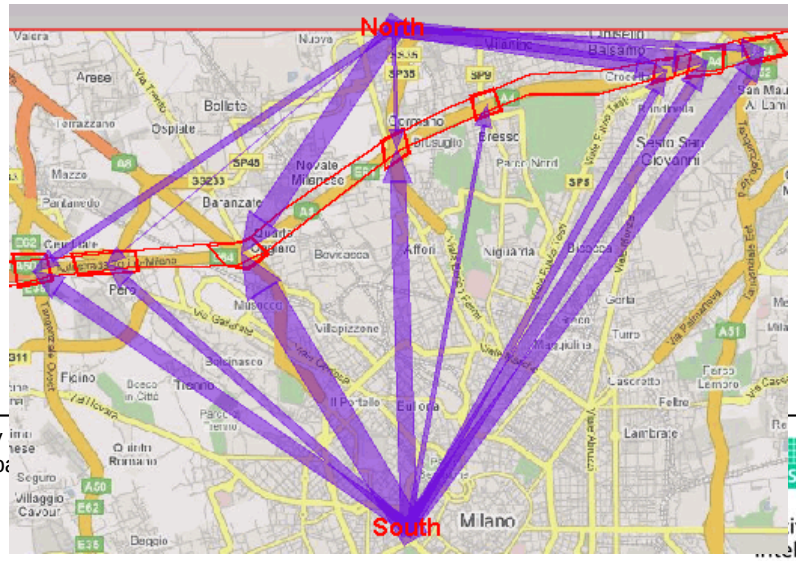
09-11h

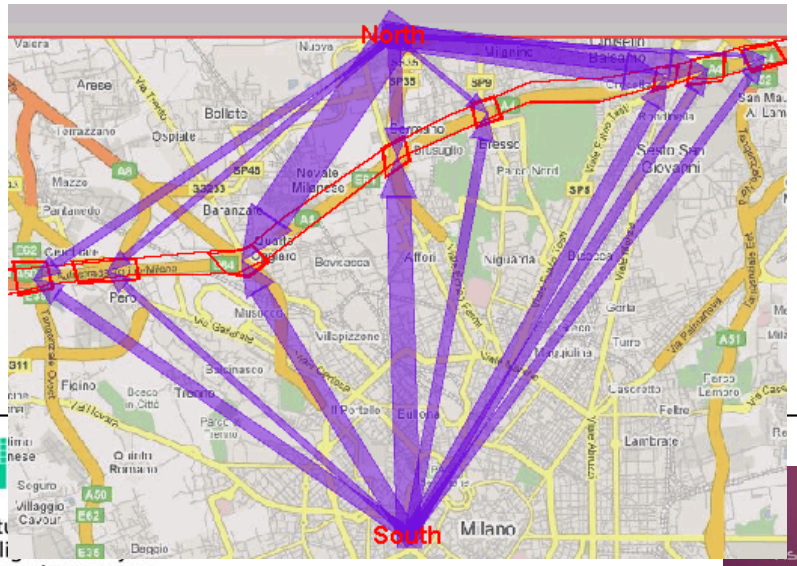
Out





In



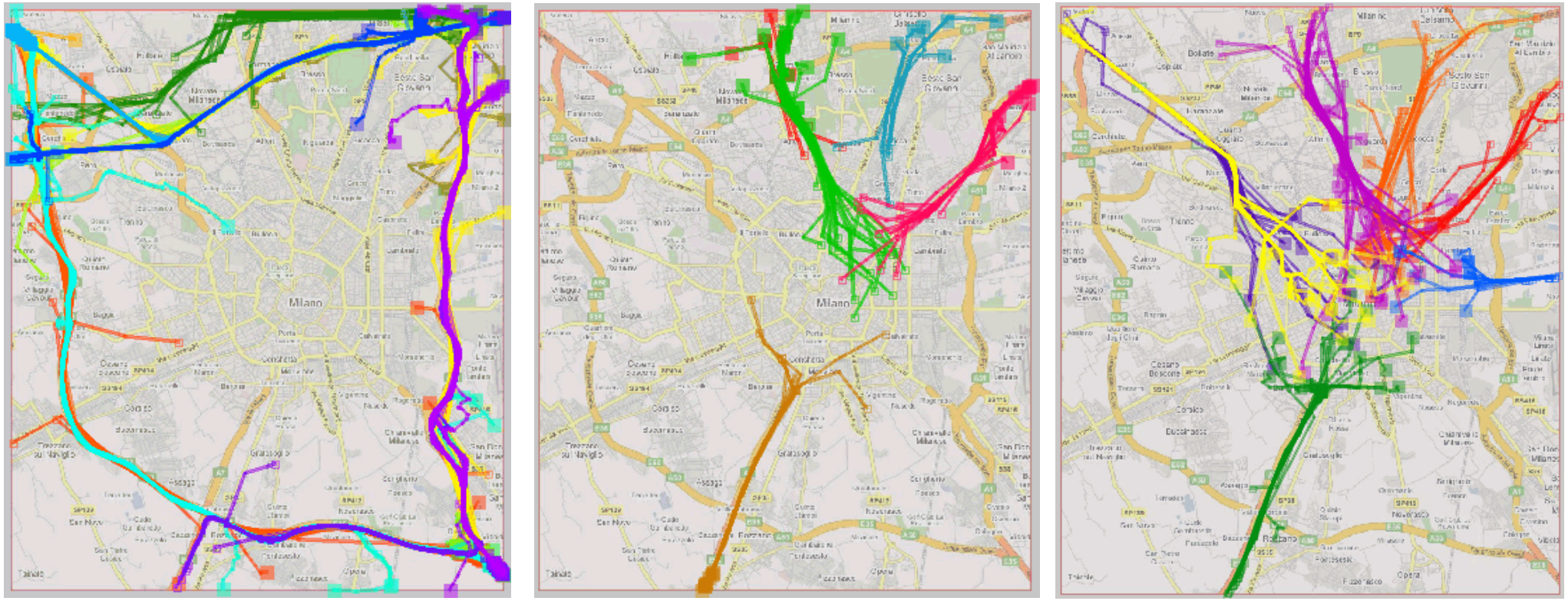


Informationssysteme

R (route-based) - aggregation

- Basic idea: group trajectories according to their **routes** (geometric similarity + spatial proximity)
- Approach: use clustering techniques with trajectory-specific distance functions
  - Rinzivillo, S., Pedreschi, D., Nanni, M., Giannotti, F., Andrienko, N., & Andrienko, G.: Visually–driven analysis of movement data by progressive clustering, Information Visualization, 7(3), 2008

Examples of clusters of trajectories



Research question: how to summarize and visualize groups of similar trajectories?

A possible approach:
R × S × S – aggregation and R × S × S × T × T – aggregation

- R – aggregation: by routes
- S × S – aggregation:
  - By pairs of generalized positions (areas)
  - ➢ Result: set of *aggregate moves* <$area_i$, $area_j$>
- T × T – aggregation:
  - By time intervals (start time + end time)
  - ➢ Result: set of *aggregate moves* <$area_i$, $area_j$, [t_k, t_{k+1}]>
- Derived characteristics of the aggregate moves: number of trajectories, statistics of the trajectory attributes (duration, distance, speed, …)

Deriving generalized positions

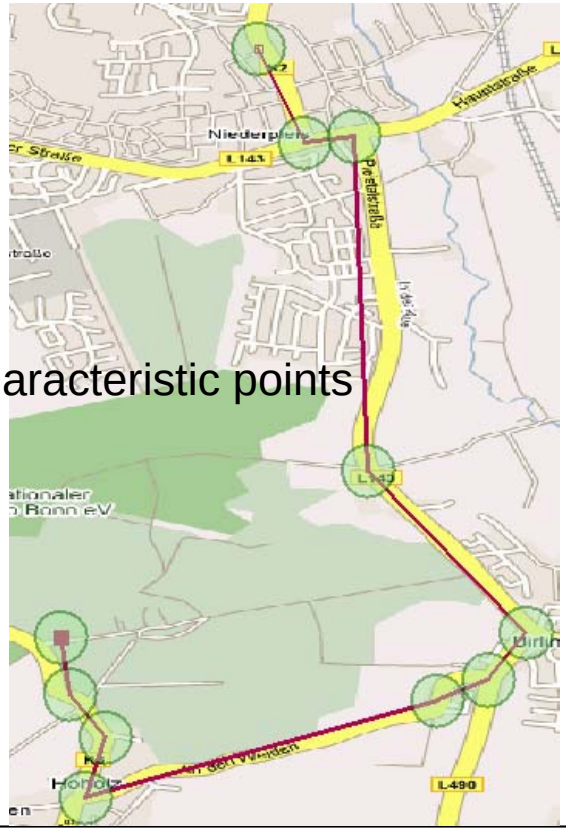
1) Extract characteristic points from trajectories, i.e. points of significant turns and stops

2) Build areas (e.g. circles) around groups of points and isolated points

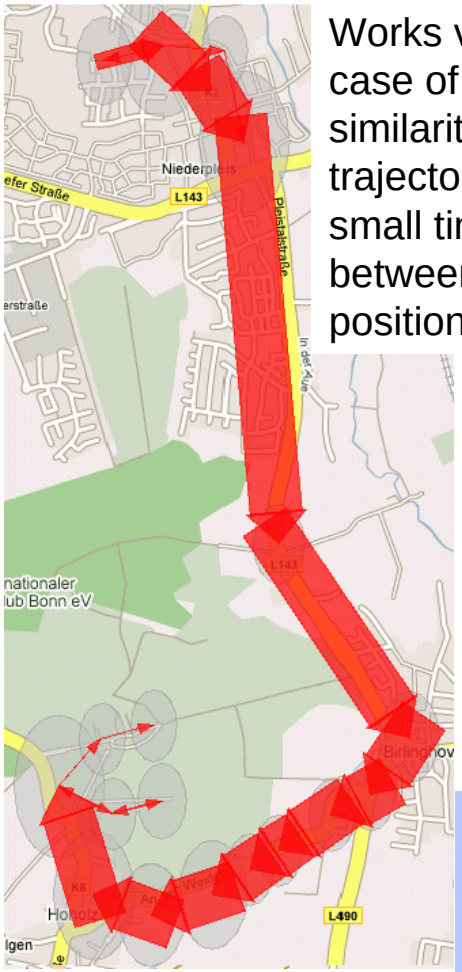
Original trajectory

Simplified trajectory

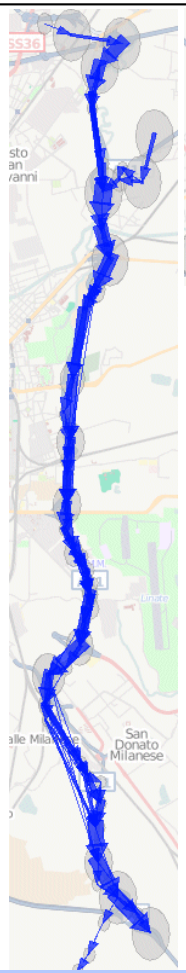
Characteristic points



R × S × S – aggregation

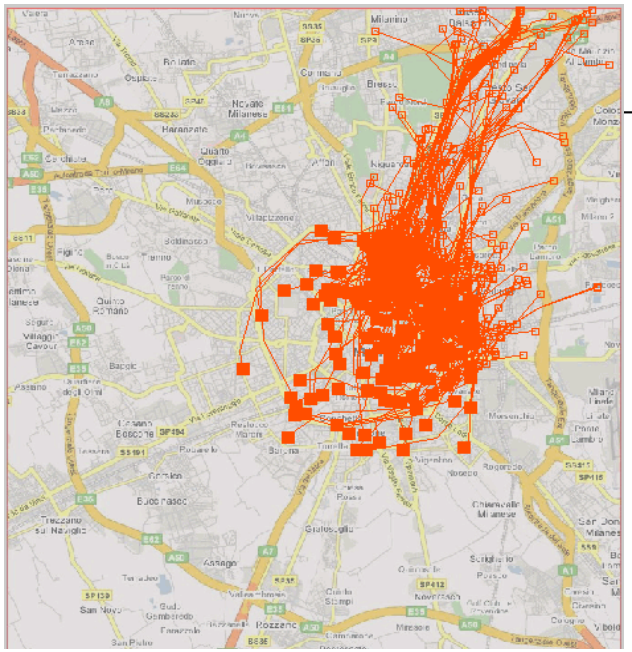


Works very well in case of very high similarity between trajectories and small time intervals between the position records

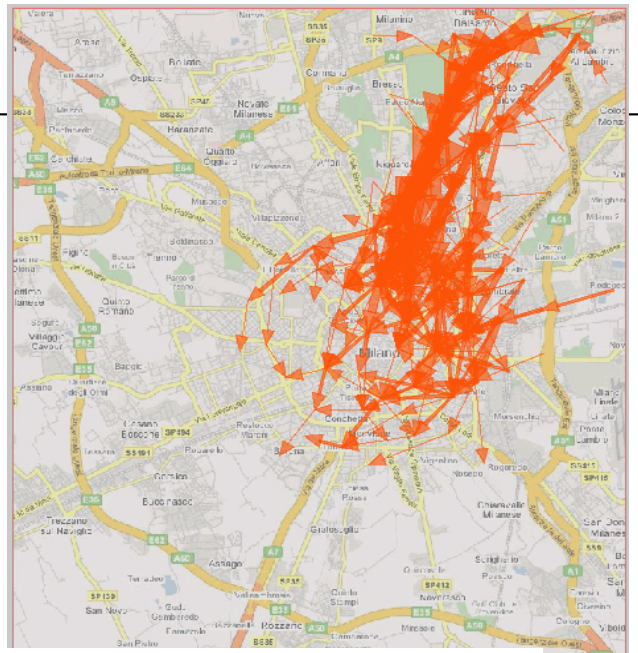


Works not so well in case of high variability of the trajectories and/or longer or unequal time intervals between the position records

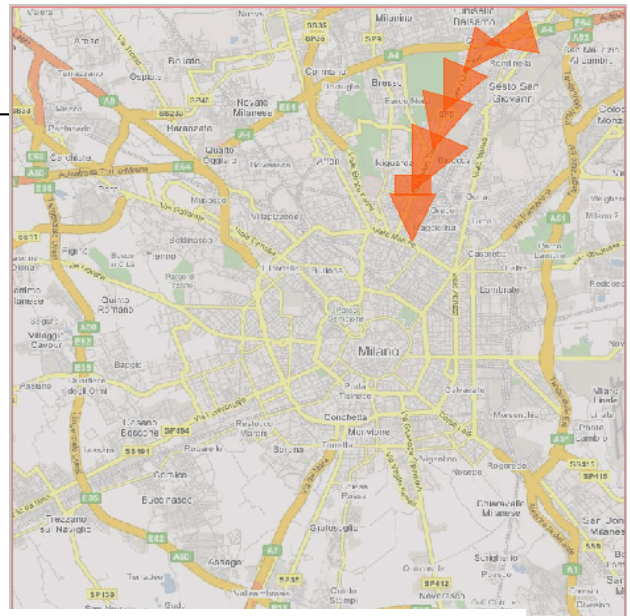
…but may still be very useful in combination with interactive filtering of the aggregate moves according to their characteristics, in particular, number of trajectories *(an example follows)*



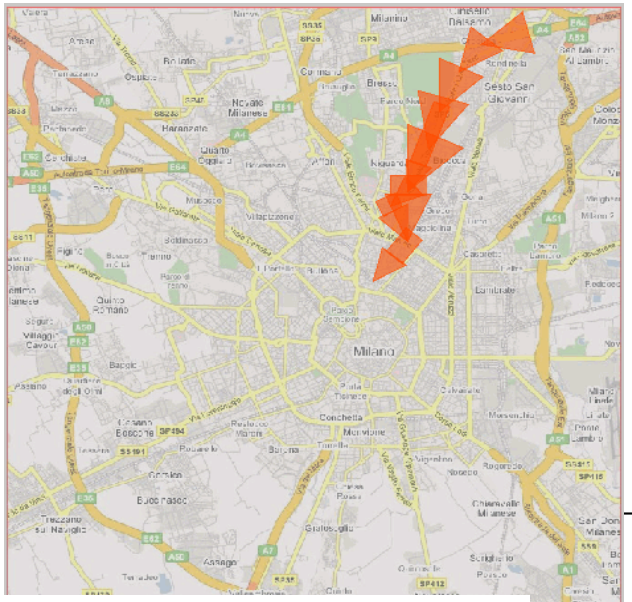
An example of a cluster of trajectories



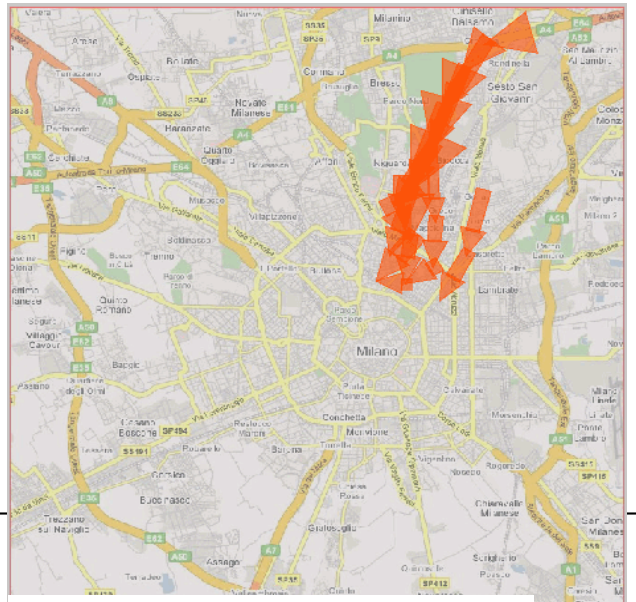
Aggregate moves; all are visible



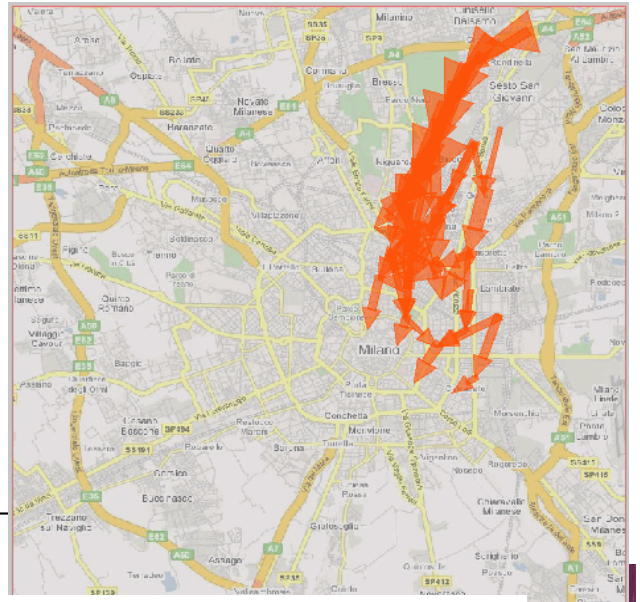
Aggregate moves occurring in at least 20 trajectories



Aggregate moves occurring in at least 15 trajectories



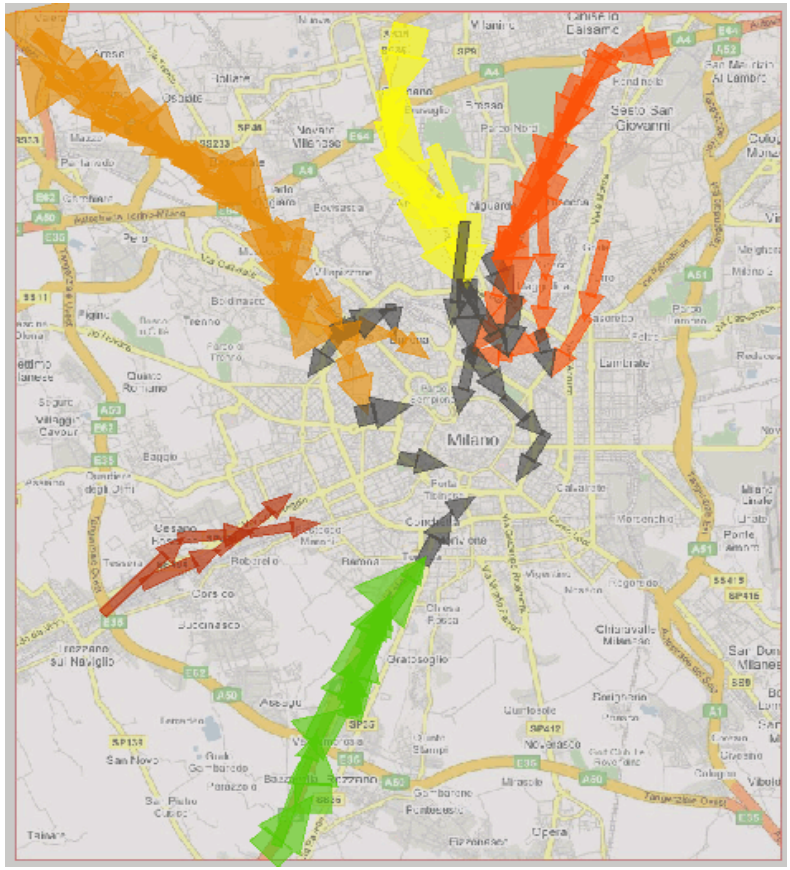
Aggregate moves occurring in at least 10 trajectories



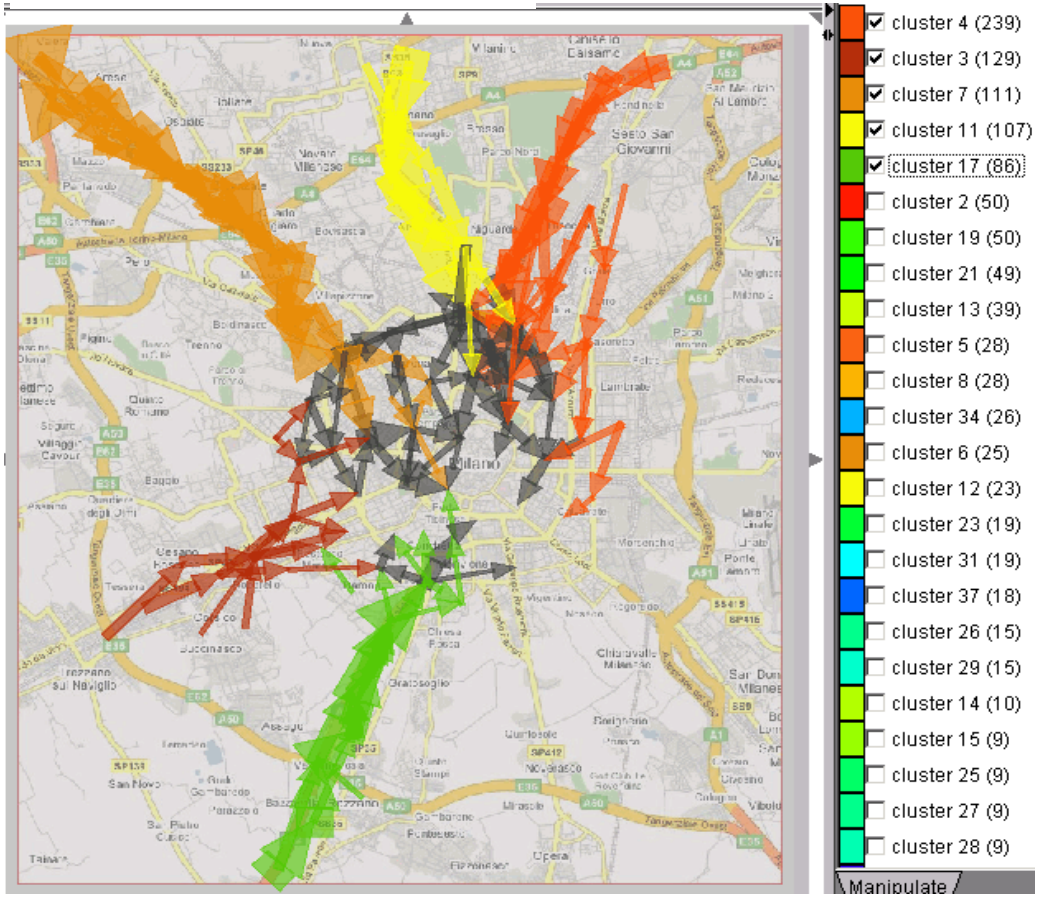
Aggregate moves occurring in at least 5 trajectories

5 biggest clusters of trajectories towards the city centre

Aggregate moves occurring in at least 10 trajectories

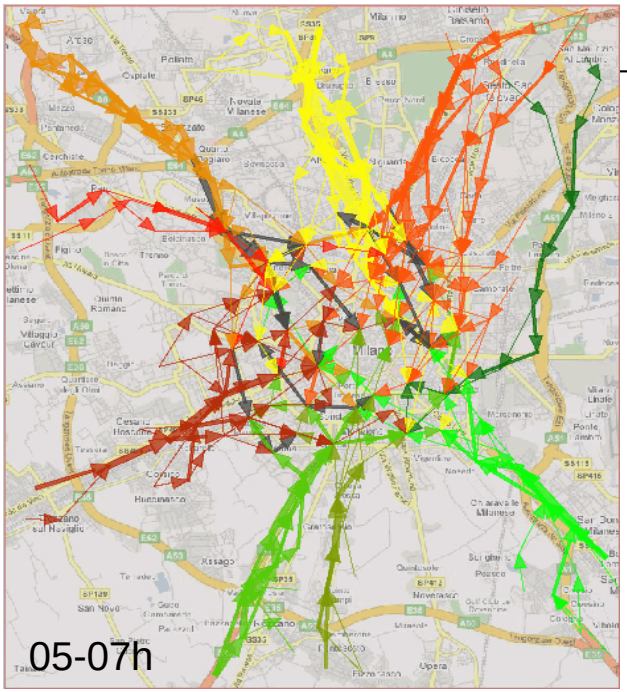


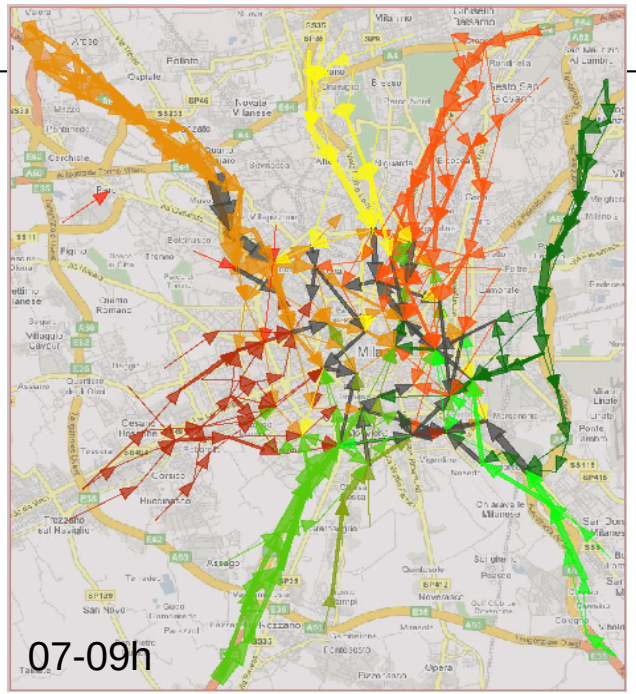
Aggregate moves occurring in at least 5 trajectories

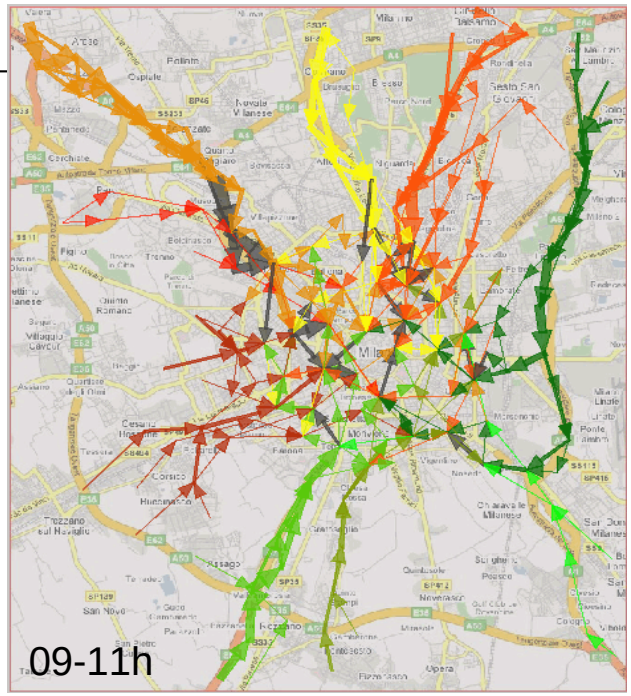


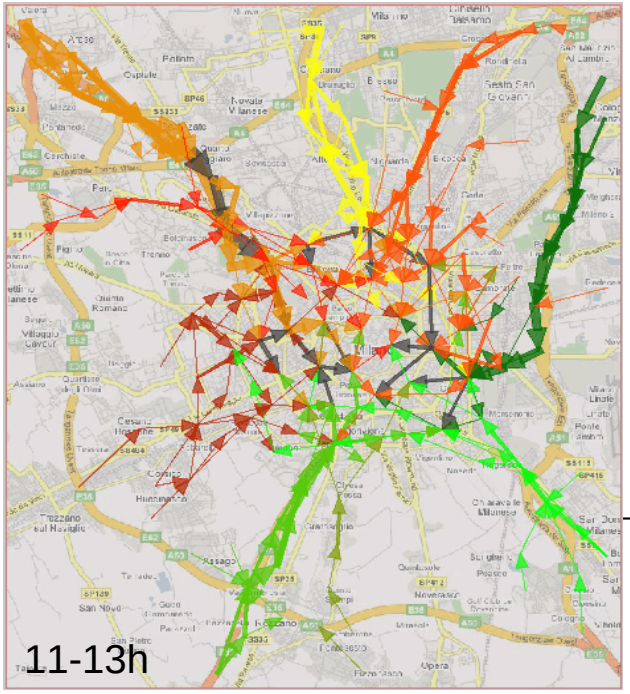
Dark grey: moves occurring in trajectories from several clusters

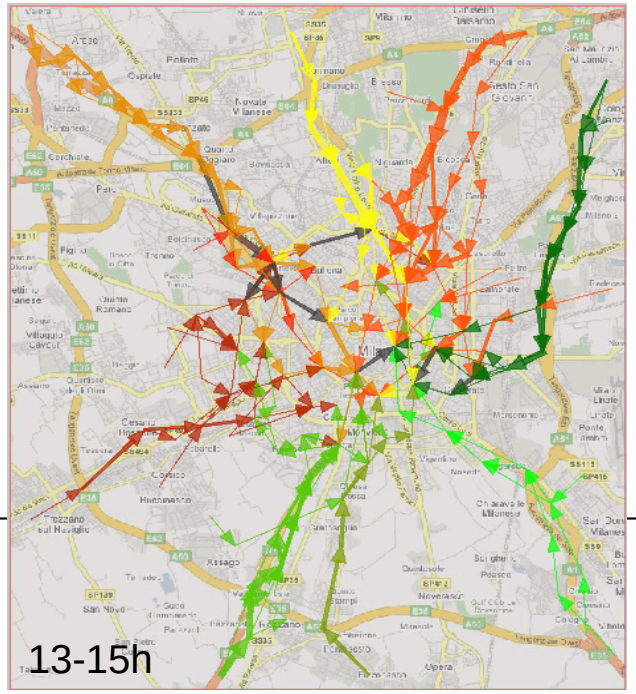
R × S × S × T × T – aggregation: the use of the most popular routes towards the centre by times of the day

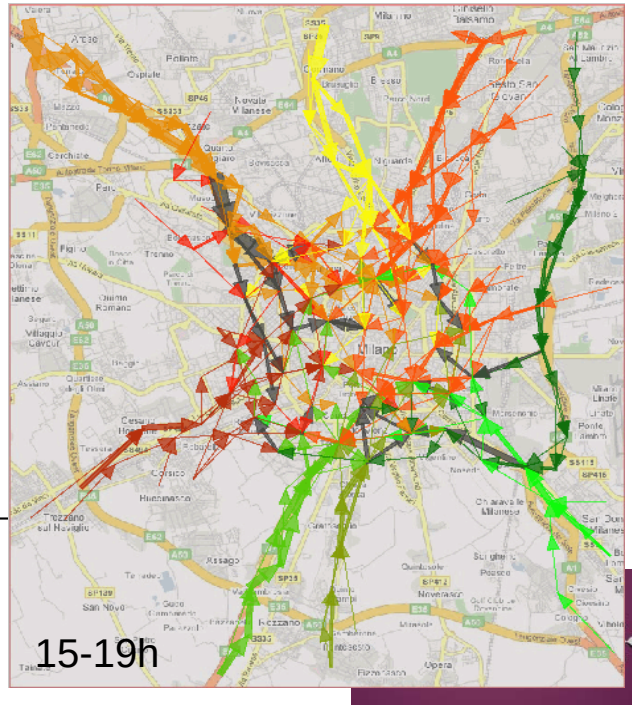






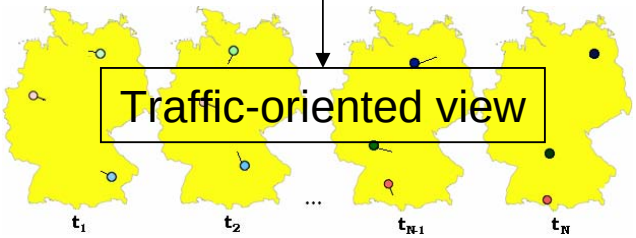


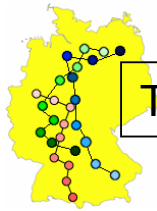




The overall scheme: aggregation

Movement data





Items to aggregate: **points** <position, time>

Spatial aggregation:
points grouped by areas

Temporal aggregation:
points grouped by time intervals

Items to aggregate: **lines** {<position, time>}
- whole trajectories or fragments *(moves)*

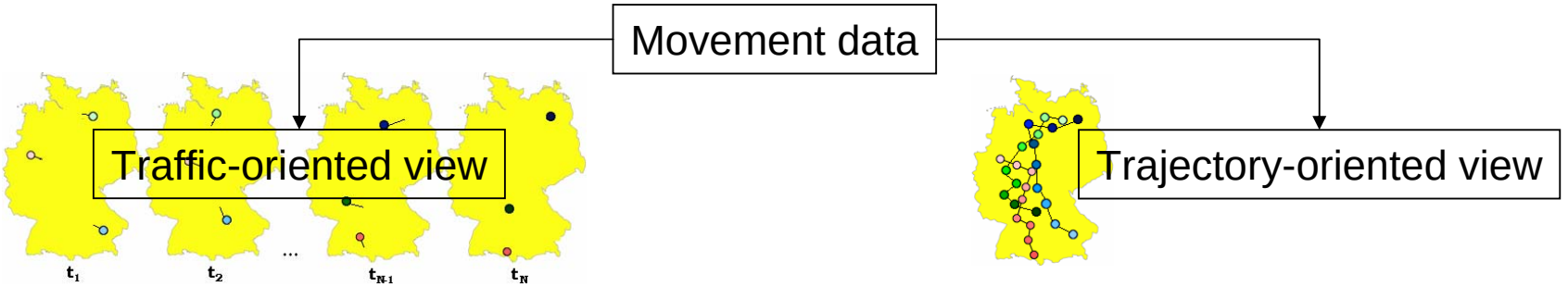
Spatial aggregation:
a) trajectories or moves grouped by areas of their start and end positions
b) trajectories or moves grouped by routes

Temporal aggregation:
moves grouped by intervals of their start and end times

Attributive aggregation:
items grouped by value intervals or by common values

Aggregate characteristics:
N of items, statistics of the movement attributes (speeds etc)

The overall scheme: visualization



- Animated map or map series: development of the traffic situation over time
- Map with diagrams located in areas: local temporal variations of the traffic intensity and characteristics; distribution over the territory
- Time histogram: overall temporal variation of the frequency distribution of the traffic characteristics on the whole territory
- Visual variables: size (in graduated symbols, elements of diagrams, segments of time histogram); colour brightness and/or hue; orientation (in directional diagrams)

- Time-controlled flow map (vectors represent aggregate moves): movement flows; temporal variation of the flows
- Time-controlled origin-destination matrix: statistics of moves between predefined areas
- Interactive filtering of aggregate moves according to their characteristics (e.g. N of trajectories)
- Visual variables: size (thickness of vectors, size of symbols in matrix cells), colour (represents trajectory groups according to the routes)