ViAMoD

Visual Spatiotemporal Pattern Analysis of Movement and Event Data

Prof. Dr. Stefan Wrobel
Dr. Natalia Andrienko
Dr. Gennady Andrienko
NN
http://visual-analytics.info

Prof. Dr. Daniel Keim
Dr. Peter Bak
Slava Kiselevich
http://infovis.uni-konstanz.de/members/keim

DFG SPP Visual Analytics kick-off meeting, Dagstuhl, December 2008

Summary

- Spatiotemporal data are generated in rapidly growing amounts. There is a high demand for scalable analysis methods, which allow a systematic analysis and have a sound theoretical basis.
- Spatiotemporal data, particularly, movement data, involve geographical space, time, and multidimensional attributes and thereby pose significant challenges for the analysis.
- We plan to develop theoretical foundations for the analysis of spatiotemporal data, which account for possible variations of the essential properties of the data. We will thereby identify the generic analysis tasks for different types of movement data and different views of movement.
- The goal of the project is to develop the appropriate analysis methods, which combine visual, interactive, and algorithmic techniques for a scalable analysis.

Motivating Applications

- Health: analyzing and predicting spread of diseases in hospitals (tracking patients) or by migrating birds
- Biology: studying behaviors of animals
- Environment protection and nature preservation: detection of illegal activities
- Social science and history: analyzing individual history, revealing social structures and patterns of interaction
- Business: transportation management, targeting outdoor advertisements, optimizing layout of trade spaces, detecting bottlenecks in logistic systems
- Mobile gaming and education: analyzing involvement of participants and usage of space
- Sport: post-game and online support for team managers, journalists, and general public
- Security and safety: improving layout of public buildings, supporting evacuation from crisis-affected areas, identifying suspicious behaviors or fraud banking transactions

Movement Data: Simple Structure

Movement data is a temporal sequence of position records:

- <time moment, spatial position, {additional attributes}>
  in case of a single moving entity
- <entity identifier, time moment, spatial position, {additional attributes}>
  in case of several moving entities

Movement Data: Simple Structure, Difficult to Analyze

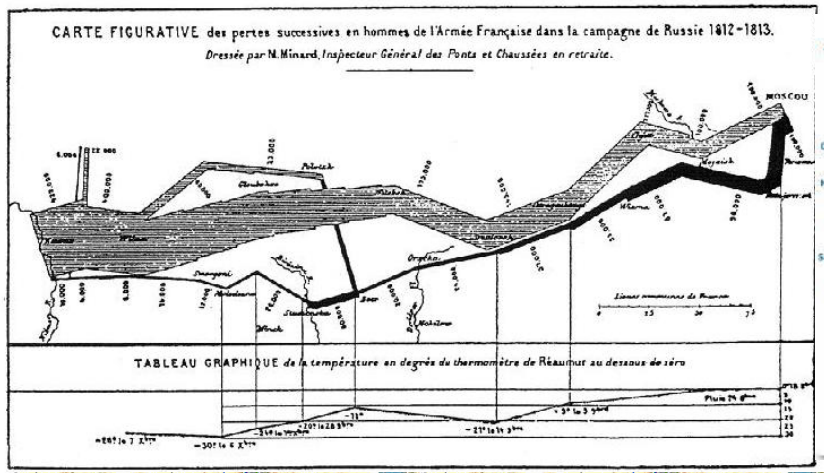
Movement data is a temporal sequence of position records:

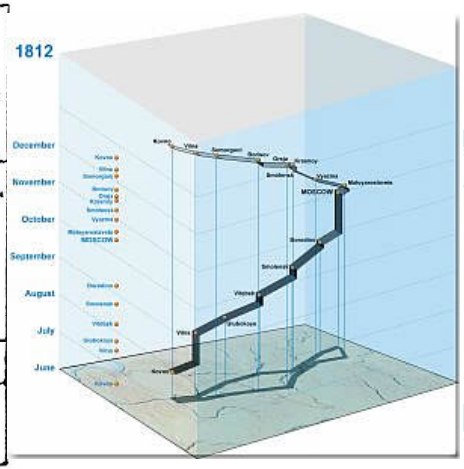
- <time moment, spatial position, {additional attributes}>
  in case of a single moving entity
- <entity identifier, time moment, spatial position, {additional attributes}>
  in case of several moving entities

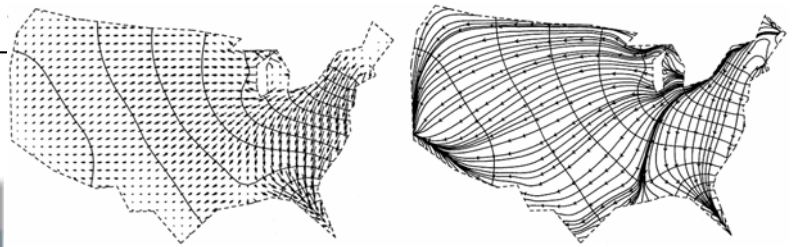
Complexities:

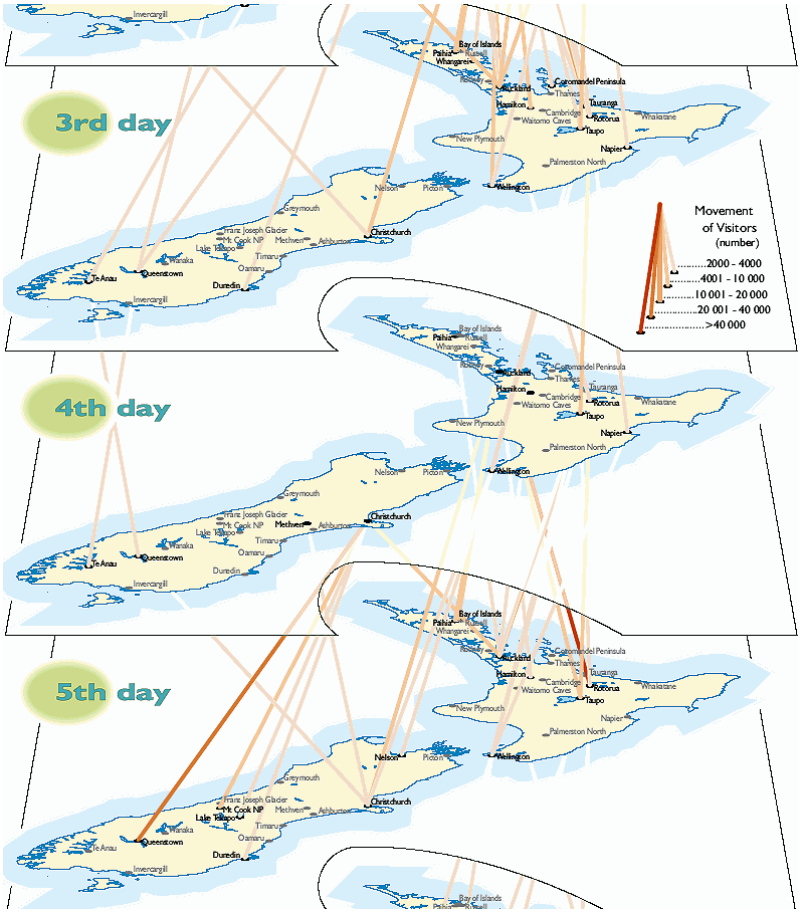
1. Amount (number of moving entities, number of records)
2. Geographic space with its structure and complexity
3. Time, linear and also multiple nested and overlapping cycles
4. Data properties:
   - imprecision (errors in location, time, attributes)
   - irregular sampling (quasi-continuous or event-based)
   - missing data

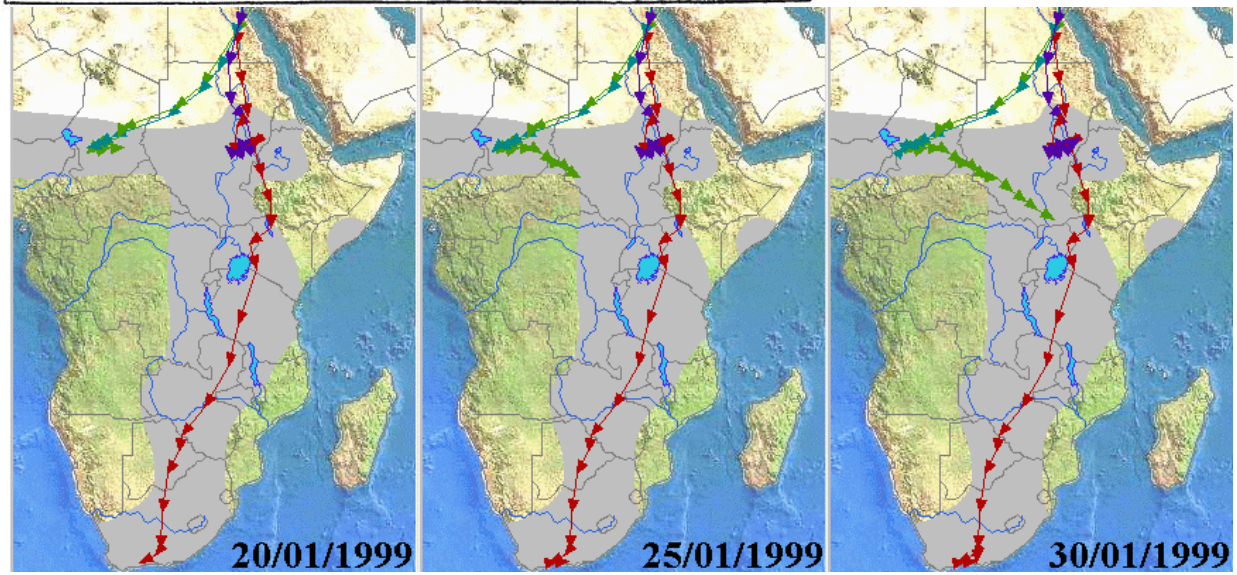
State of the art - visualization









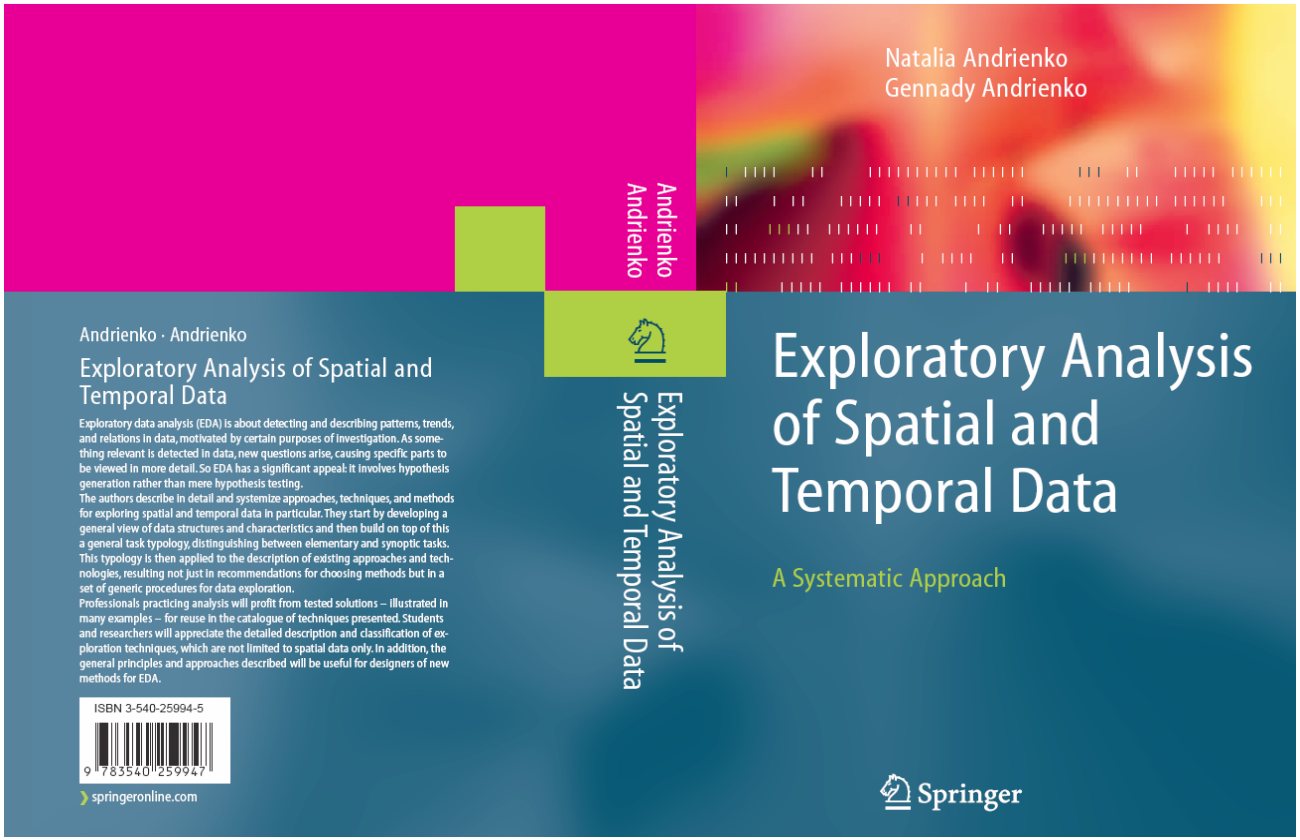


State of the art – data mining

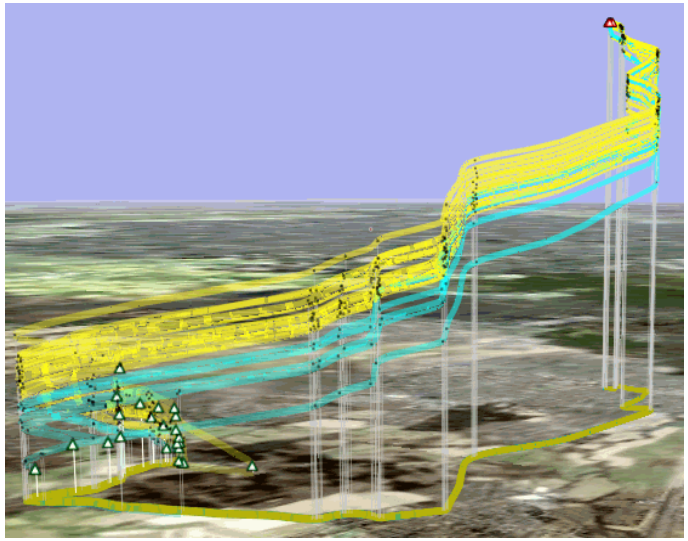
- Spatial data mining: usually feature extraction from spatial data followed by application of regular data mining methods
- Distance functions for trajectories – used for clustering
- Ad hoc methods for specific kinds of patterns:
  - T-patterns (same sequences of visited places with similar transition times),
  - Relative motion patterns (flock, leadership etc.)
- Location prediction by applying statistical models
- Classification of movement trajectories

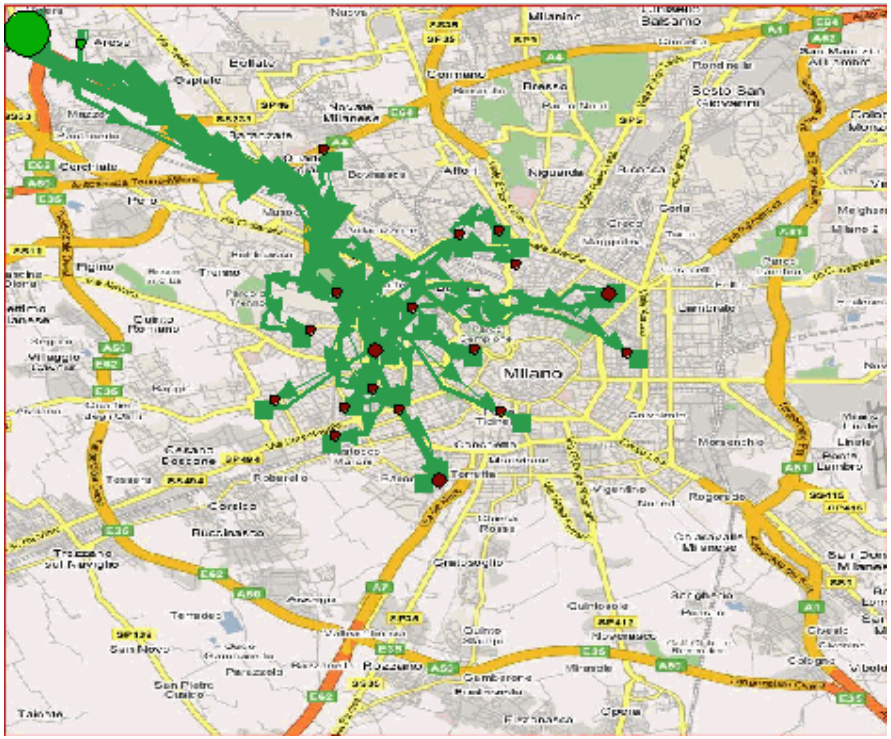
Basic Techniques – Univ. Bonn and Fraunhofer IAIS
Exploratory Analysis of Spatial and Temporal Data

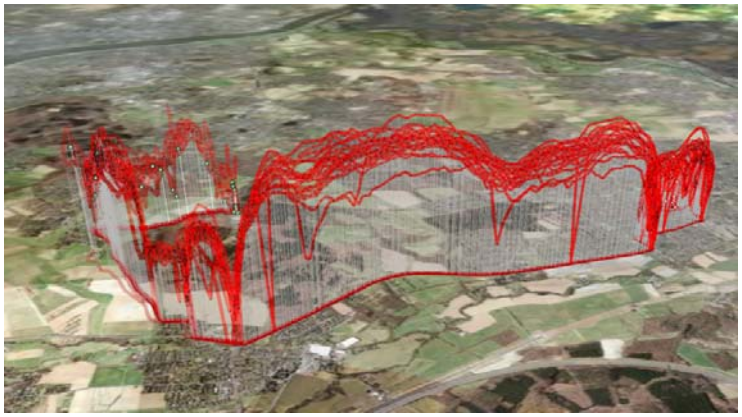
- A system of visualization and interaction techniques supporting exploration of different types of spatial and spatio-temporal data
- A taxonomy of generic tasks in EDA defined on the basis of a formal data model
- A systematic survey of the state of the art in the methods for EDA
  - Visualization and display manipulation
  - Data manipulation
  - Querying
  - Computational analysis
- A system of generic principles and procedures for EDA



Basic Techniques – Univ. Bonn and Fraunhofer IAIS
Analysis of trajectories



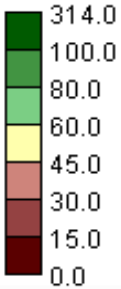


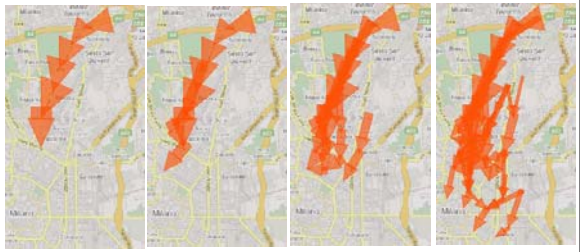


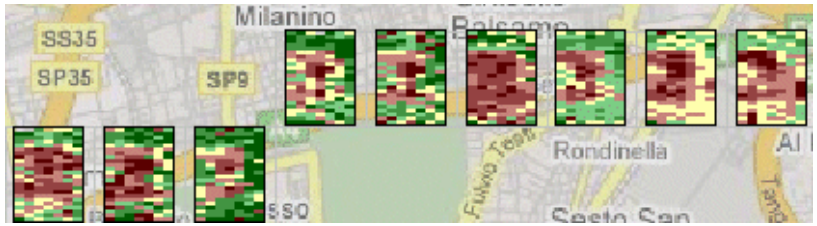
Algorithms, other details:
G.Andrienko, N.Andrienko, S.Wrobel
Visual Analytics Tools for Analysis of Movement Data
ACM SIGKDD Explorations, v.9(2), December 2007

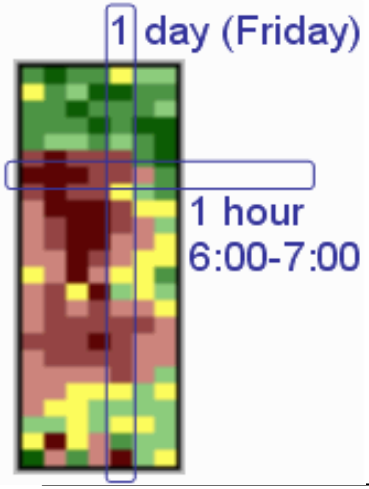
Basic Techniques – Univ. Bonn and Fraunhofer IAIS
Analysis of city traffic

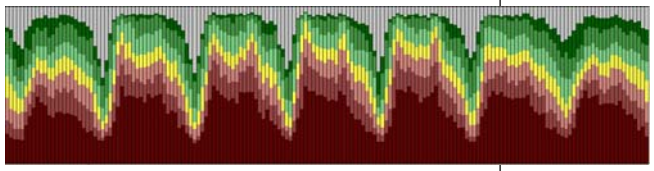
Car data aggregated by grid
Attribute: median speed
Parameters:
Hour: from 0 to 23
Day of week: from 1 to 7

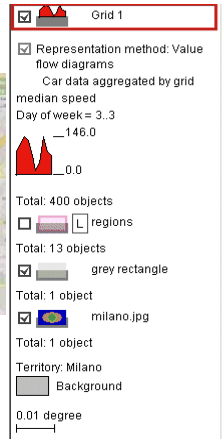


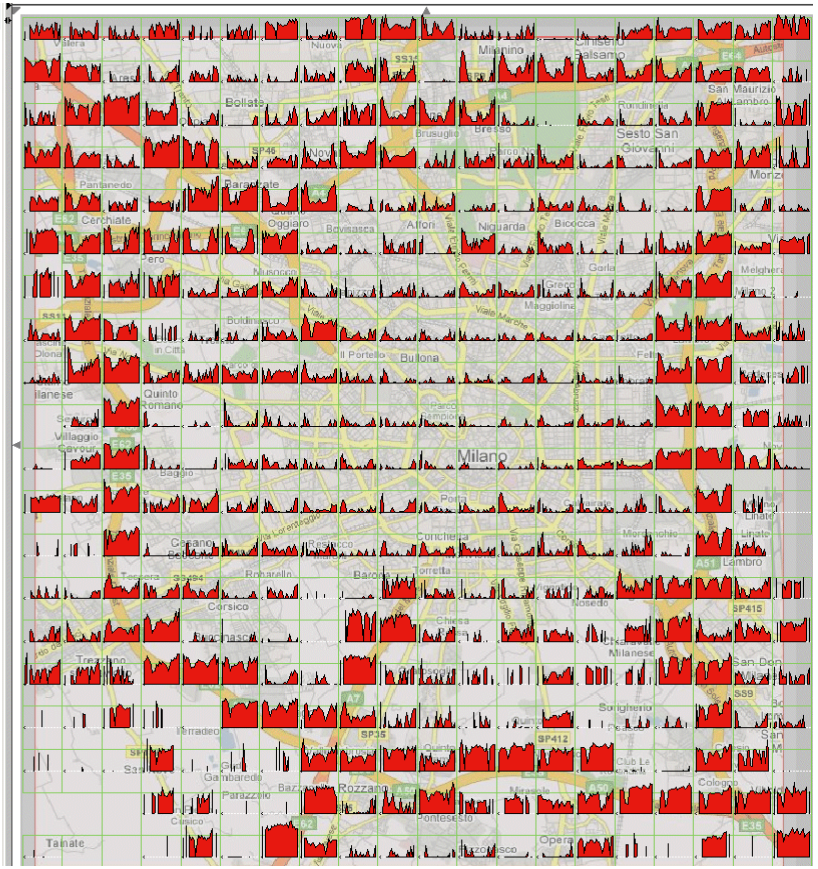






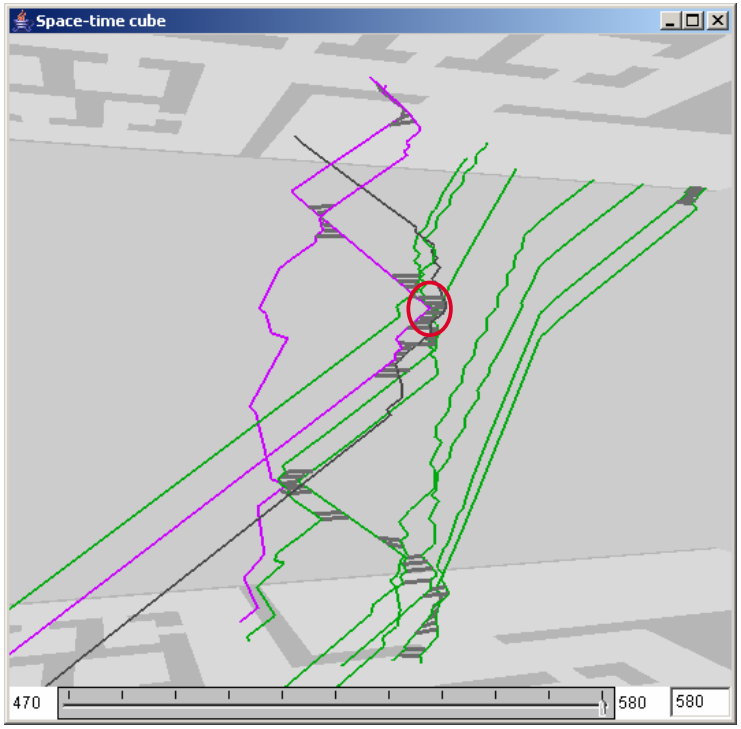


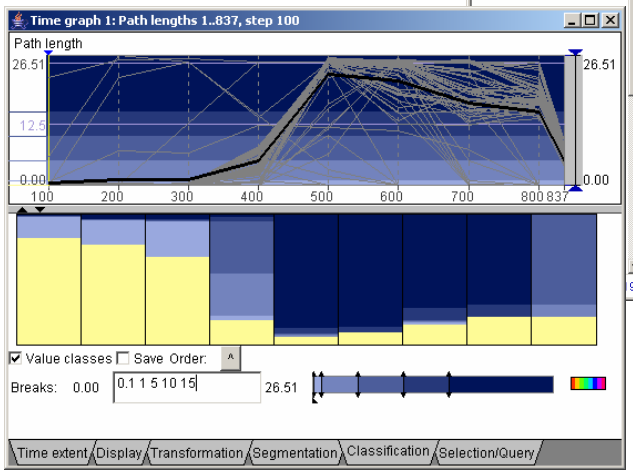


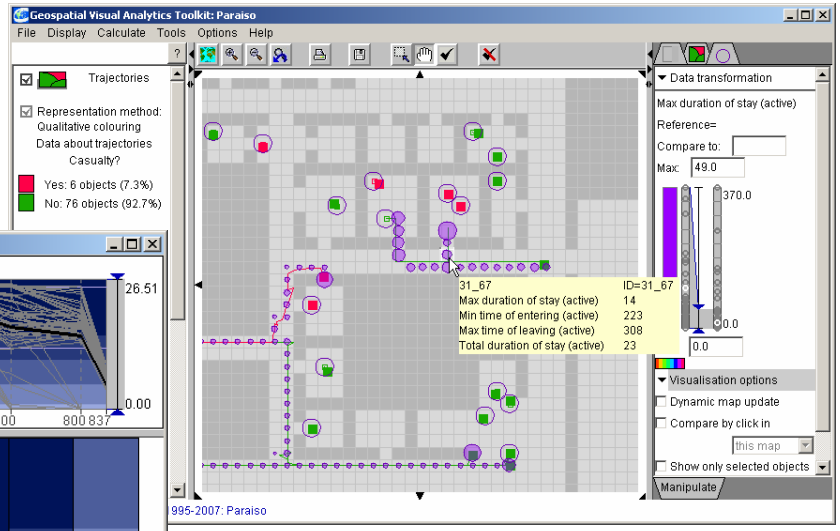


Algorithms, other details:
Gennady Andrienko, Natalia Andrienko
Spatio-temporal aggregation for visual analysis of movements
IEEE Visual Analytics Science and Technology (VAST 2008)
Proceedings, IEEE Computer Society Press, 2008, pp.51-58

Basic Techniques – Univ. Bonn and Fraunhofer IAIS
Analysis of movements by complementary tools





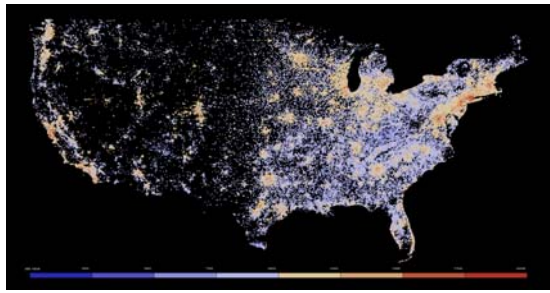


IEEE VAST 2008 CHALLENGE

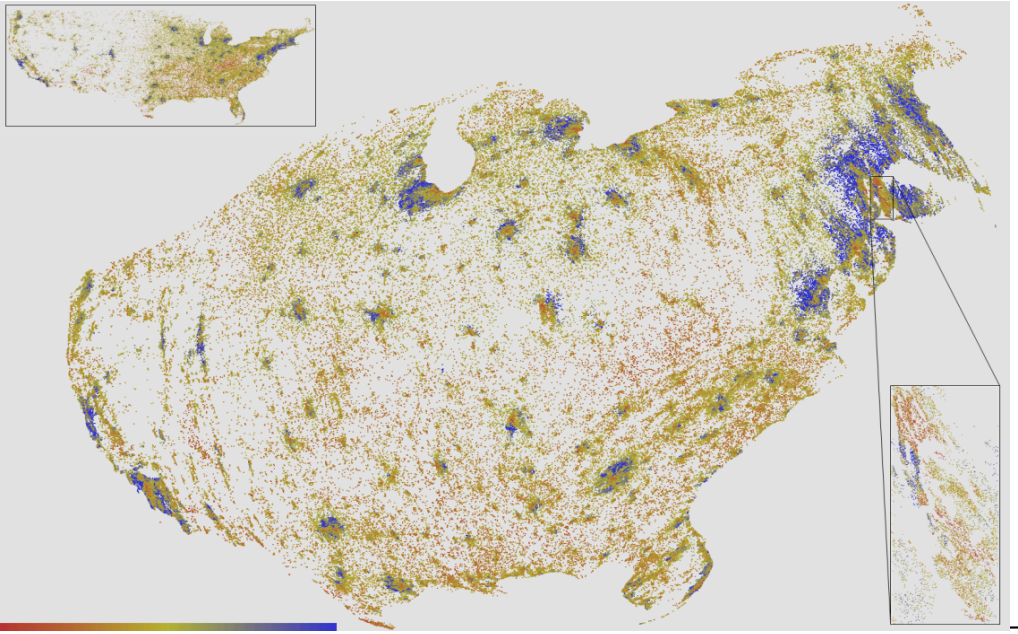
Algorithms, other details:
Natalia Andrienko, Gennady Andrienko
Evacuation Trace Mini Challenge Award: Tool Integration. Analysis of Movements with Geospatial Visual Analytics Toolkit
IEEE Visual Analytics Science and Technology (VAST 2008)
Proceedings, IEEE Computer Society Press, 2008, pp.205-206

Basic Techniques: Univ. Konstanz – Spatial Data Analysis

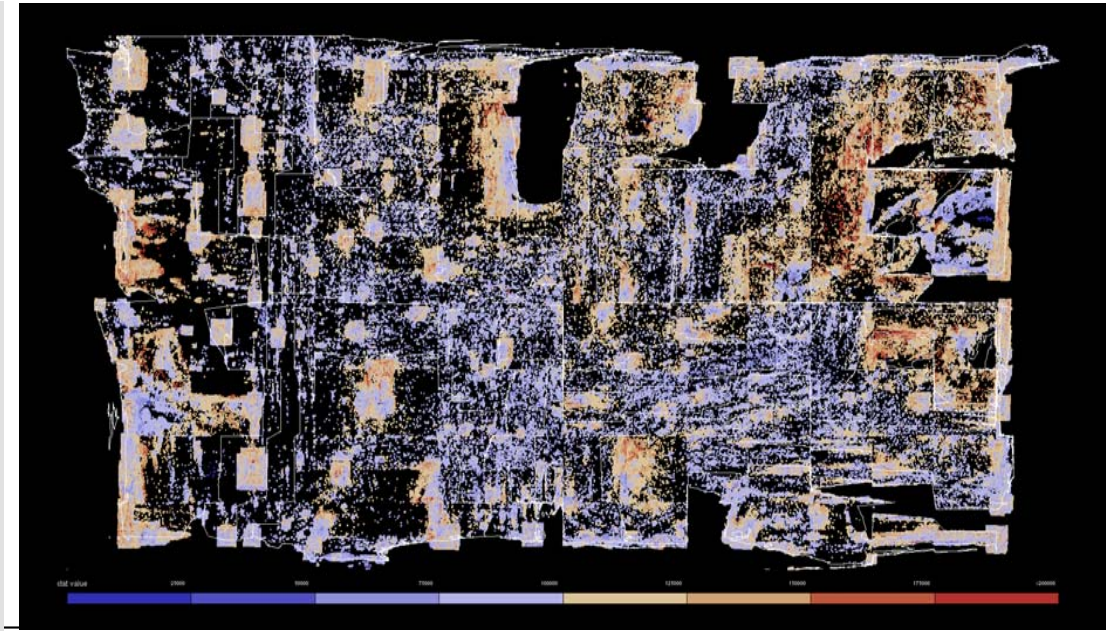
- ✓ Pixel based geographic data-representations.
- ➢ Making information in large datasets visible.



Density Equalizing Distortions

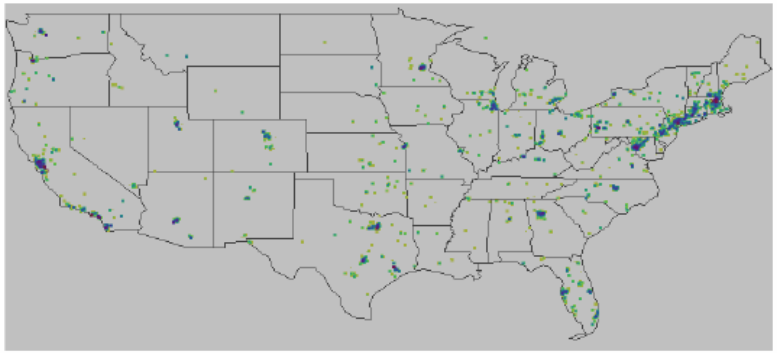


Pixel placement and cartograms

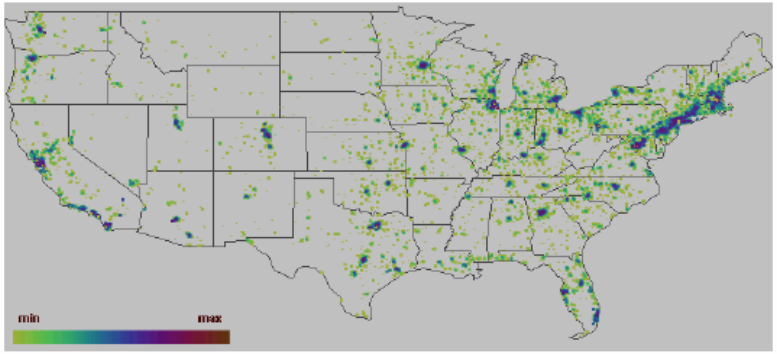


Basic Techniques: Univ. Konstanz – Geo-related Temporal Data Analysis

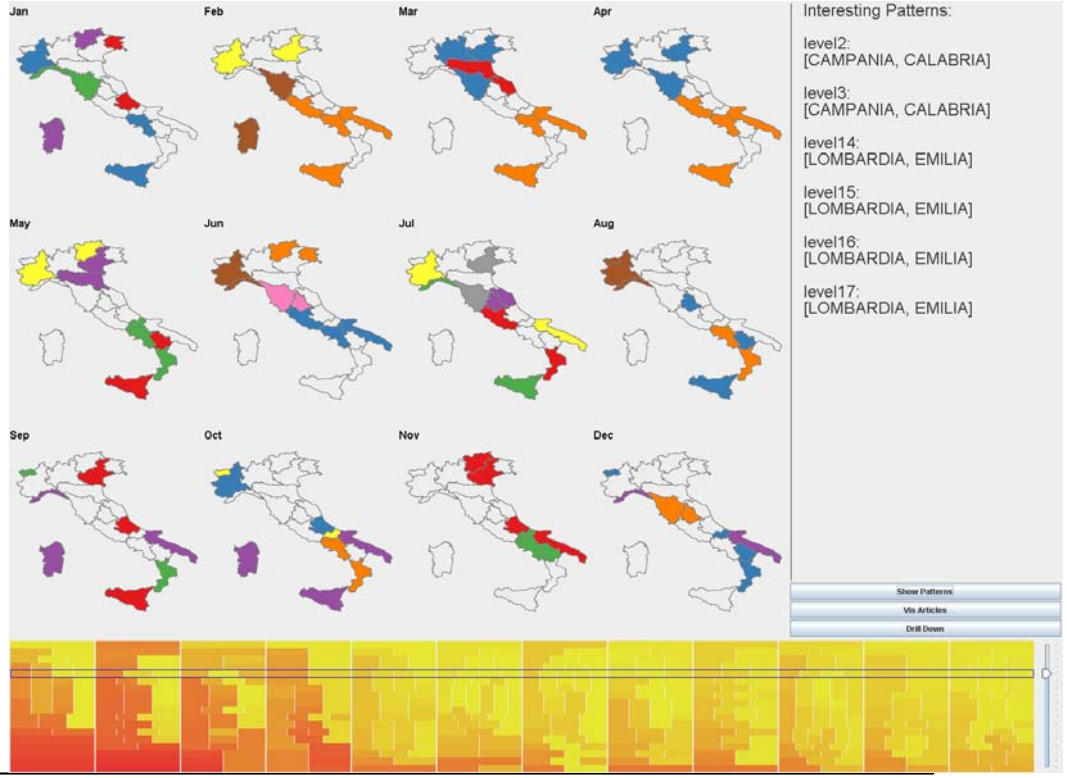
- ✓ Investigating changes over time.
- ➢ Using small multiples at different observations in time.



Industry development ('89-'03)

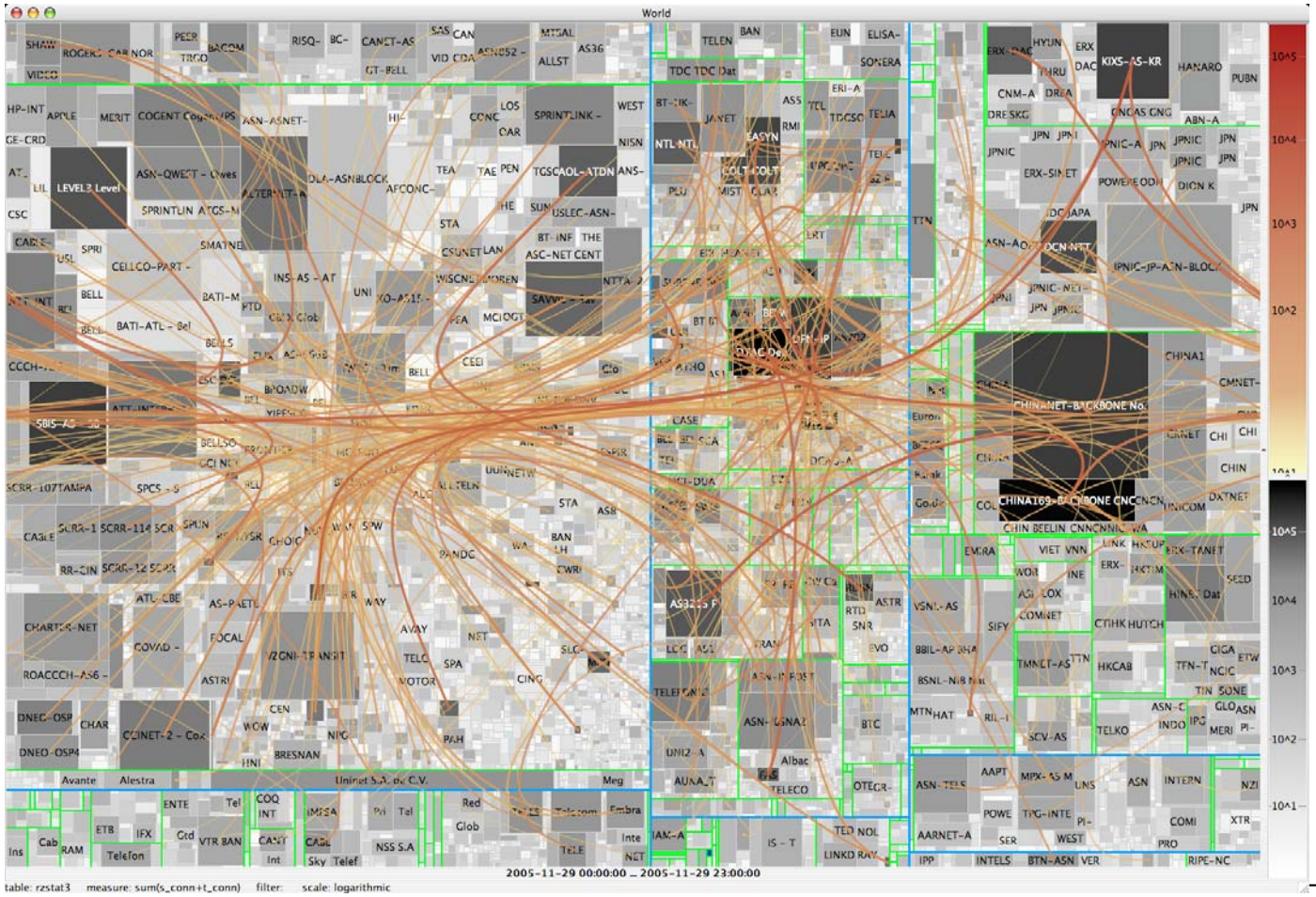


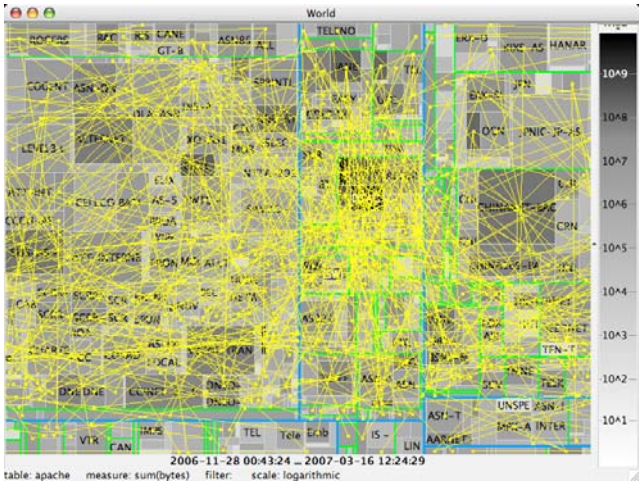
Evolution of geo-spatial patterns.



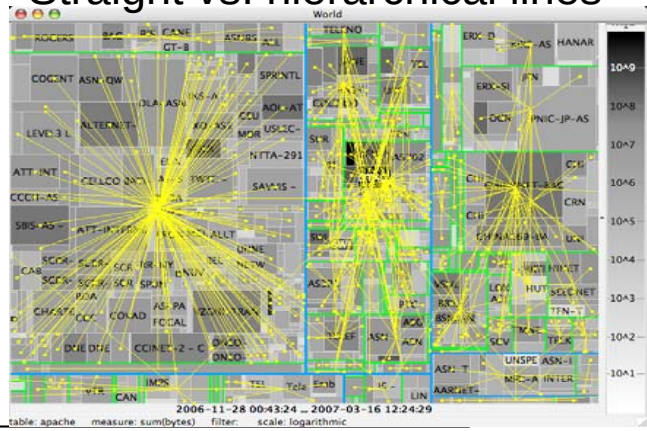
Basic Techniques: Univ. Konstanz – Traffic Analysis

✓ Investigating internet traffic by using Edge Bundles [4]





Straight vs. hierarchical lines



What is missing and why this research is needed

- Lack of an appropriate theoretical basis
  - a great part of the research goes along the way of importing and adapting existing methods for the analysis of geographical data, time-series data, item sequences;
  - the other part is concerned with designing ad hoc methods for specific data and applications
- Typical assumption: data represent continuous space-time paths, interpolation is used for obtaining intermediate positions
- Little has been done on joint analysis of movement data and multidimensional attributes of the moving entities and of the environment
- Scalability
  - In most cases analysis is done in RAM

Our goals

The project aims at advancing the state of the art by

1. developing theoretical foundations for the analysis of movement data;
2. addressing various types of movement data;
3. developing methods for joint analysis of movement data and multidimensional attributes, both static and dynamic;
4. finding approaches to overcome the scalability limitations.

Object and level of analysis

Object of analysis:

- Movement of a single entity
- Movement of several or multiple entities
  - Unrelated entities
  - Related entities

Level of analysis:

1. Motion, i.e. the process of changing the spatial position
2. Trips, i.e. travelling from one place to another
3. Activities of the moving entities

- *Differ in the amount of semantics involved in the analysis*

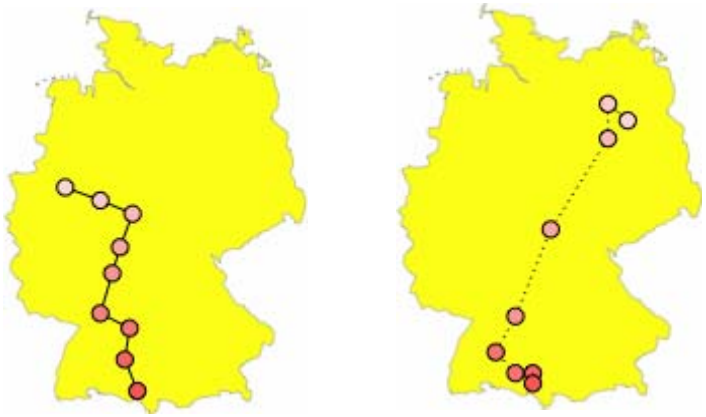
Structure of movement data

Movement data is a temporal sequence of position records:

- <entity identifier, time moment, spatial position, {additional attributes}>

Typical procedures for collecting movement data:

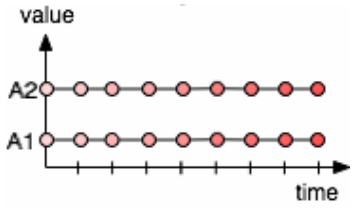
- time-based
- change-based
- location-based
- event-based



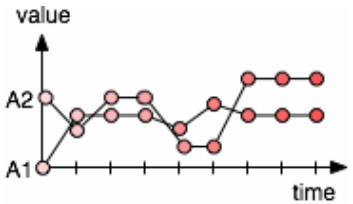
quasi-continuous Vs. discontinuous trajectories

Multidimensional attributes related to space and time

- Static attributes: are a function of the entity; do not change with the entity's movement.
  Static attributes may (partly) explain why and how an entity moves in space and predict the entities' future movement.



- Dynamic attributes are a function of the entity's movement; they change over time as a consequence of the movement.
  Analysis questions might then be concerned with correlations between space, time, and the multidimensional attributes of the entities.



Trajectory Analysis Vs. Traffic Analysis

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

A general formal model of movement of multiple entities: a two-argument function from entity and time to space

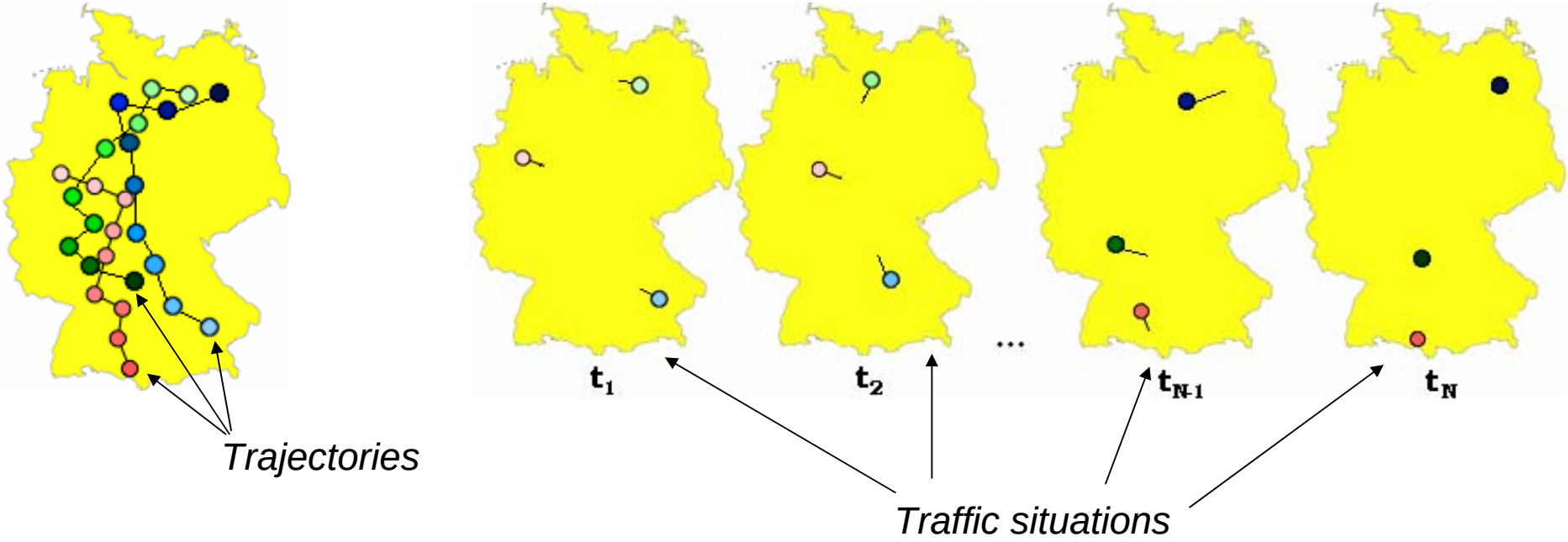
Two possible decompositions into single-argument functions

Trajectory-oriented view

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

Traffic-oriented view

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



Research questions

1. Q1: How do the possible analysis tasks depend on the characteristics of movement data (quasi-continuous or event-based, static or dynamic attributes)?
2. Q2: How do the tasks of trajectory analysis differ from those of traffic analysis?
3. Q3: What are the implications of the different characteristics of movement data for the choice of suitable analysis methods?
4. Q4: What visual analytics methods and procedures are appropriate for supporting trajectory analysis and traffic analysis with the different varieties of movement data, in terms of continuity and temporal dependency of the multidimensional attributes?

Work Packages

Attributes / Data	Static	Dynamic
Quasi-continuous	Work Package 2 (e.g., Car Movements in a City)	Work Package 4 (e.g., Physiological Measures in Sport)
Event-based	Work Package 3 (e.g., Package Tracking)	Work Package 5 (e.g., Credit Card Transactions)

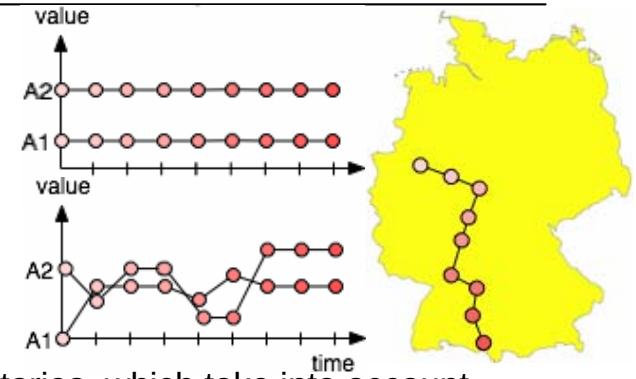
Work Packages

1 (Theoretical foundations)
+
6 (evaluation)

WP1: Theoretical Foundations

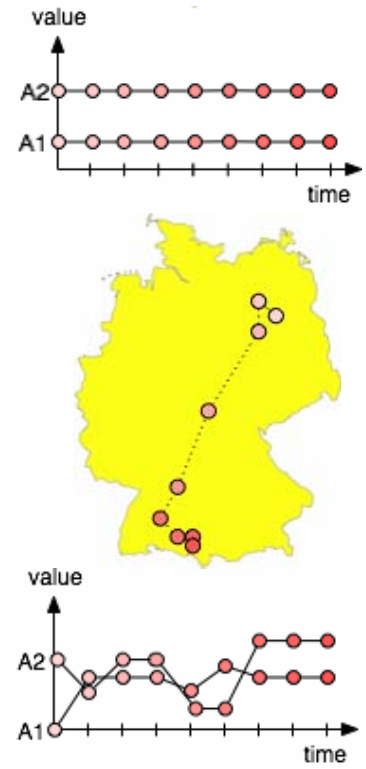
- investigate in depth the distinctions between quasi-continuous and event-based movement data and draw implications concerning the possible methods of data processing and analysis;
- identify the generic analysis tasks for different varieties of movement data in terms of continuity (quasi-continuous or event-based) and temporal dependence of the attributes (static or dynamic); describe the forms of the results of these tasks;
- conceptually define the methods required for accomplishing the tasks that have been identified, i.e. specify the essential qualities of the methods and their inputs and outputs;
- in case of finding common methods for different types of movement data, describe the distinctions in the application of these methods to the different data types;
- determine possible and/or necessary interactions between the methods;
- define analytic procedures where several methods are complementary for accomplishing a task.

WP2,4: Analysis of Quasi-Continuous Data



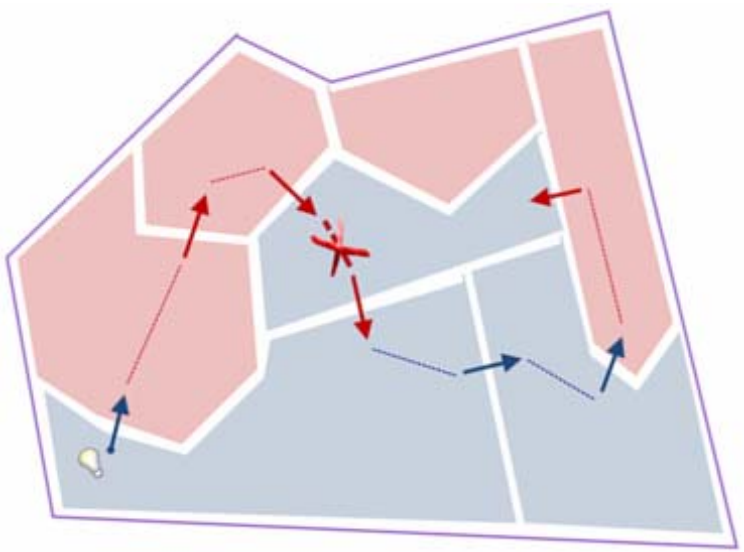
- Trajectory-oriented view:
  - interactive tools to divide the entire trajectories of the entities into suitable portions according to various criteria;
  - a library of distance functions for assessing similarities between trajectories, which take into account various aspects of trajectories: spatial positions, directions, routes, times, speeds, stops, and non-spatial dynamic attributes;
  - clustering algorithms capable of using these distance functions (different methods will be required for space-based clustering and for attribute-based clustering);
  - computational methods for deriving summarized profiles of groups of similar trajectories;
  - visualization methods for presenting summarized trajectories and separate trajectories (e.g. atypical) in geographical space, in time, and in multi-attribute space;
  - interactive tools for display linking (e.g. through brushing), filtering, and selection.
- Traffic-oriented view:
  - appropriate aggregation procedures
  - how to detect proper representatives of an aggregate
  - how to visualize aggregates
  - combining aggregation with clustering and other methods

WP3,5: Analysis of Event-based Data

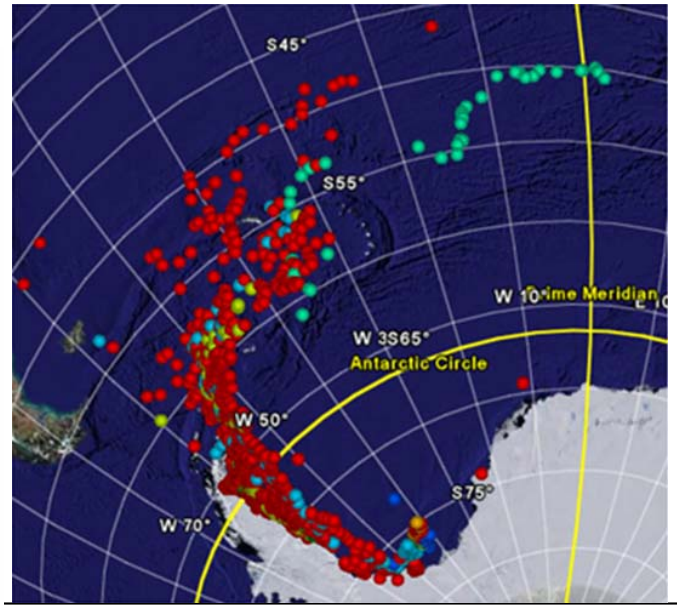


- ✓ Data: Types of data & conceptual modeling
- ✓ Methods: Automatic vs. visualization guided
- ✓ Tasks: Correlations & predictions

Coverage by
Cell-phone providers



Iceberg movements



Collaborative and community activities

- VisMaster, EU-funded Coordination Action
  - D.Keim is scientific coordinator of the project
  - G.Andrienko is coordinator of WG on spatio-temporal aspects
- Commission on GeoVisualization of ICA
  - G.Andrienko is chairing the commission
- Special issue of Cartography and GIScience on Geospatial Visual Analytics, 2009
- VA is selected to be one of the 12 Fraunhofer Frontline Themes
  - S.Wrobel is coordinating this theme