ARTS1422 Data Visualization

Lecture 15

Visual Analysis of Trajectories
&
More than Design: Evaluation

Quan Li

Spring 2024
2024. 04.25

Outline

- 1 Introduction
- 2 Trajectory Data
- 3 Trajectory Pattern Analysis
- 4 Trajectory Visualization Techniques
- 5 Applications
- 6 Conclusions

1 Introduction

- 1.1 Motivation
- 1.2 Research Problems
- 1.3 Challenges
- 1.4 Applications Overview

1.1 Motivation

- Wireless Exploration

1.2 Research Problems

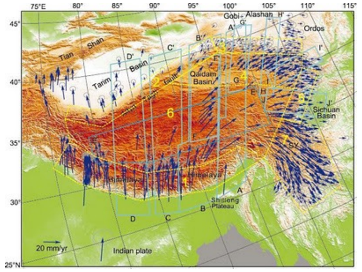
- Wireless technologies are improving rapidly and generating tremendous amounts of trajectory data
- It is important to provide visualization tools to help analyzing the trajectory data generated by mobile devices
  - Provide intuitive views
  - Rich user interactions

1.3 Challenges

- Uncertainty
  - Sampling rate could be inconstant
    - Data can be sparse
- Noise
  - Erroneous points (e.g., a point in a river)
- Scalability
  - Visual clutter problems

1.4 Applications Overview

- Trajectory data visualization has many important, real-world applications driven by real need
  - Disaster Forecasting
    - Earthquake prediction
  - Traffic Analysis and Management (Traffic control)
    - To alert people about traffic jams, accidents
    - Identify/predict low traffic regions in a city
  - Analysis of Movements of People
  - Location-based Service

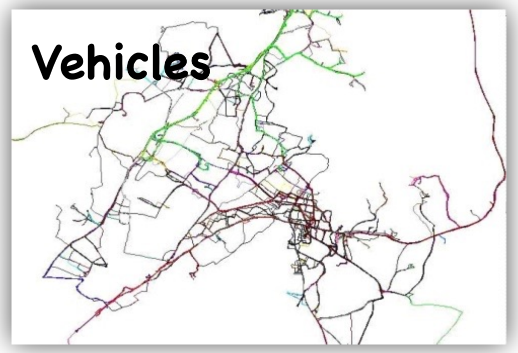


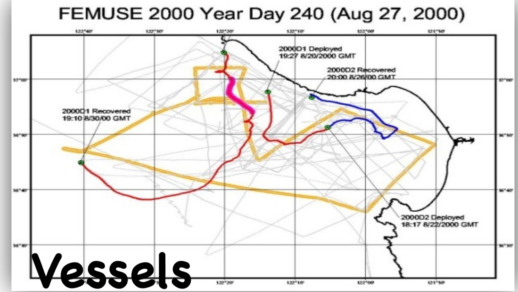
2 Trajectory Data

- 2.1 Data Collection
- 2.2 Data Characteristics
- 2.3 Data Preprocessing
- 2.4 Data Calibration

2.1 Data Collection

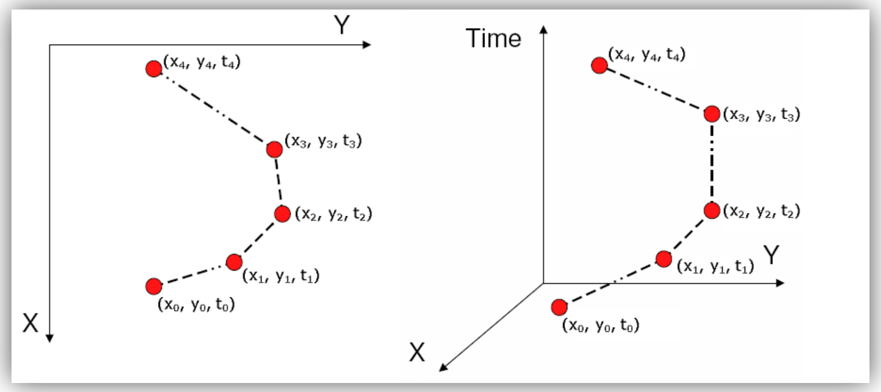
- Trajectory data of moving objects can be collected from various application domains
  - Transportation management
  - Urban planning
  - Tourism
  - Location-based services
  - Flight safety
  - Marine safety





2.2 Data Characteristic (1/2)

- *"A trajectory is the path made by the moving entity through the space it moves."*
  - N. Andrienko et al., 2008
- A trajectory can be defined as a function from the temporal :



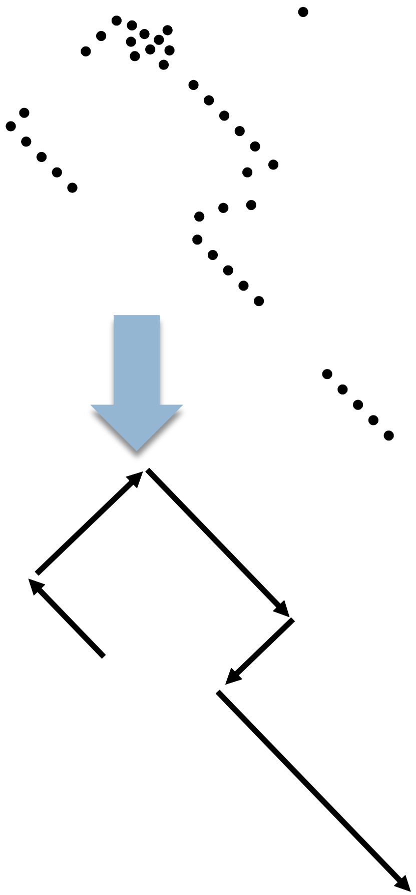
2.2 Data Characteristic (2/2)

Characteristics
- Formal Definition
- Basic Characteristics

- Geometric Shape
- Length (distance traveled)
- Duration (time traveled)
- Speed
  - Mean, median, and maximal speed
  - Periods of constant speed, acceleration, deceleration
- Direction:
  - Periods of straight, curvilinear and circular movement;
  - Major turns (‘turning points’) in: time, position, angle, initial and final directions, and speed in the moment of the turn;

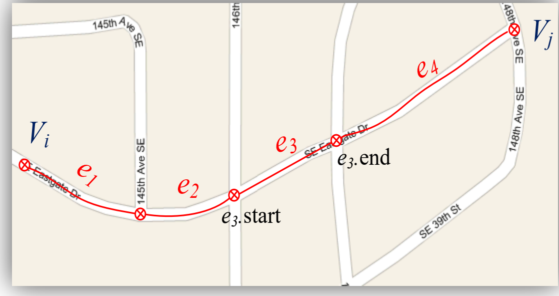
2.3 Data Preprocessing

- **Sampling**
  - Typically represented as a set of localization points of a tracked device
- **Interpolation**
  - Finding a suitable curve to connect the sample points



2.4 Data Calibration

- Map Matching
  - This algorithm utilizes positioning data and spatial road network data to identify the correct link on which a vehicle is travelling and determine the location of a vehicle on that link.
    - Quddus et al., 2007



3 Trajectory Pattern Analysis

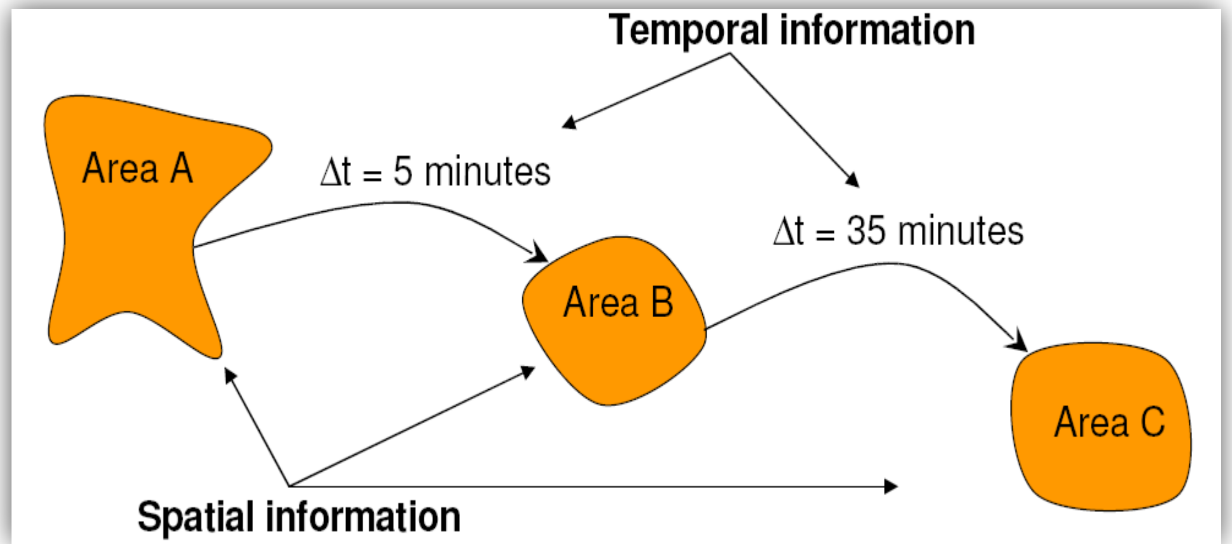
- 3.1 Pattern Mining
- 3.2 Clustering
- 3.3 Classification
- 3.4 Outlier Detection

3.1Pattern Mining (1/3)

上海科技大学 ShanghaiTech University

- Pattern Mining
  - Frequent Pattern
  - Periodic Pattern
  - Trajectory Join

- Trajectory Pattern
  - A trajectory pattern should describe the movements of objects in both space and time



(Giannotti et al., 2007)

3.1Pattern Mining (2/3)

- Pattern Mining
  - Frequent Pattern
  - Periodic Pattern
  - Trajectory Join

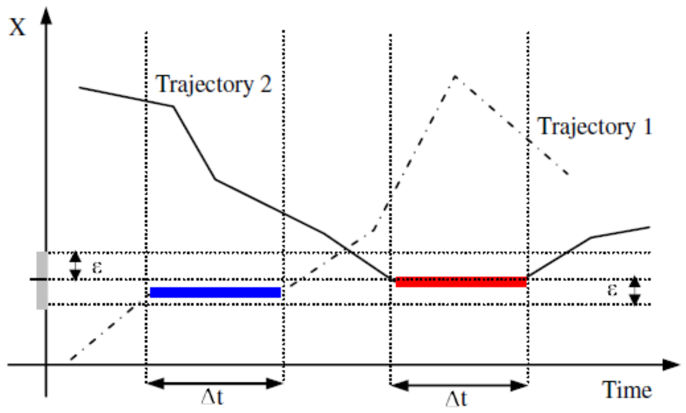
- In many applications, objects follow the same routes (approximately) over regular time intervals

(Mamoulis et al., 2004)

3.1Pattern Mining (3/3)

Pattern Mining
Frequent Pattern
Periodic Pattern
Trajectory Join

- To identify all pairs of similar trajectories between two datasets
  - e.g., identify the pairs of trucks that were constantly within a 1 mile of each other ***this morning***

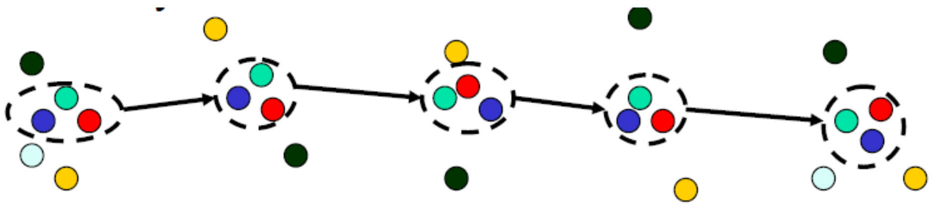


(Bakalov et al., 2005)

Definition: Two trajectories match if there exist time intervals of the same length δt such that the distance between the locations of the two trajectories during these intervals is no more than the spatial threshold ε

3.2 Clustering (1/3)

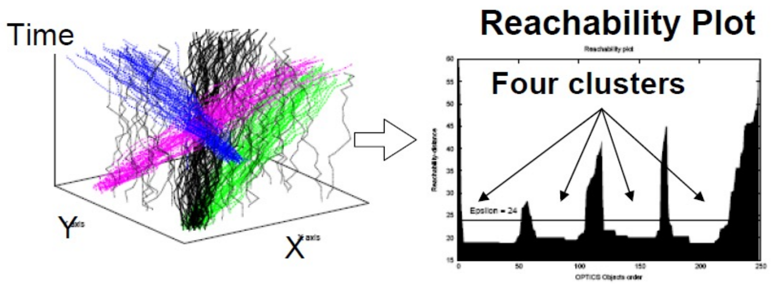
- A *moving cluster* is a set of objects that move close to each other over a long time interval
- Formal Definition [Kalnis et al. 05]:
  - A *moving cluster* is a sequence of (snapshot) clusters c_1, c_2, …, c_k such that for each timestamp i ($1 \le i < k$), $|c_i \cap c_{i+1}| / |c_i \cup c_{i+1}| \ge \theta$ ($0 < \theta \le 1$)



3.2 Clustering (2/3)

- Clustering
  - Moving Object Clustering
  - Density-Based Trajectory Clustering
  - Partition-and-Group Framework

- Define the distance between *whole* trajectories
  - A trajectory is represented as **a sequence of location and timestamp**
  - The distance between trajectories is the average distance between objects for every timestamp
- OPTICS algorithm is used for trajectories

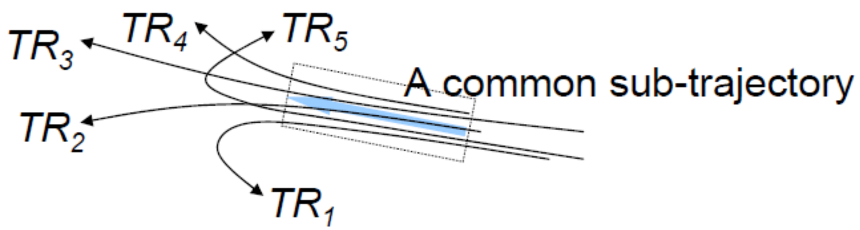


(Nanni & Pedreschi, 2006)

3.2 Clustering (3/3)

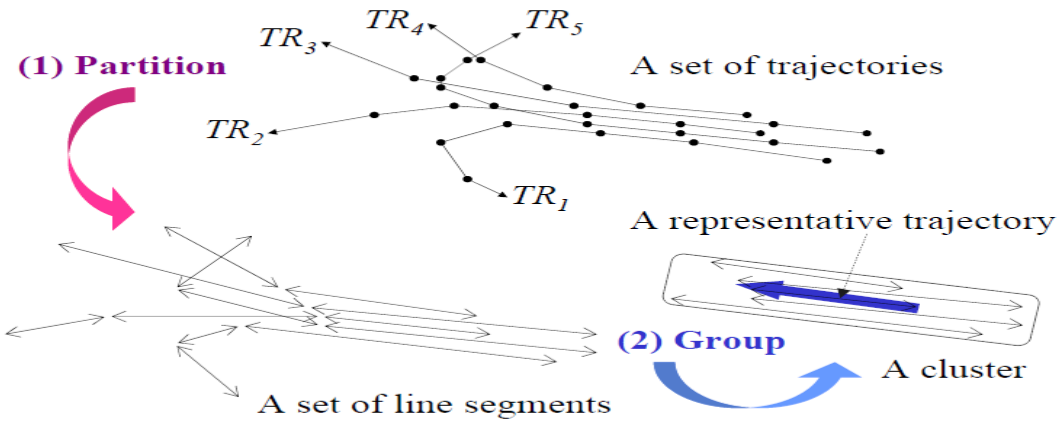
Clustering
Moving Object Clustering
Density-Based Trajectory Clustering
Partition-and-Group Framework

- The ***partition-and-group framework*** is proposed to discover common ***sub***-trajectories

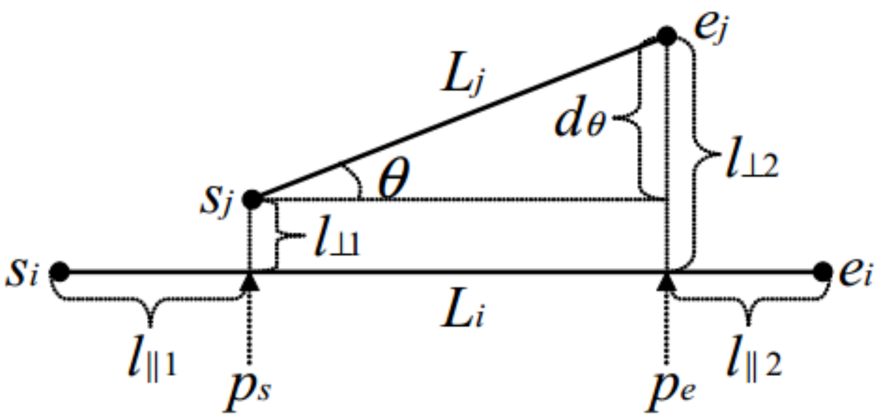


(Lee et al., 2007)

- Two phases: ***partitioning*** and ***grouping***



Distance functions



$$d_\perp = \frac{l_{\perp 1}^2 + l_{\perp 2}^2}{l_{\perp 1} + l_{\perp 2}}$$

$$d_\parallel = \mathrm{MIN}(l_{\parallel 1}, l_{\parallel 2})$$

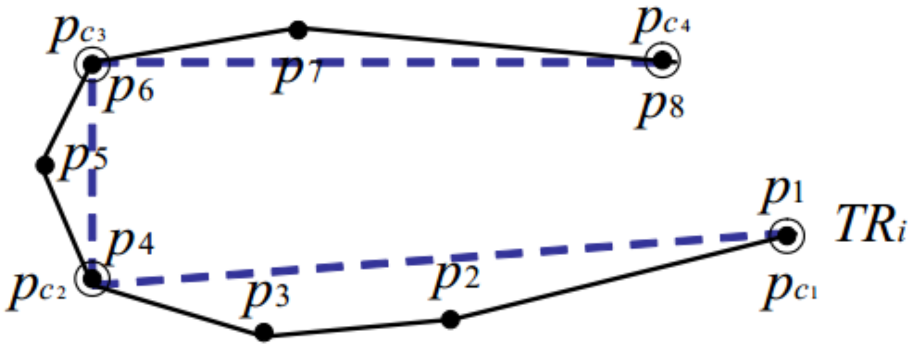
$$d_\theta = \|L_j\| \times \sin(\theta)$$

$$p_s = s_i + u_1 \cdot \overrightarrow{s_i e_i}, \quad p_e = s_i + u_2 \cdot \overrightarrow{s_i e_i},$$
$$\text{where} \quad u_1 = \frac{\overrightarrow{s_i s_j} \cdot \overrightarrow{s_i e_i}}{\|\overrightarrow{s_i e_i}\|^2}, \quad u_2 = \frac{\overrightarrow{s_i e_j} \cdot \overrightarrow{s_i e_i}}{\|\overrightarrow{s_i e_i}\|^2} \tag{4}$$

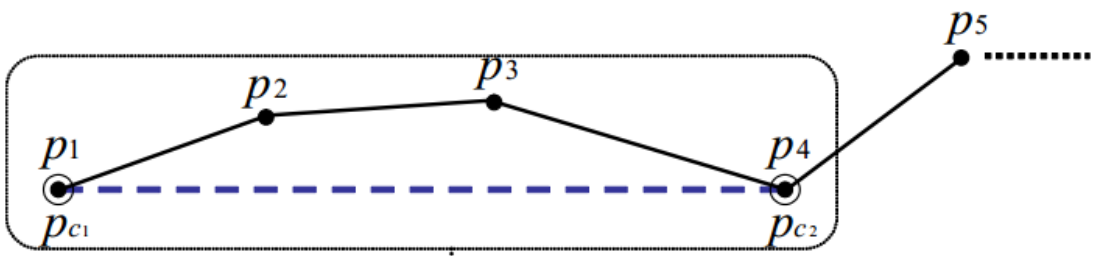
$$cos(\theta) = \frac{\overrightarrow{s_i e_i} \cdot \overrightarrow{s_j e_j}}{\|\overrightarrow{s_i e_i}\| \|\overrightarrow{s_j e_j}\|} \tag{5}$$

$$dist(L_i, L_j) = w_\perp \cdot d_\perp(L_i, L_j) + w_\parallel \cdot d_\parallel(L_i, L_j) + w_\theta \cdot d_\theta(L_i, L_j)$$

Partitioning

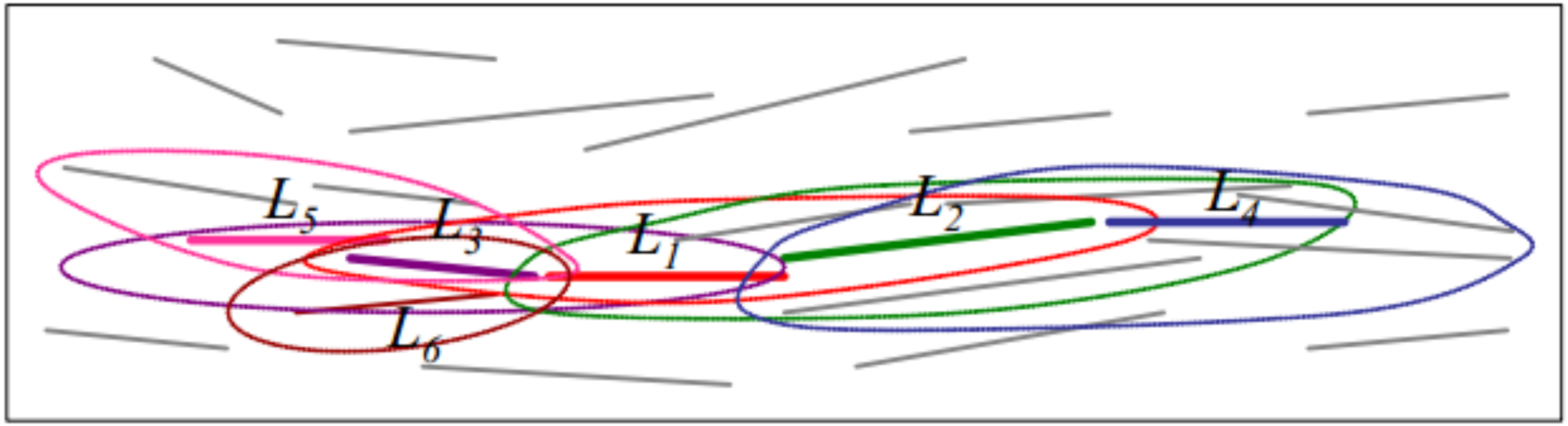


- ➢ Desirable Properties
  - – Preciseness
  - – Conciseness
- ➢ Using the MDL Principle



$$
\begin{aligned}
L(H) &= \log_2(len(p_1p_4)) \\
L(D|H) &= \log_2(d_\perp(p_1p_4, p_1p_2) + d_\perp(p_1p_4, p_2p_3) + d_\perp(p_1p_4, p_3p_4)) + \\
&\quad \log_2(d_\theta(p_1p_4, p_1p_2) + d_\theta(p_1p_4, p_2p_3) + d_\theta(p_1p_4, p_3p_4))
\end{aligned}
$$

Grouping: DBSCAN



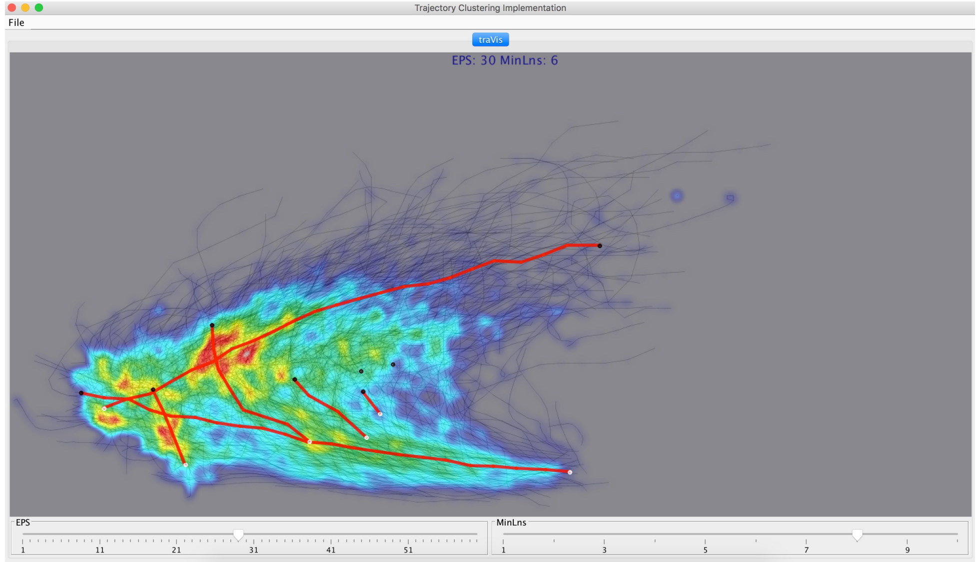
$$L_6 \quad L_5 \leftrightarrow L_3 \leftrightarrow L_1 \leftrightarrow L_2 \leftrightarrow L_4$$

Figure 10: Density-reachability and density-connectivity.

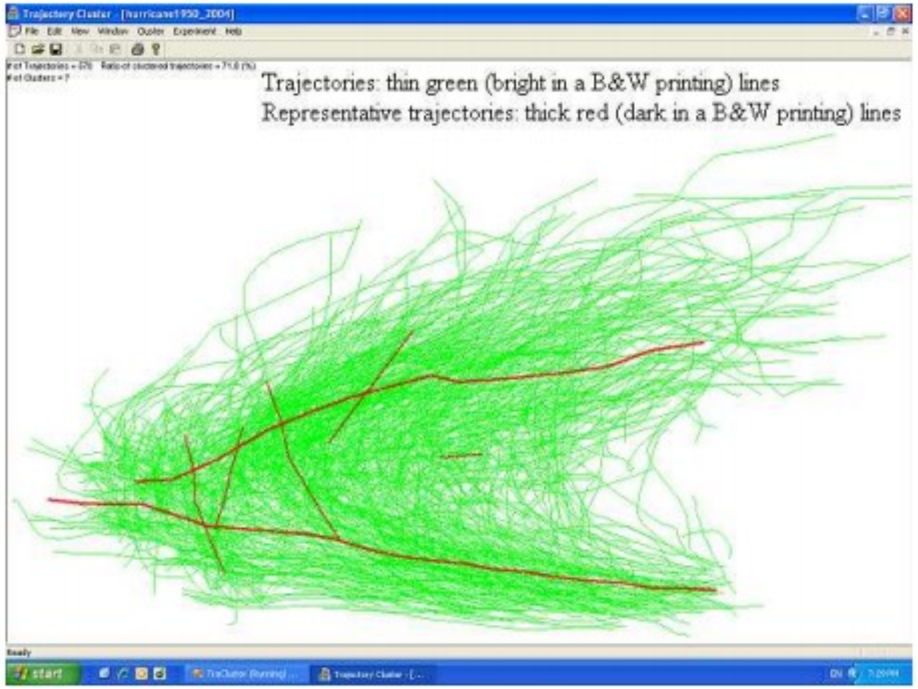
Experiments

- Data set used in the paper
  - Hurricane data (608 trajectories, 17736 points)
  - Elk1993 data (33 trajectories, 47199 points)
  - Deer1995 data (32 trajectories, 20064 points)
- Data set collected by ourselves
  - Gesture trajectory (641 trajectories, 47107 points)

- Hurricane Track Data ($\varepsilon = 30, MinLns = 6$)

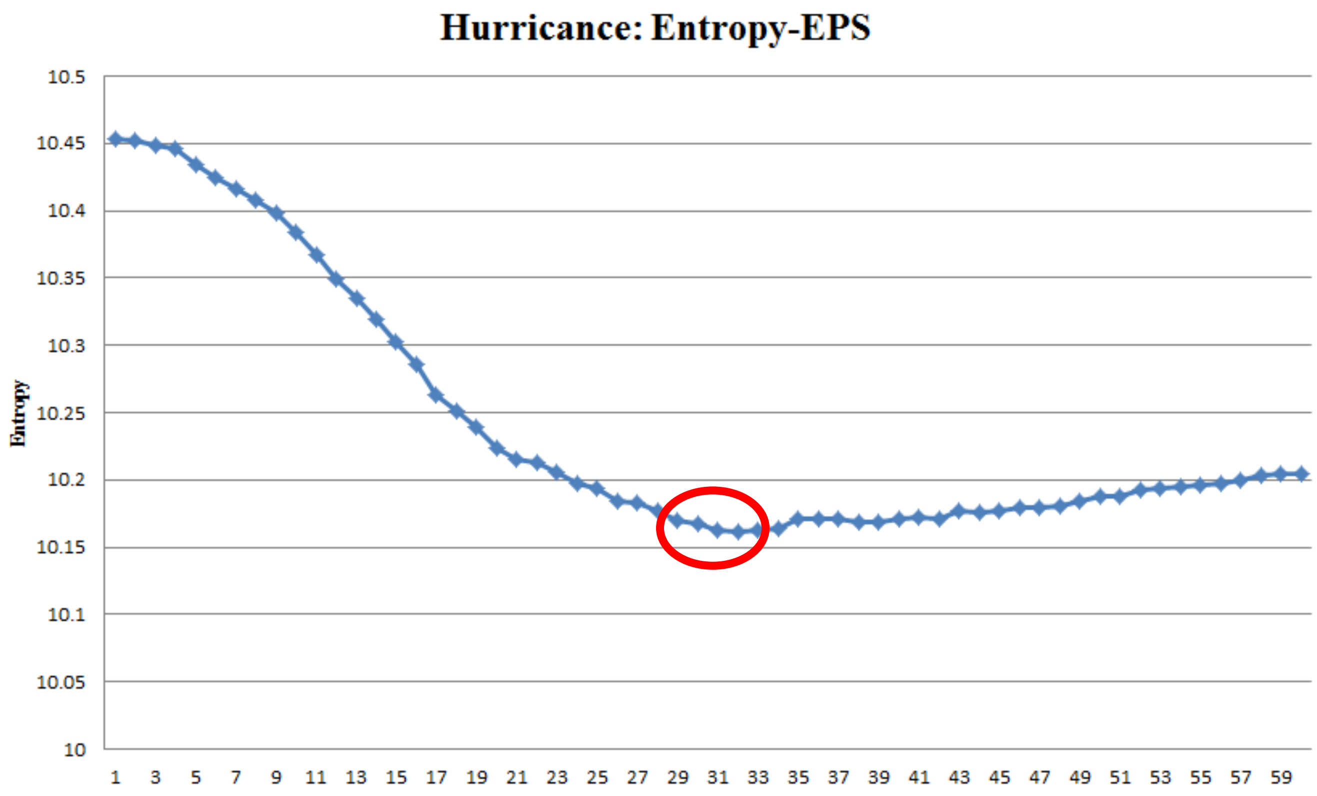


Screenshot from our result:
608 trajectories, 17736 points

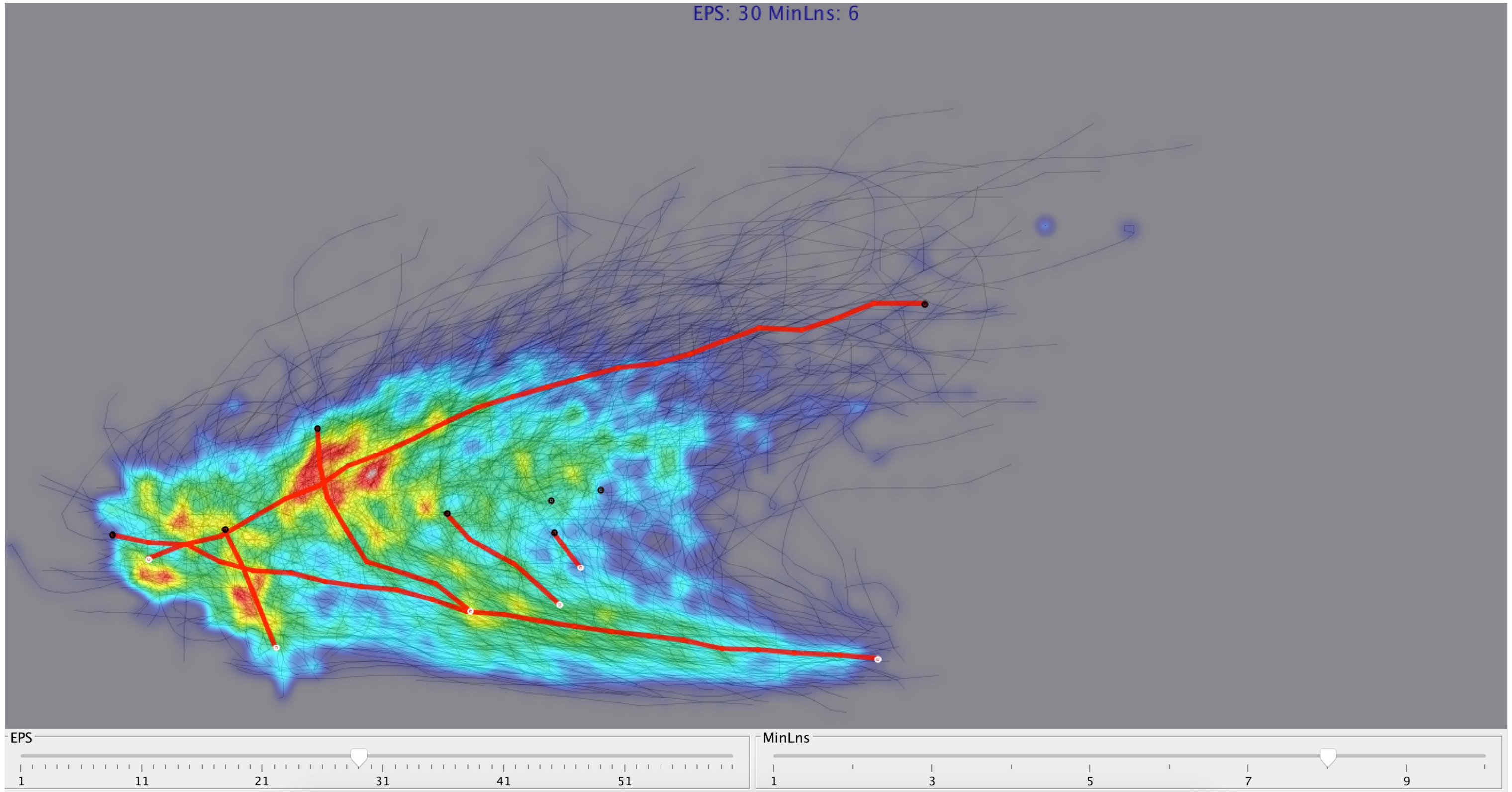


Screenshot from paper:
570 trajectories, 17736 points

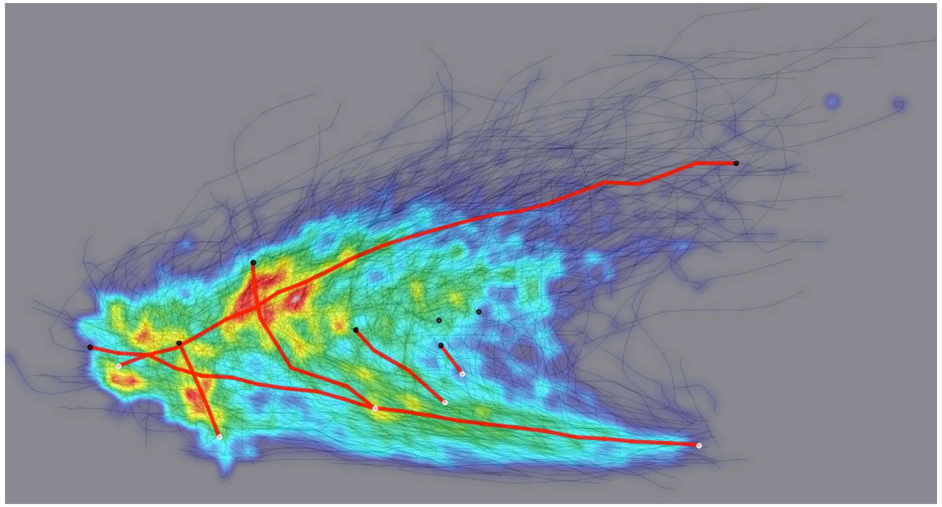
- Choose the parameter for Hurricane Data
- Using entropy theory to select ε: minimum entropy



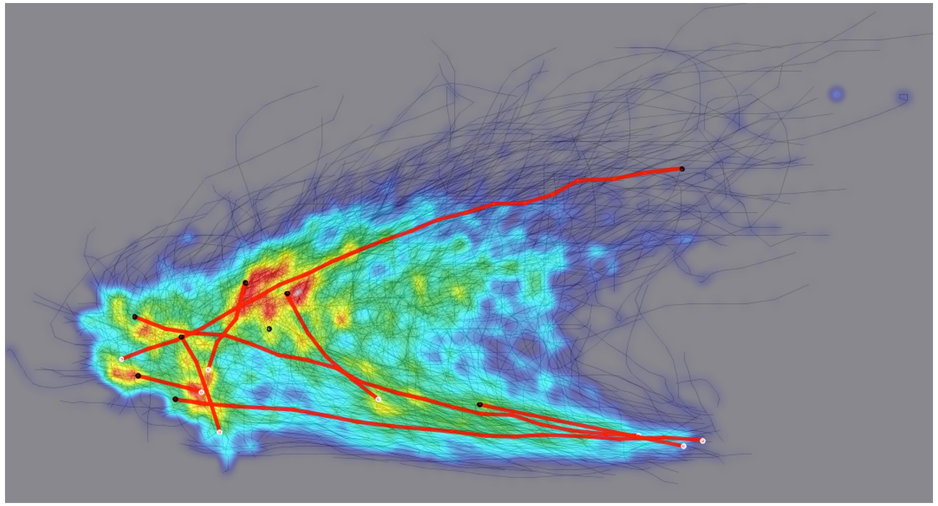
➢Hurricane Track Data($\varepsilon = 30, MinLns = 6$)



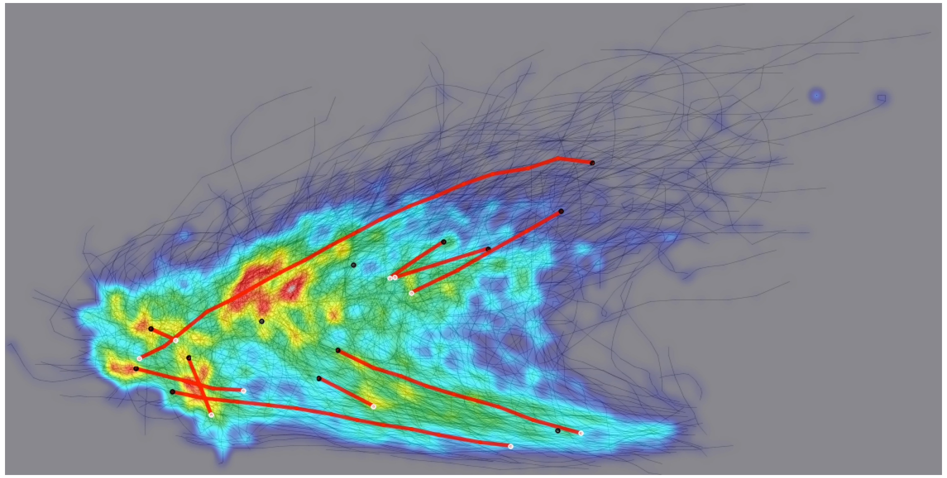
Red: Representative Trajectories of clusters by TRACLUS algorithm



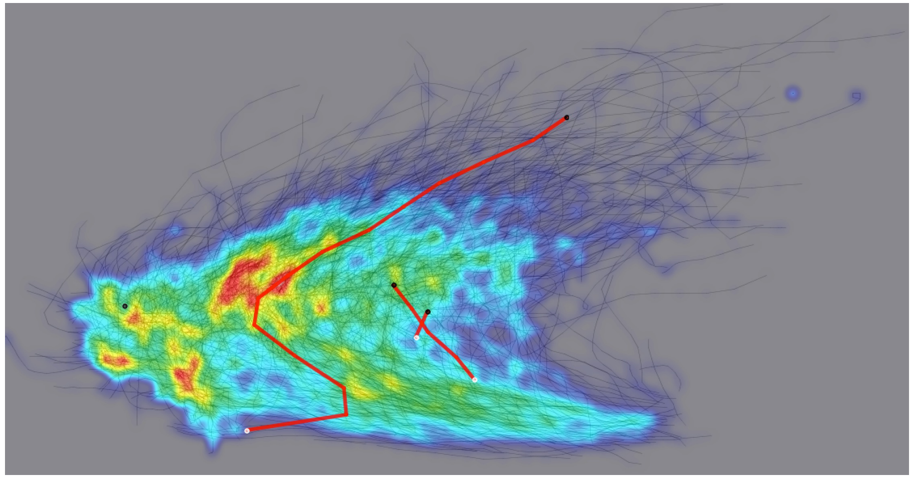
$\varepsilon = 30, MinLns = 6$



$\varepsilon = 28, MinLns = 6$

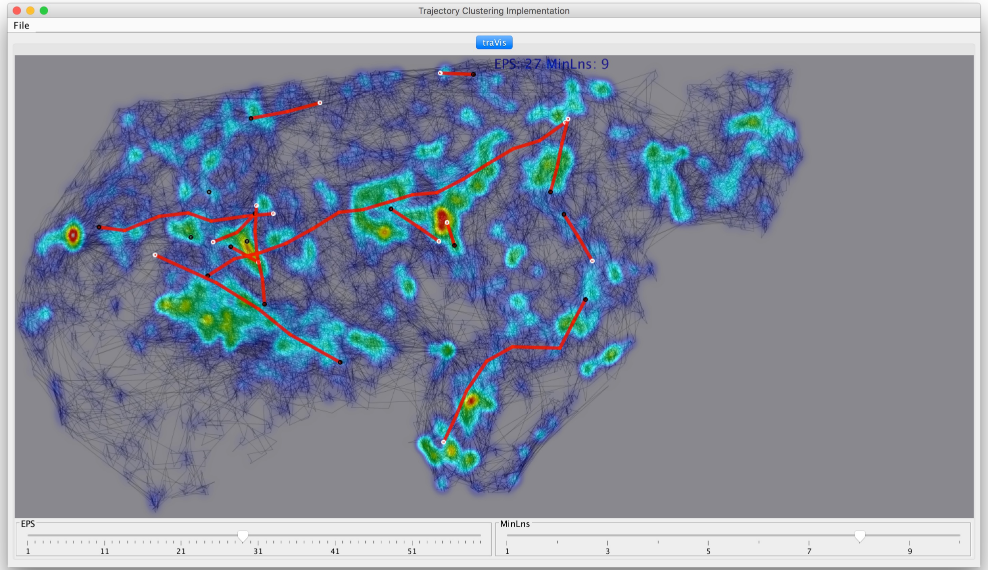


$\varepsilon = 25, MinLns = 6$

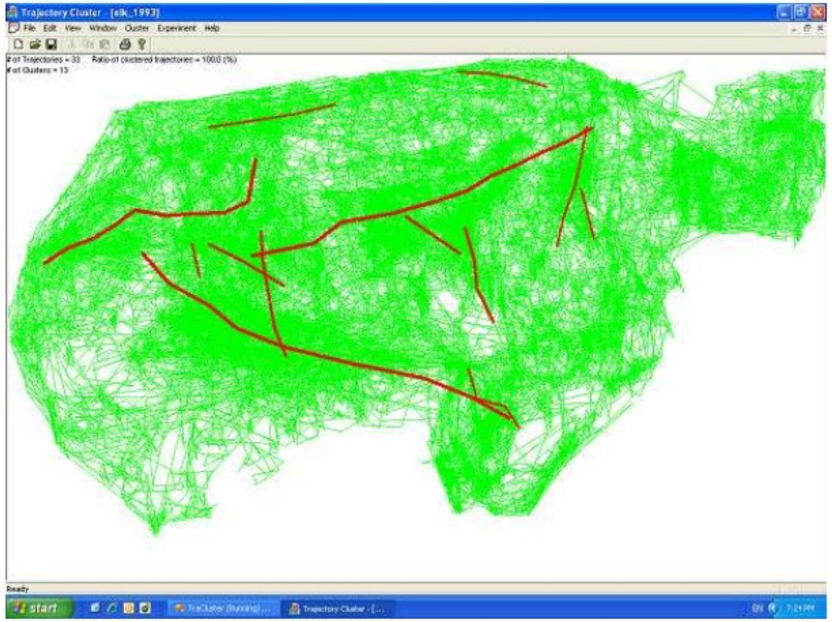


$\varepsilon = 35, MinLns = 6$

Animal movement Data (Elk1993) ($\varepsilon = 27, MinLns = 9$)

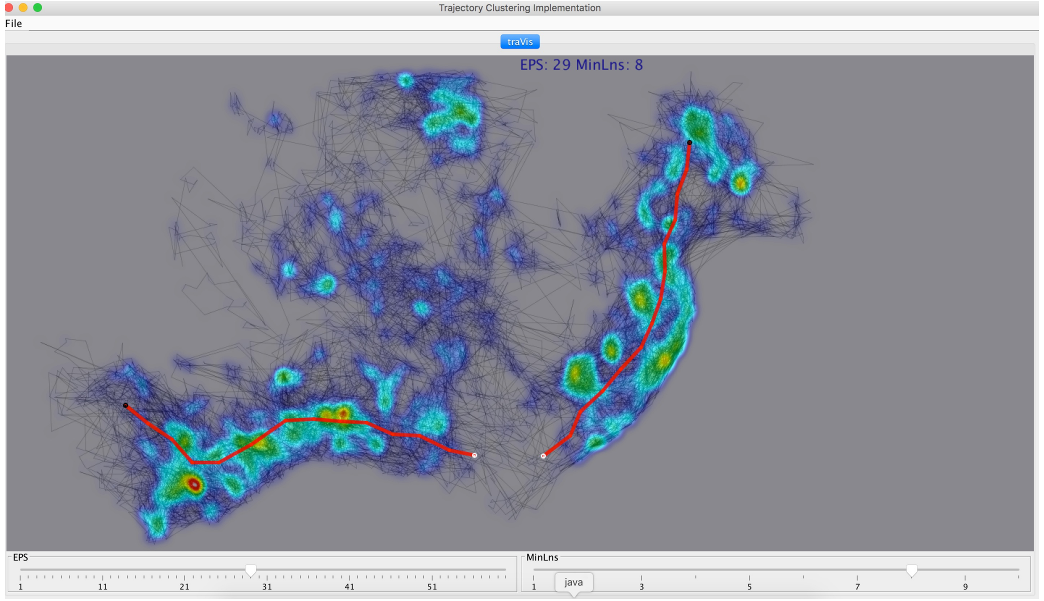


Screenshot from our result:
33 trajectories,
47199 points

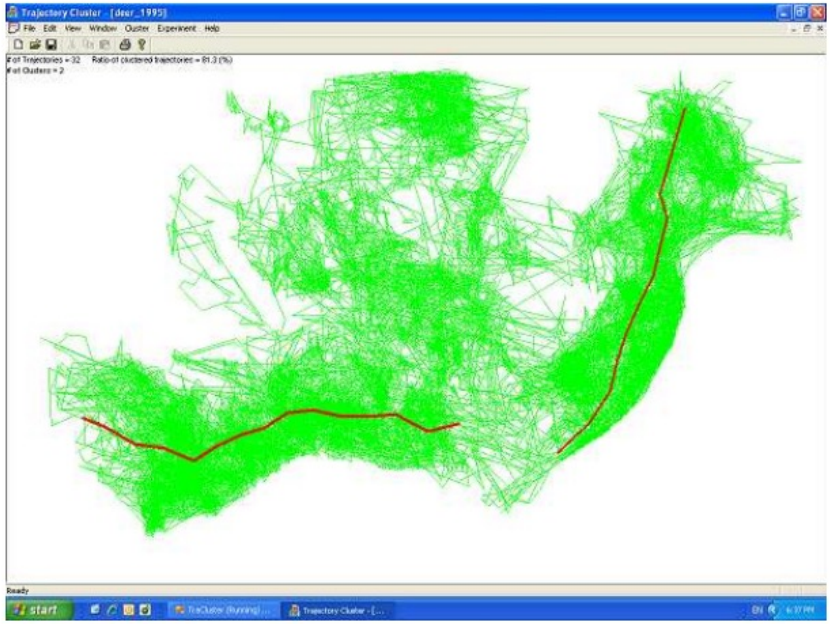


Screenshot from paper:
33 trajectories,
47204 points

Animal movement Data (Deer1995) ($\varepsilon = 29, MinLns = 8$)



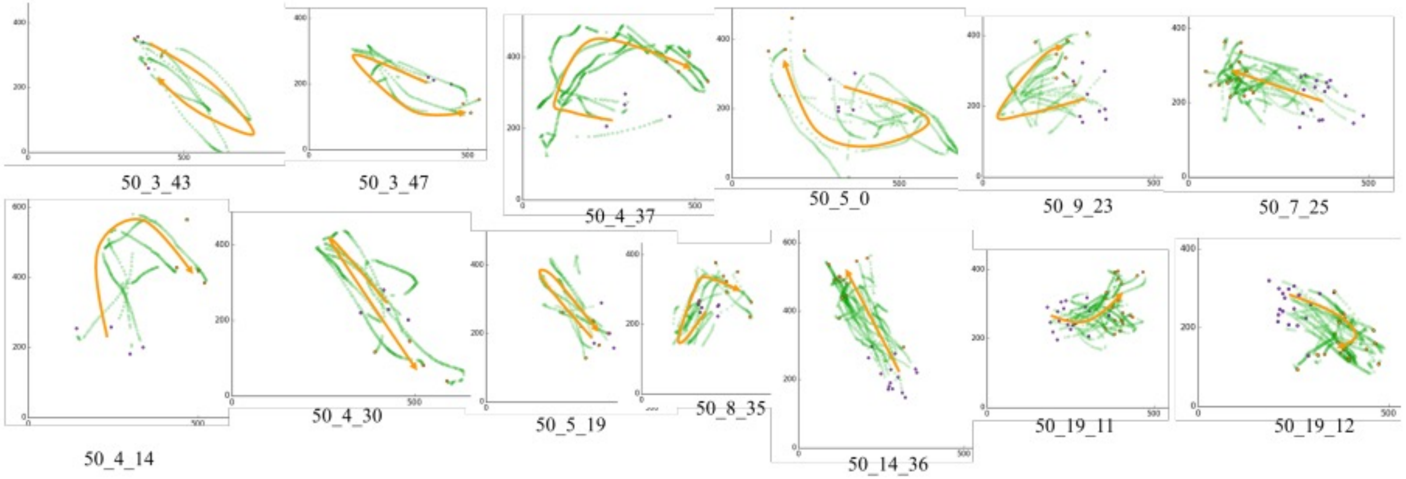
Screenshot from our result



Screenshot from paper

32 trajectories, 20065 points

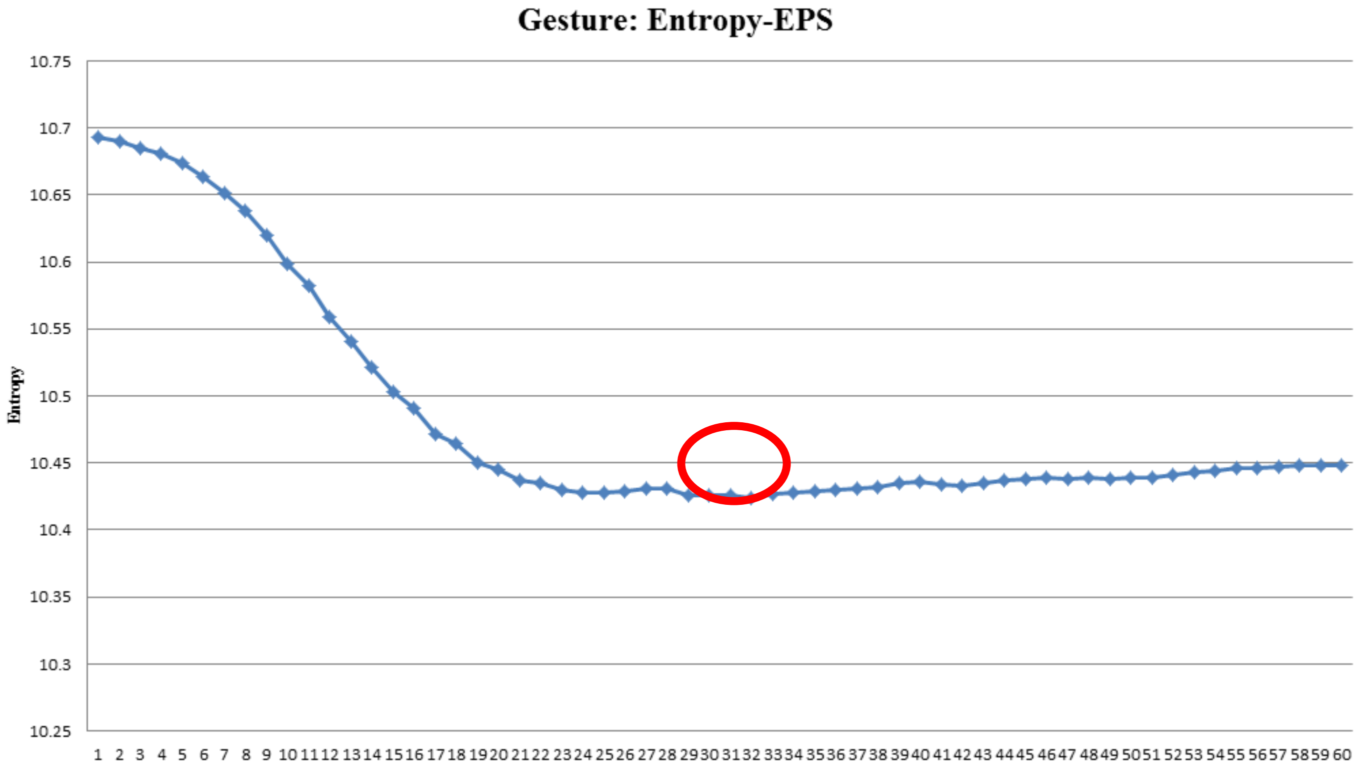
Application: Gesture Trajectory Analysis of a mobile game



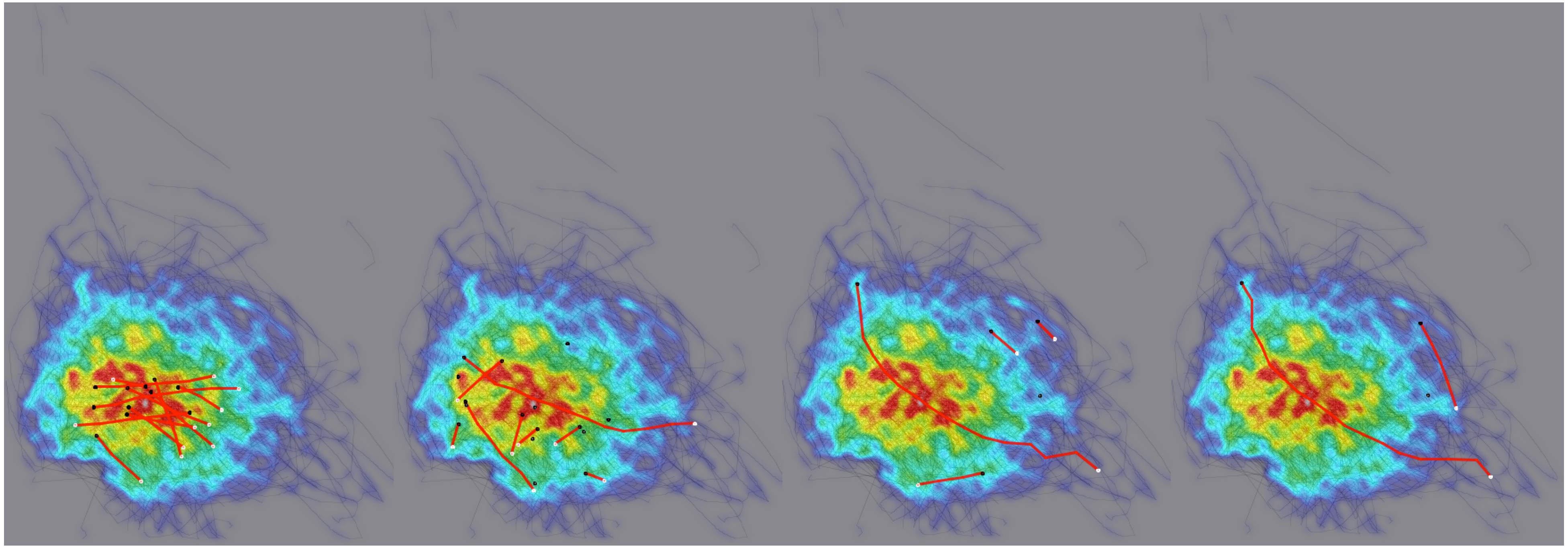
(treat trajectory as a whole)

Result

➢ Choose the parameter for Gesture Trajectory



➢ Apply TRACLUS algorithm to gesture trajectory



$(\varepsilon = 20, \mathrm{MinLns} = 6)$ $(\varepsilon = 25, \mathrm{MinLns} = 6)$ $(\varepsilon = 32, \mathrm{MinLns} = 6)$ $(\varepsilon = 36, \mathrm{MinLns} = 6)$

Blue: Original trajectories

Heat map: Density distribution of trajectories

Red: Representative trajectories of clusters by TRACLUS algorithm

3.3 Classification

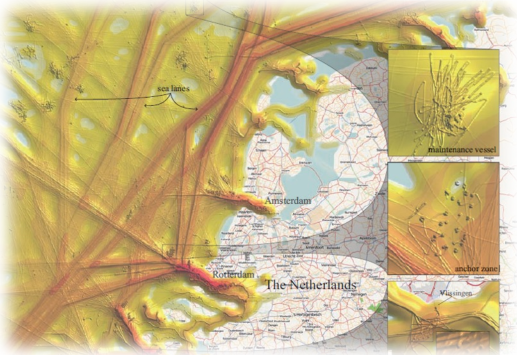
- **Predict** the class **labels** of moving objects based on their trajectories and other features
- Machine learning techniques
  - Studied mostly in pattern recognition, bioengineering, and video surveillance
  - The hidden Markov model (HMM)
- Trajectory-based classification (TraClass): Trajectory classification using hierarchical region-based and trajectory-based clustering

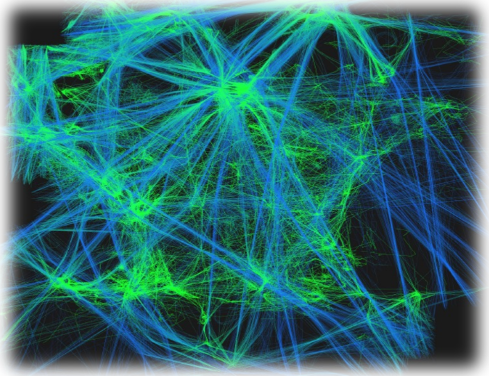
3.4 Outlier Detection

- Detect trajectory outliers that are different from or inconsistent with the remaining set of trajectories
- ***Whole*** Trajectory Outlier Detection
  - An unsupervised method
  - A supervised method *based on classification*
- ***Partial*** Trajectory Outlier Detection
  - The Partition-and-Detect Framework

4 Trajectory Visualization Techniques

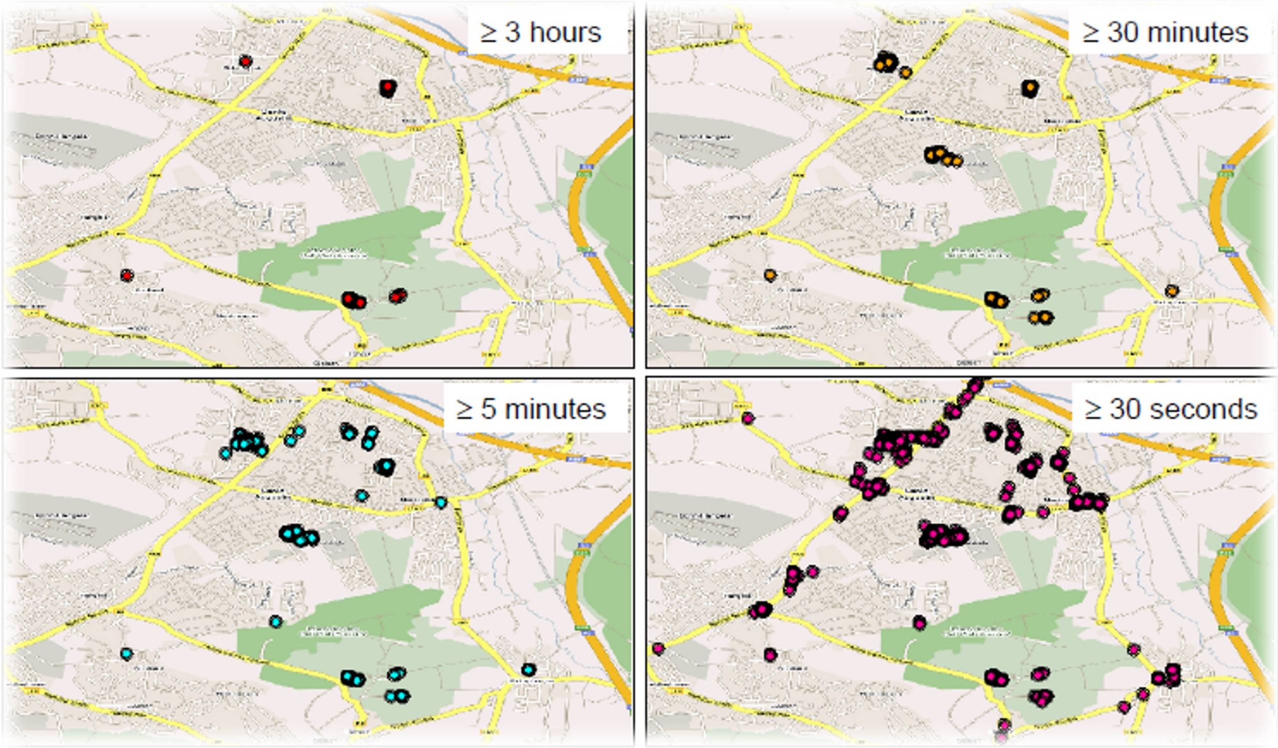
- 4.1 Trajectory Spatial Attribute
- 4.2 Trajectory Temporal Attribute
- 4.3 Other Trajectory Attributes
- 4.4 User Interaction
- 4.5 Clutter Reduction
- 4.6 Visual Analysis System





4.1 Visualization of Trajectory Spatial Attribute (1/3)

- Points



(Andrienko et al., 2000)

4.1 Visualization of Trajectory Spatial Attribute (2/3)

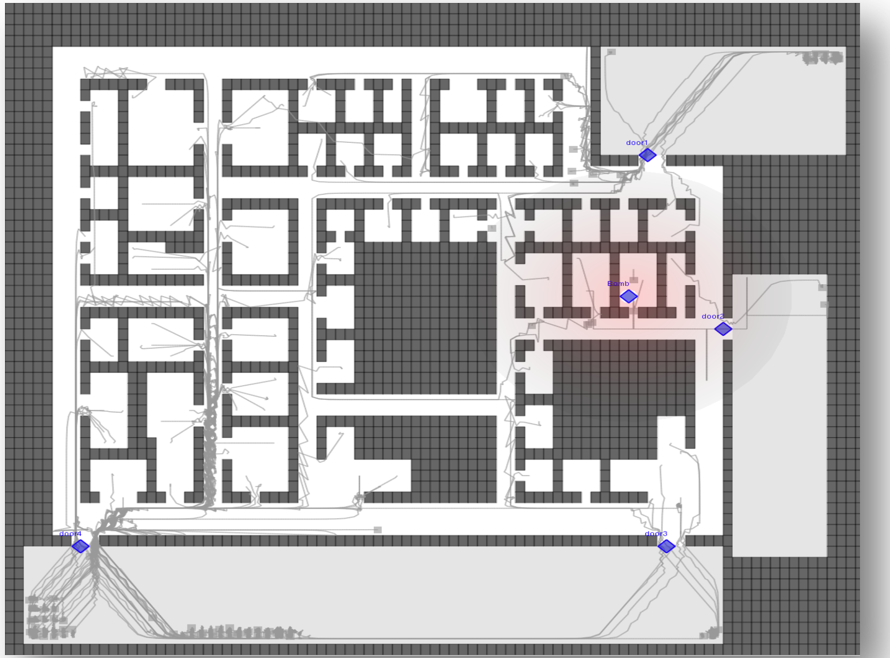
- Curves

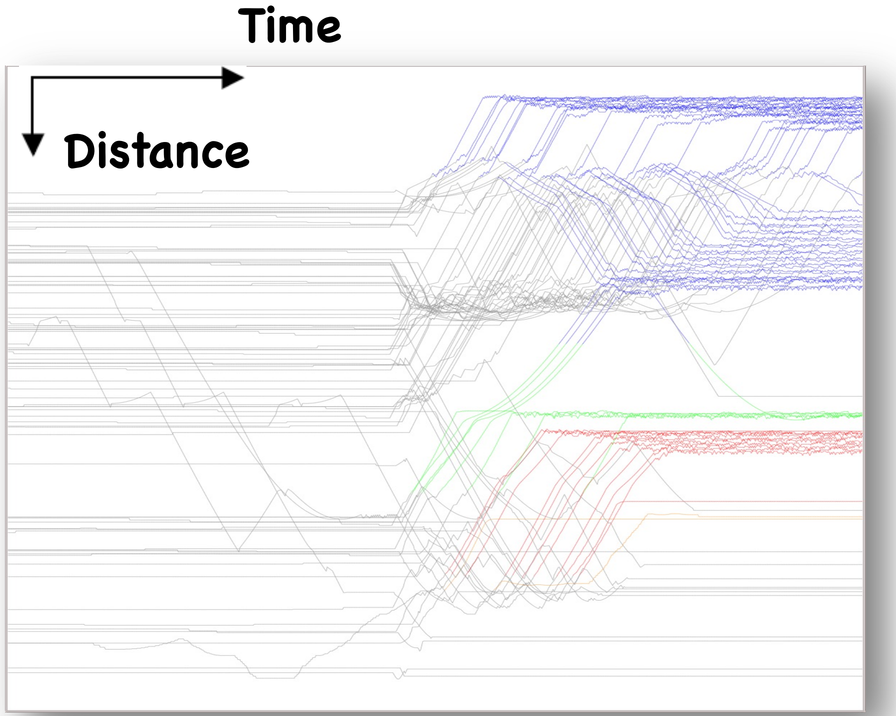
Tracks of 35 storks during 8 years, about 2,000 positions

(Andrienko et al., 2000)

4.1 Visualization of Trajectory Spatial Attribute (3/3)

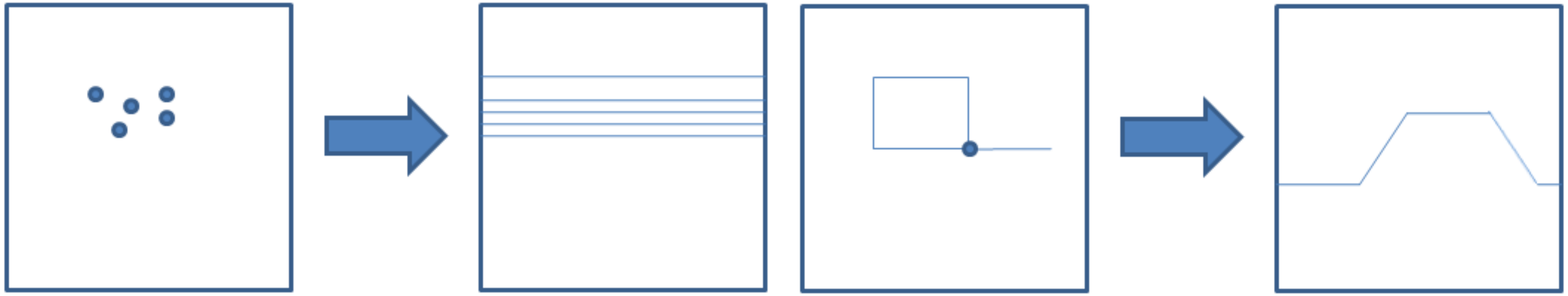
- Proximity-based Approach





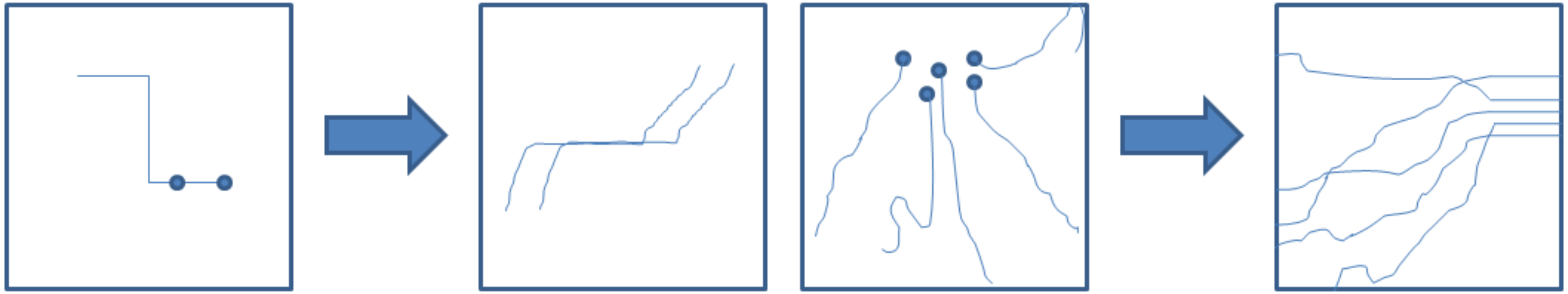
(Crnovrsanin et al., 2009)

4.1 Visualization of Trajectory Spatial Attribute (3/3)



Spatial concentration

Coincidence

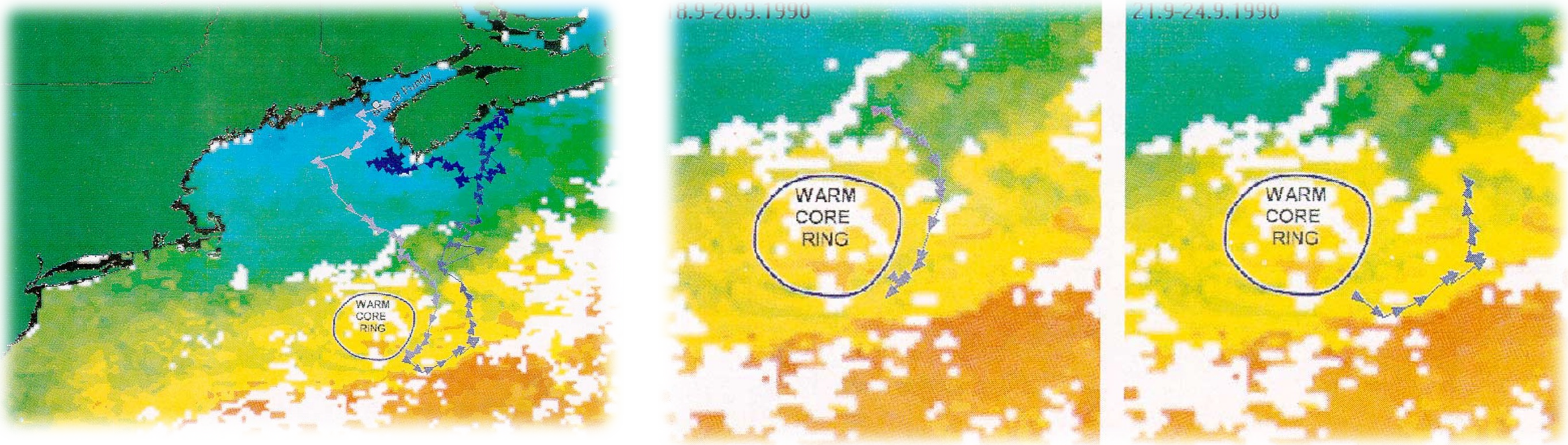


Convergence

Divergence

4.2 Visualization of Trajectory Temporal Attribute (1/4)

- Animations
  - Animation of White Whale Feeding Movement

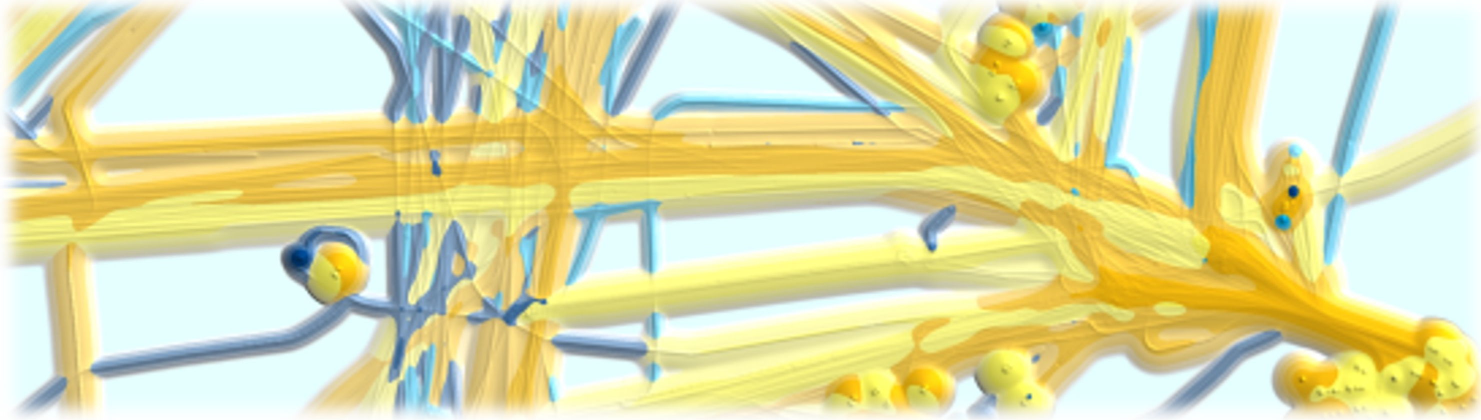


(Andrienko et al., 2000)

4.2 Visualization of Trajectory Temporal Attribute (2/4)

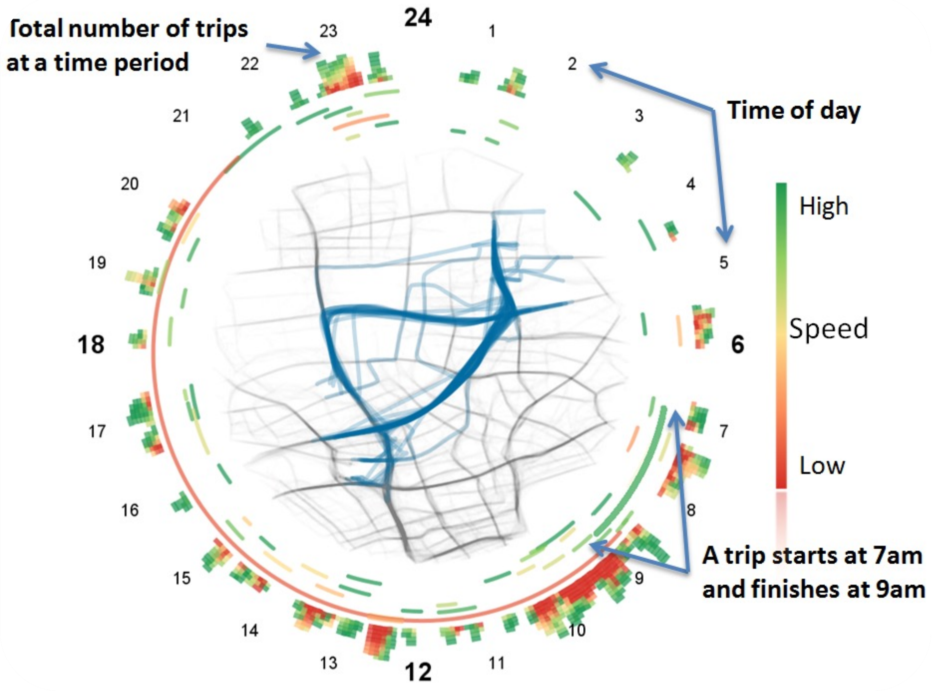
- Color
  - Night: dark blue
  - Morning: bright yellow
  - Afternoon : dark yellow
  - Evening: bright blue

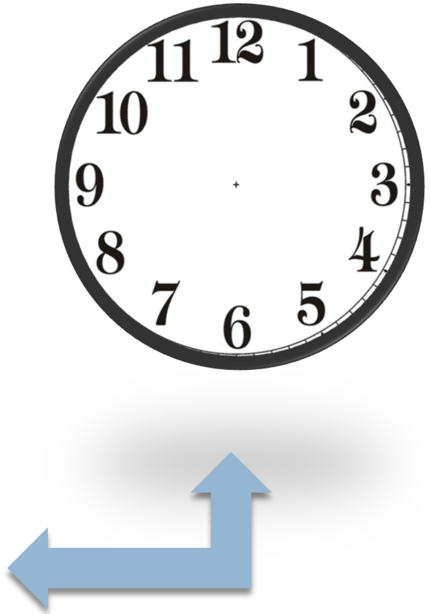
(Willems et al., 2011)



4.2 Visualization of Trajectory Temporal Attribute (3/4)

- Layout/Shape



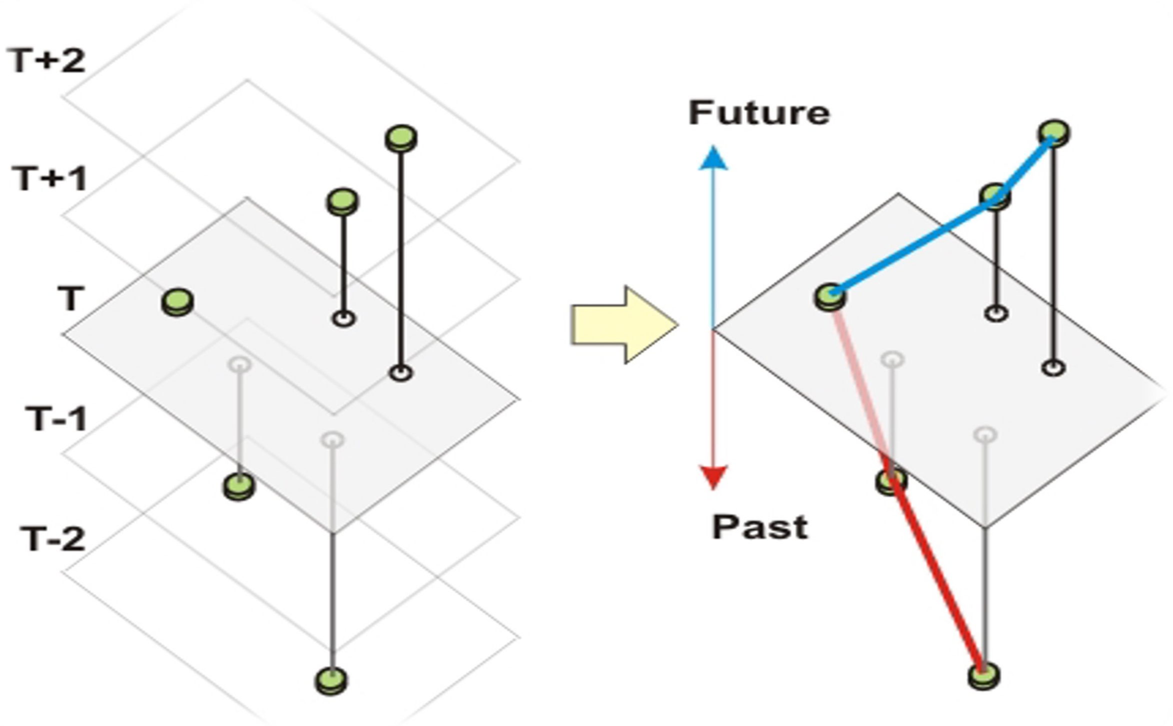


(Liu et al., 2010)

4.2 Visualization of Trajectory Temporal Attribute (4/4)

上海科技大学 ShanghaiTech University

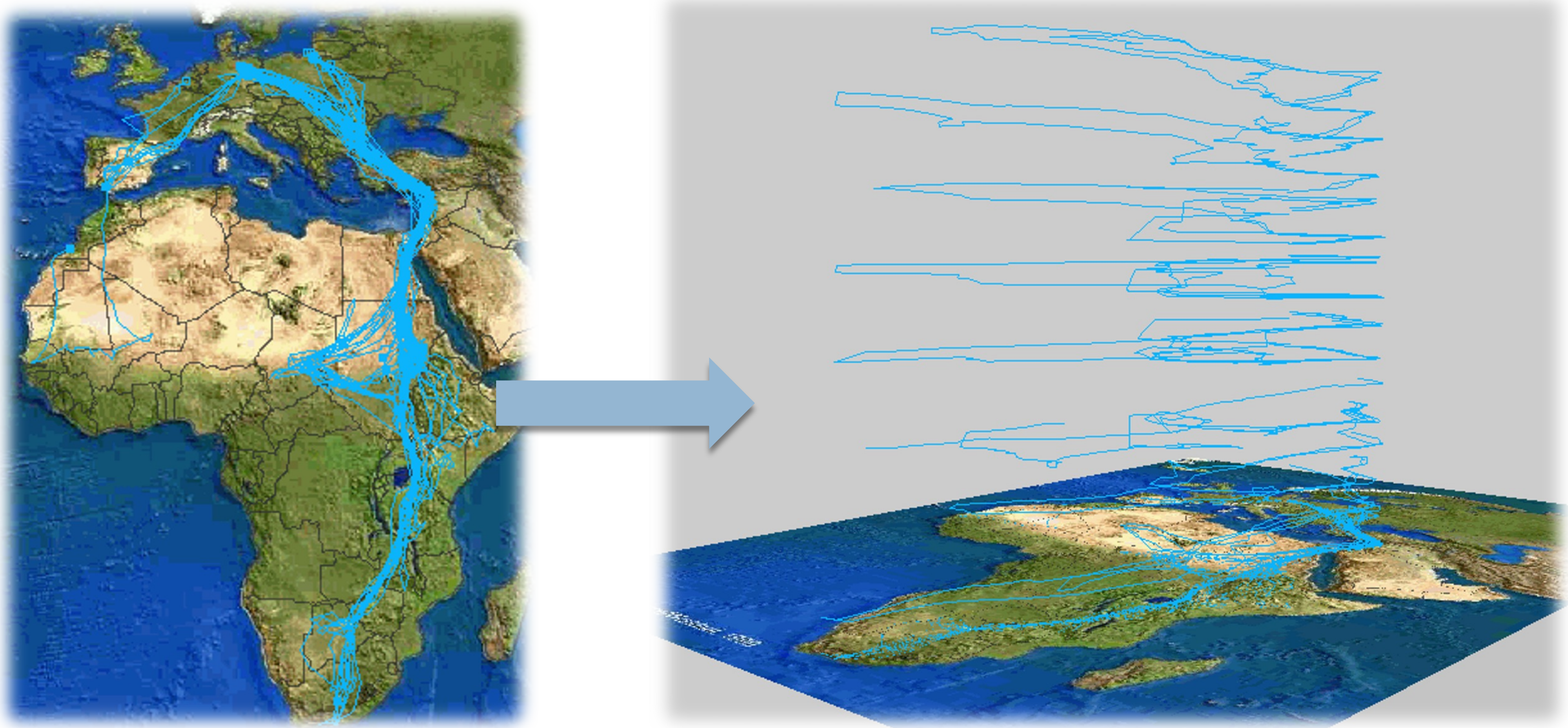
- 3D Views/Z-axis



(Thomas Kapler & William Wright, 2005)

4.2 Visualization of Trajectory Temporal Attribute (4/4)

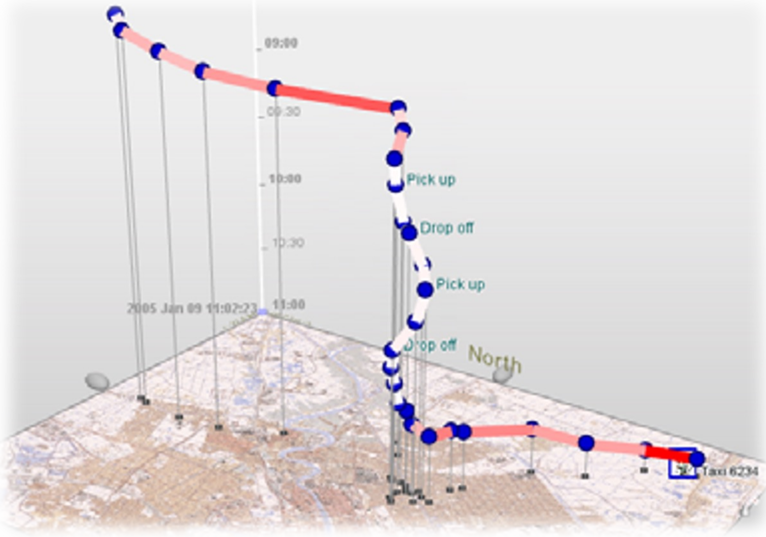
- 3D Views/Z-axis



4.3 Visualization of Other Trajectory Attributes (1/3)

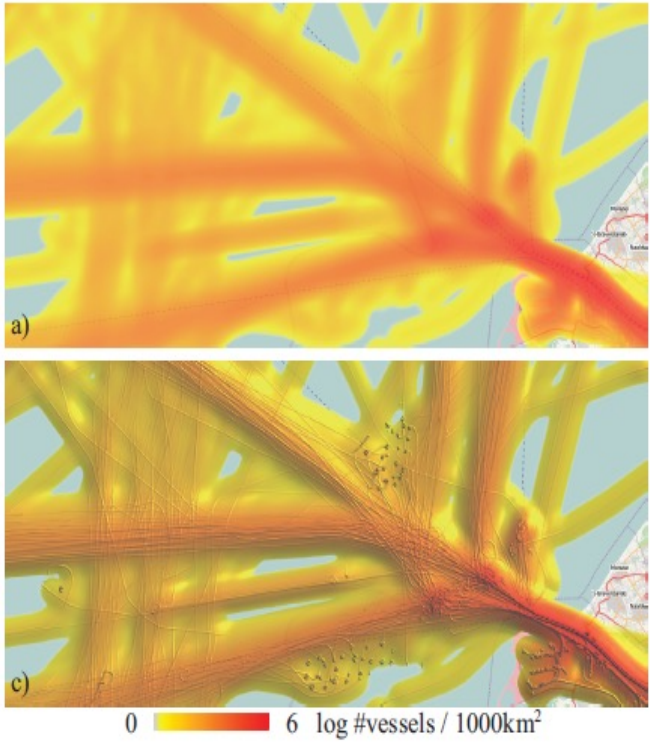
- Speed
  - Annotated snapshot result of the velocity annotation function.
  - Path of movement is annotated with indicator of speed
  - Red is fast; white is slow.

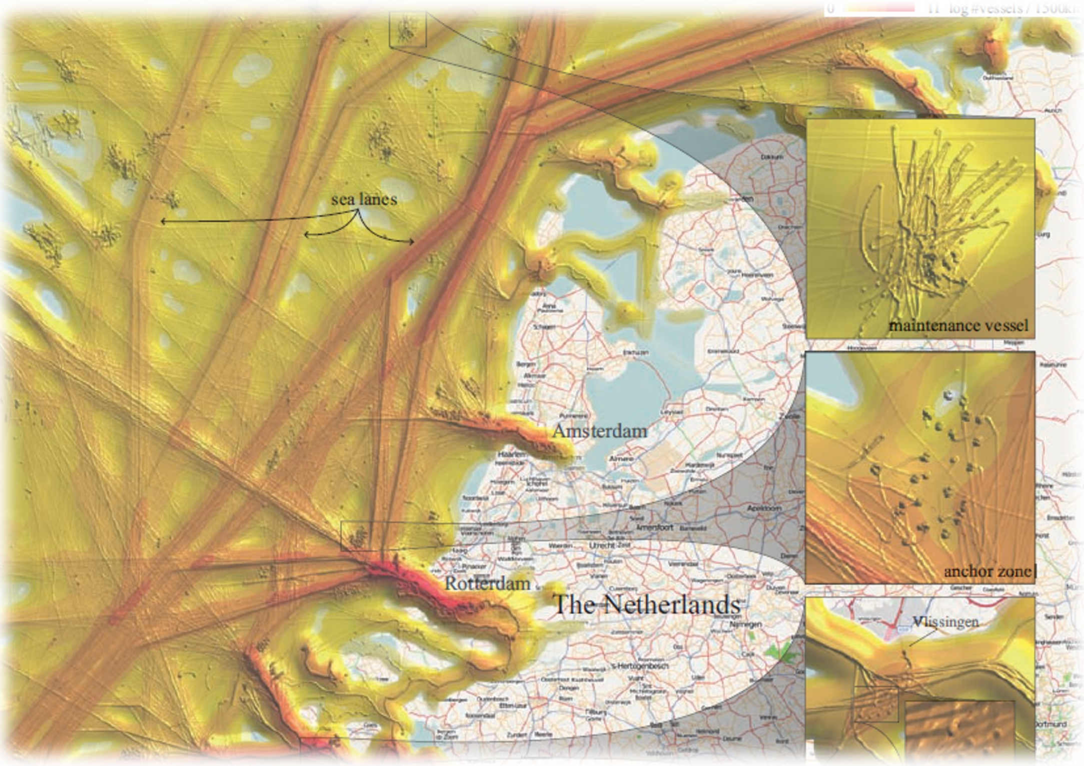
(Eccles et al., 2007)



4.3 Visualization of Other Trajectory Attributes (2/3)

- Density (Entity #)





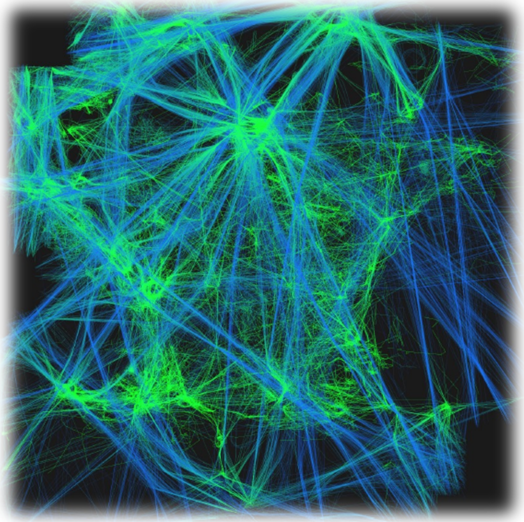
(Willems et al., 2009)

4.3 Visualization of Other Trajectory Attributes (3/3)

上海科技大学 ShanghaiTech University

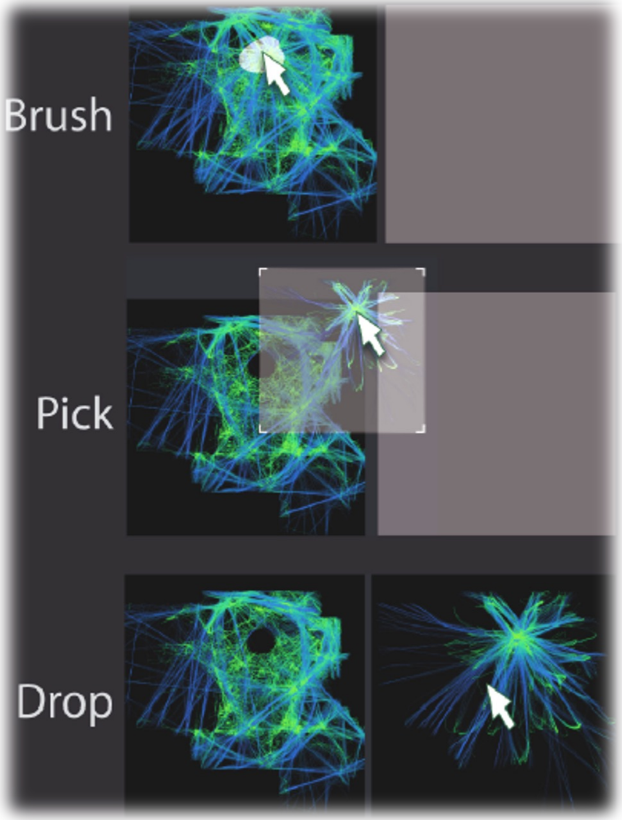
- Height
  - One day record of air traffic over France.
  - The color gradient from green to blue represents the ascending altitude of aircraft
    - Green: the lowest
    - Blue: the highest

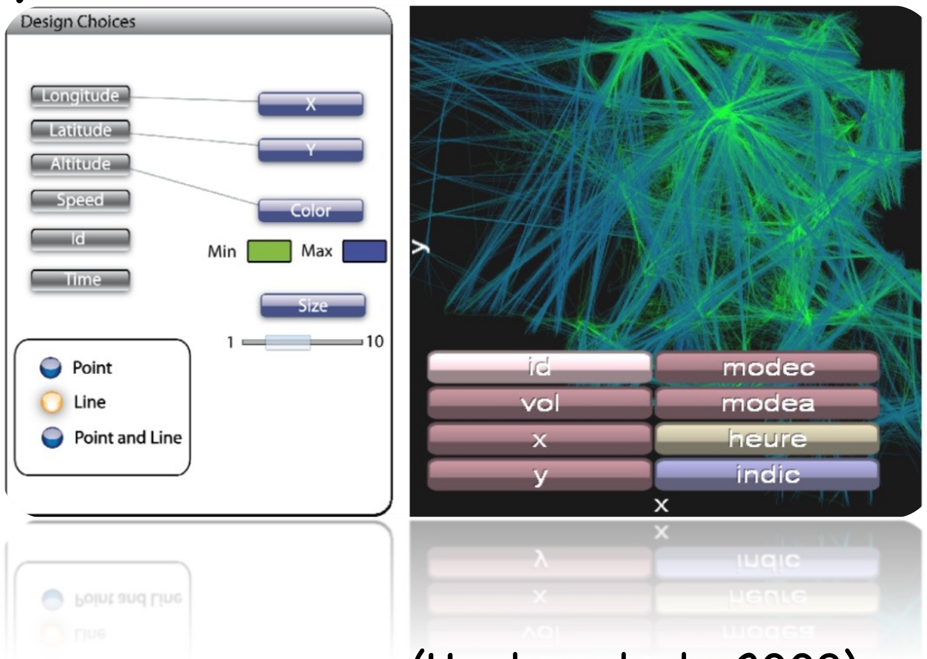
(Herter et al., 2009)



4.4 User Interaction

- Interactive Air Traffic System

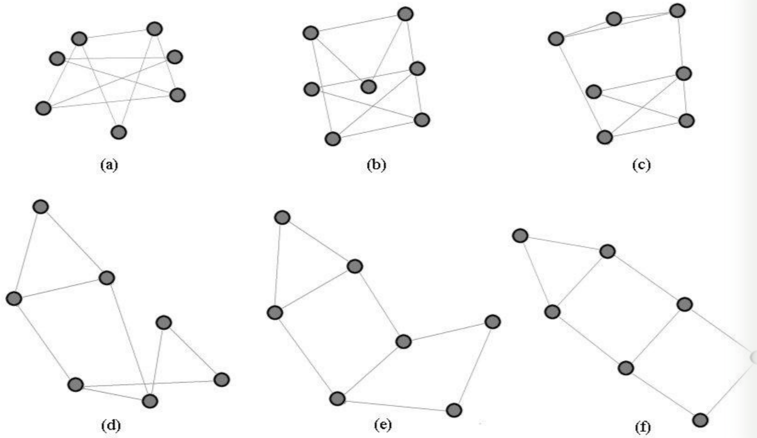


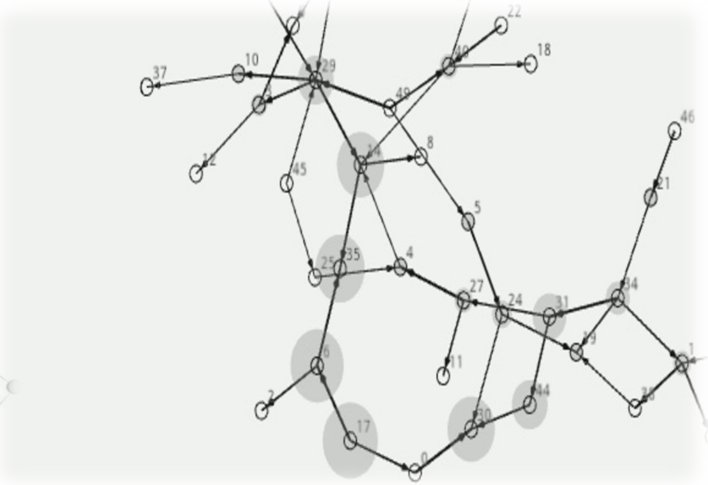


(Herter et al., 2009)

4.5 Clutter Reduction (1/2)

- Force-directed layout
  - Edges are viewed as elastic, attracting adjacent nodes
  - Nodes as charged particles, repelling other nodes

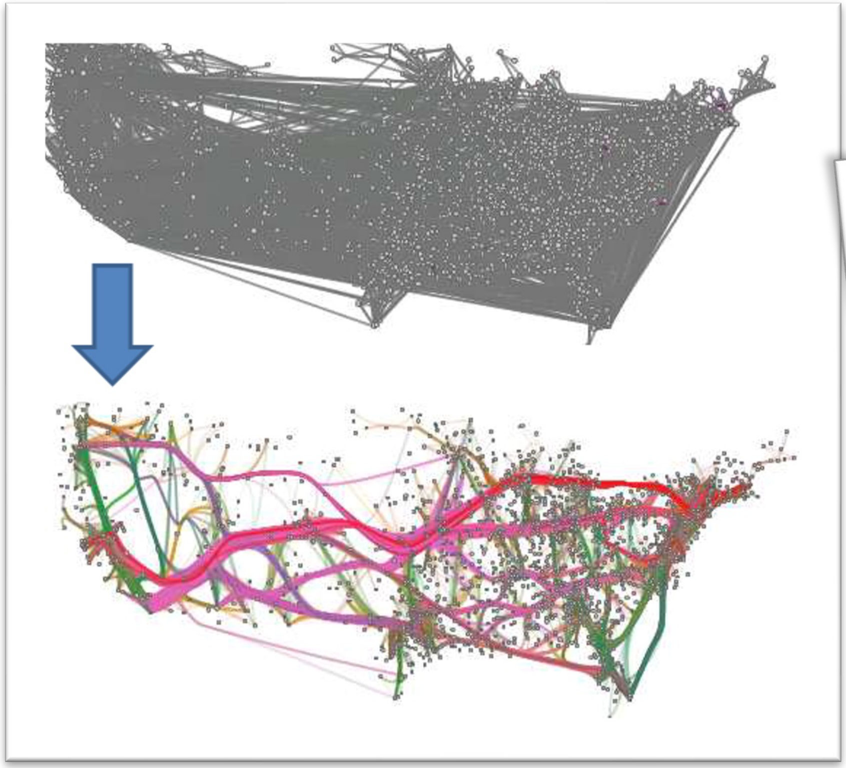




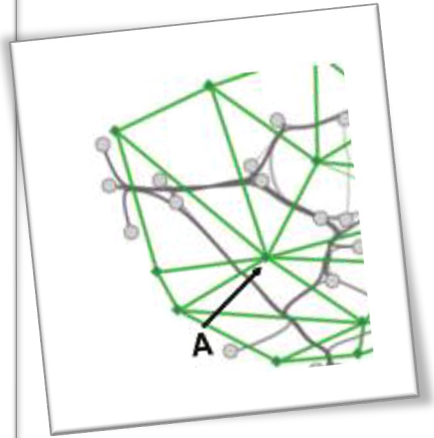
(Heer & Boyd, 2005)

4.5 Clutter Reduction (2/2)

- Geometry based edge clustering
  - Edges routed through a control mesh

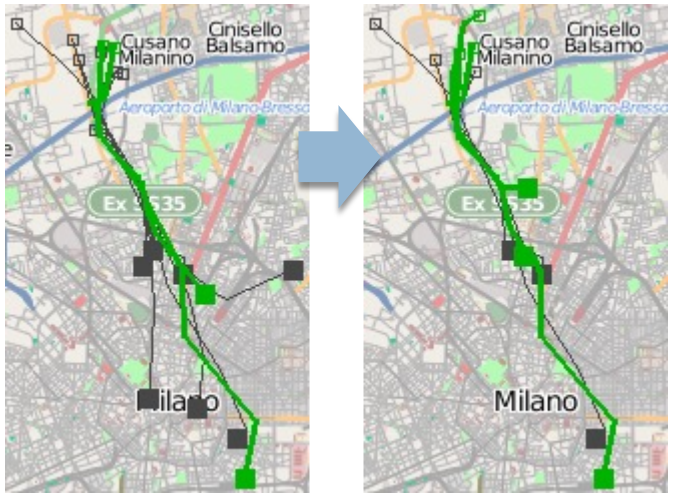


(Cui et al., 2008)



4.6 Visual Analysis System (1/2)

- An interactive visual clustering framework to find clusters in very large sets of complex objects
- To apply the clustering algorithm to a small subset of objects and then to attach the remaining objects to the clusters

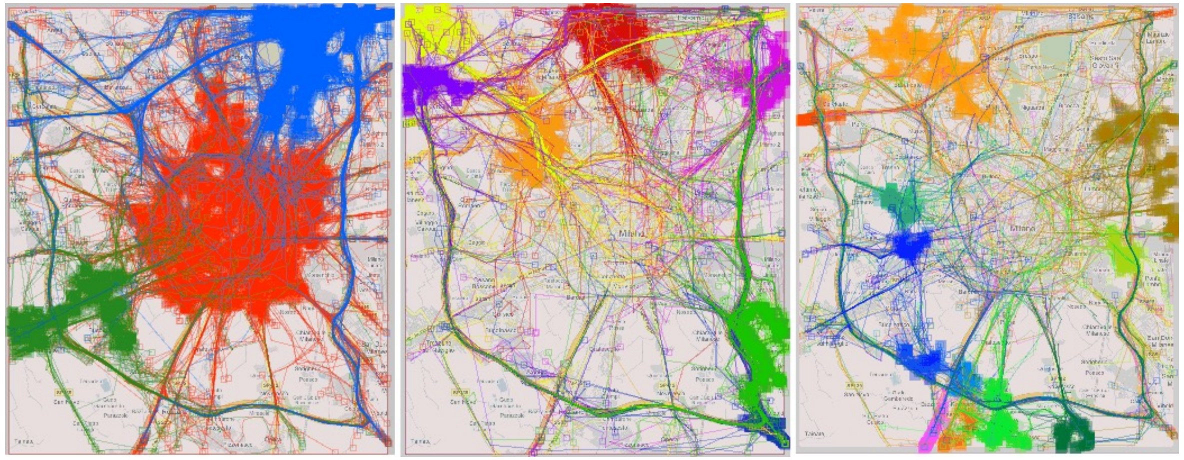


(Andrienko et al., 2009)

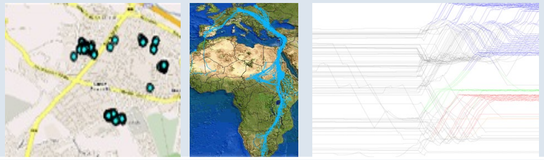
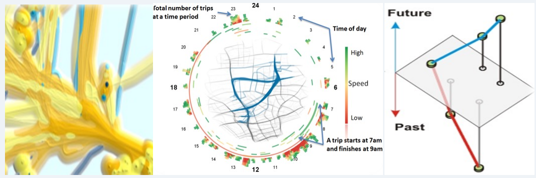
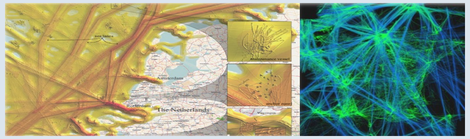
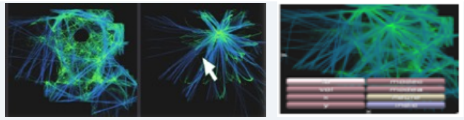
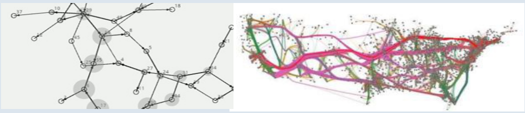
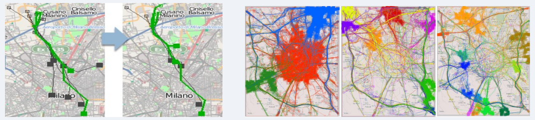
4.6 Visual Analysis System (2/2)

- Progressive clustering
  - The trajectories are very roughly clustered according to the spatial proximity of their starts and ends then further clustering with more sophisticated distance functions.

(Andrienko et al., 2009)



Comparison

	Techniques	Visualizations	Features
Spatial	Points Curves Proximately		1. Simple 2. Visual clutter
Temporal	Animation Color Shape 3D		1. Clear overview 2. Instant display vs durative display
Other	Color Height + Illumination		Clear visual patterns
User Interaction	Focus + Overview Brushing		User involved in the loop
Clutter Reduction	Force-directed Model Edge Bundling		1. Obscured semantics 2. Loss of information
Visual Analysis System	Interactive Progressive		Visual comparison and refinement

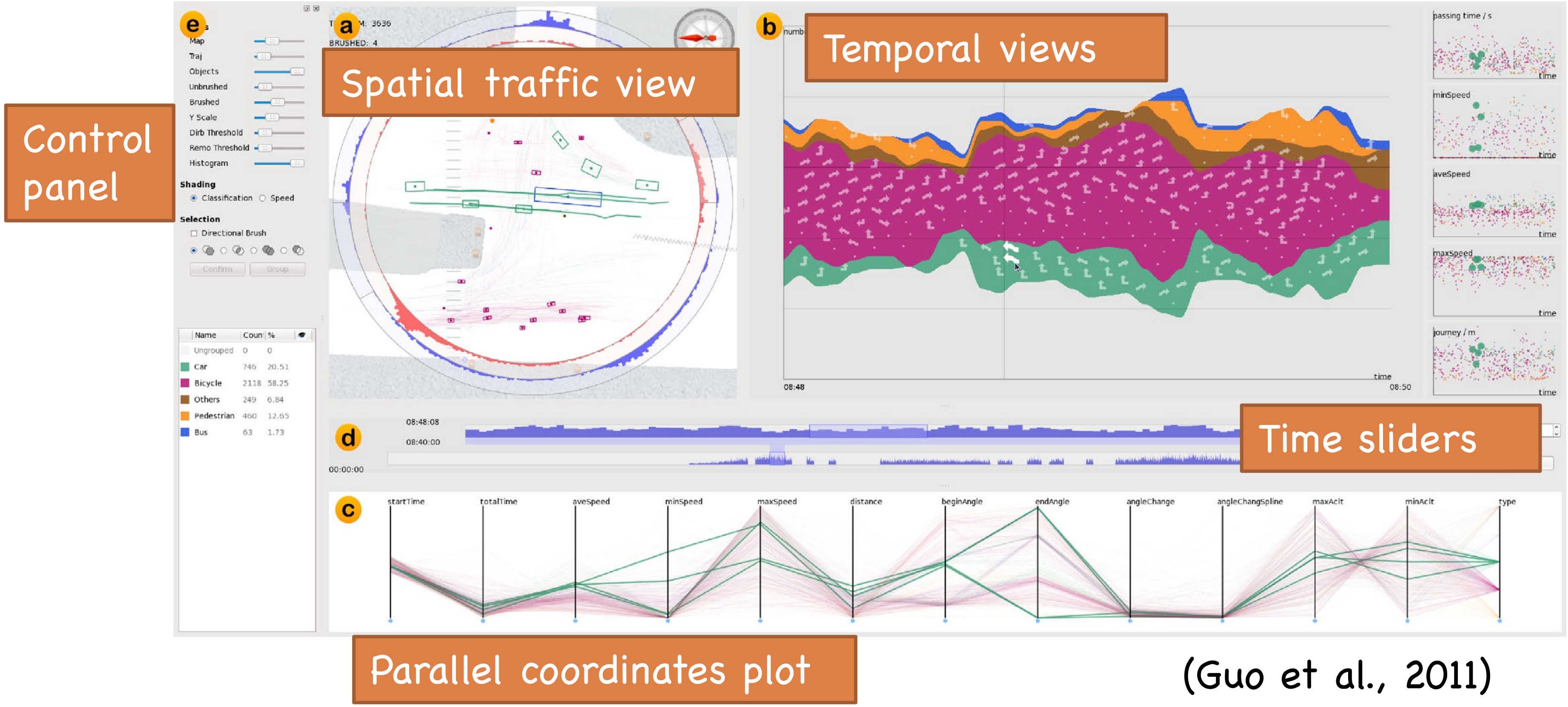
上海科技大学 ShanghaiTech University

5 Applications

- 5.1 Transportation Management
- 5.2 Mobility Intelligence of Drivers
- 5.3 Route Suggestions

5.1Transportation Management

- Triple Perspective Visual Trajectory Analytics (TripVista)

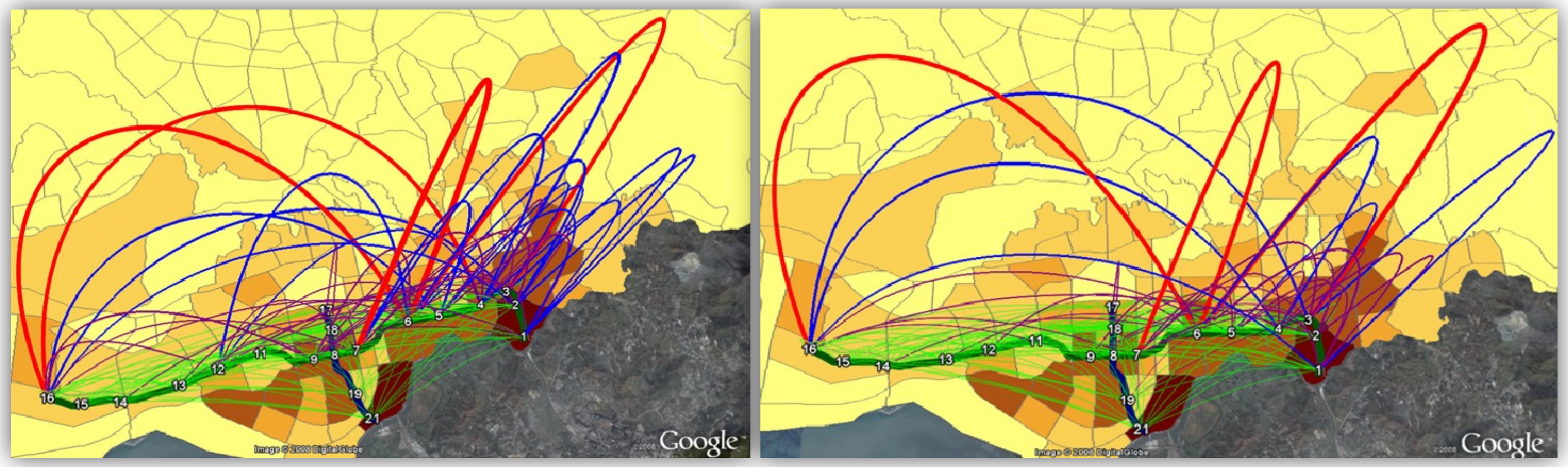


(Guo et al., 2011)

[redacted]

5.2 Mobility Intelligence of Drivers

- Urban Mobility Landscape System (UMLS)
  - To analyze taxi and bus trajectories, and metro records
  - To discover residents' daily mobility patterns

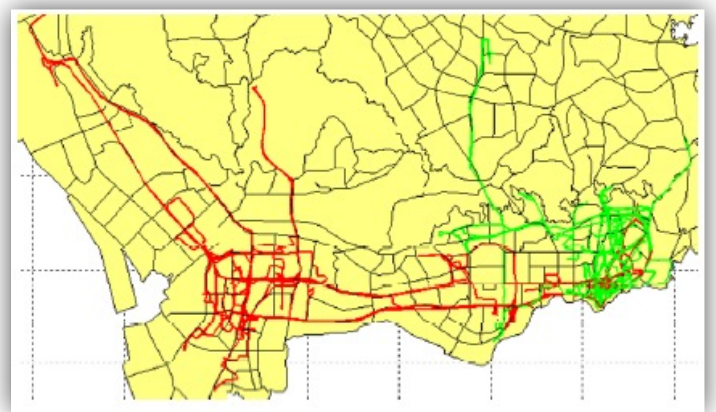


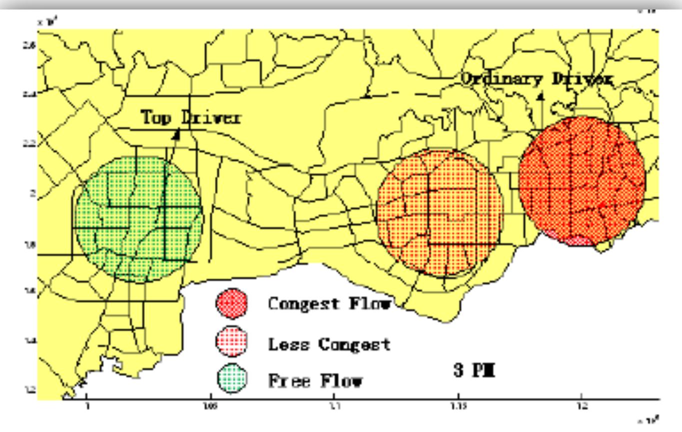
(Liu et al., 2009)

5.2 Mobility Intelligence of Drivers

- Uncover Taxi Driver's Mobility Intelligence
  - Comparing the operating behaviors of several top-income taxi drivers with several lower income drivers
  - Top drivers operate in regions with better traffic conditions

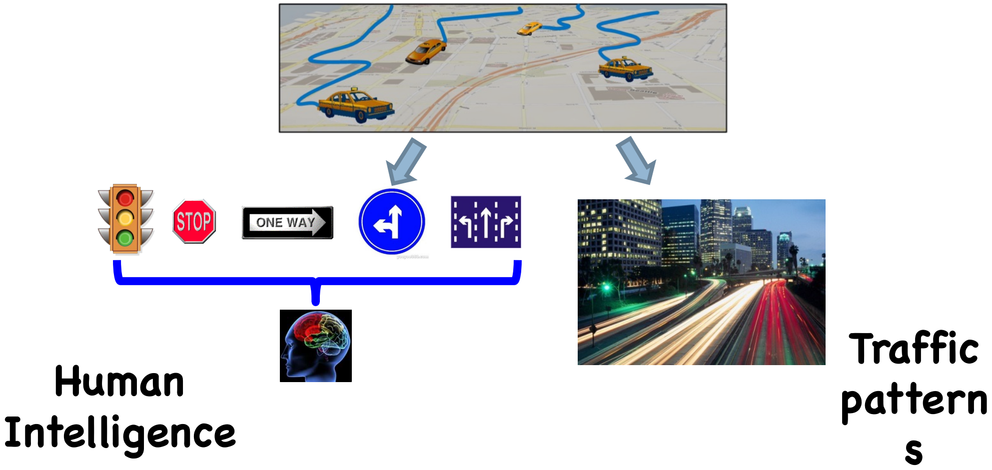
(Liu et al., 2009)





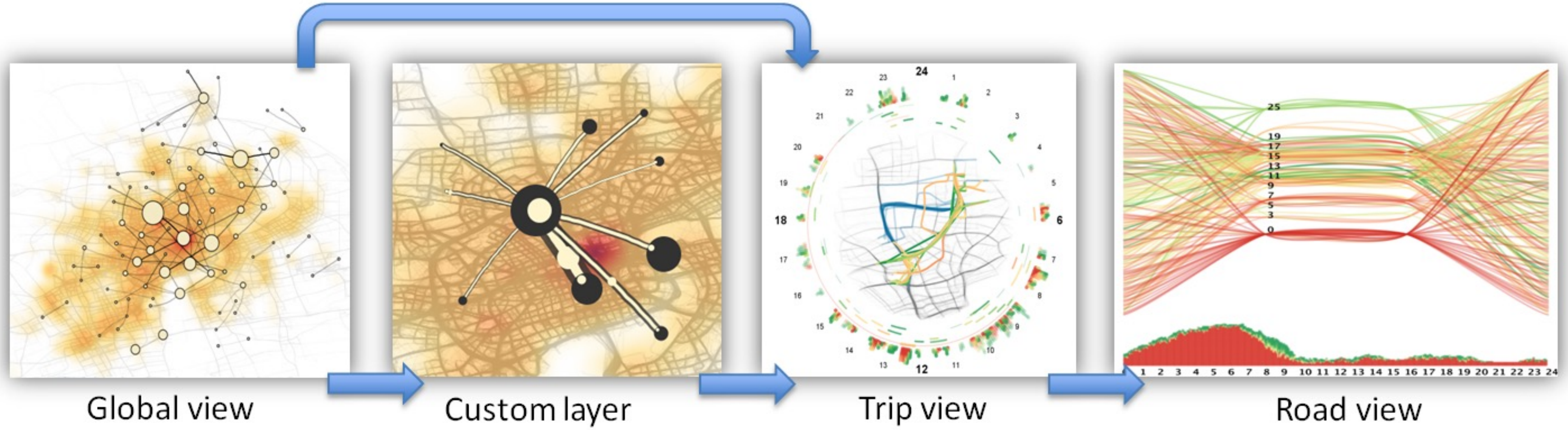
5.3 Route Suggestions

- T-Drive
  - A smart driving direction service based on GPS traces of a large number of taxis
  - Find out the practically fastest driving directions with less online computation according to user queries



5.3 Route Suggestions

- Route Diversity Analysis System



(Liu et al., 2010)

6 Conclusions

- 6.1 Conclusions
- 6.2 Discussion
- 6.3 Future Work

6.1 Conclusions

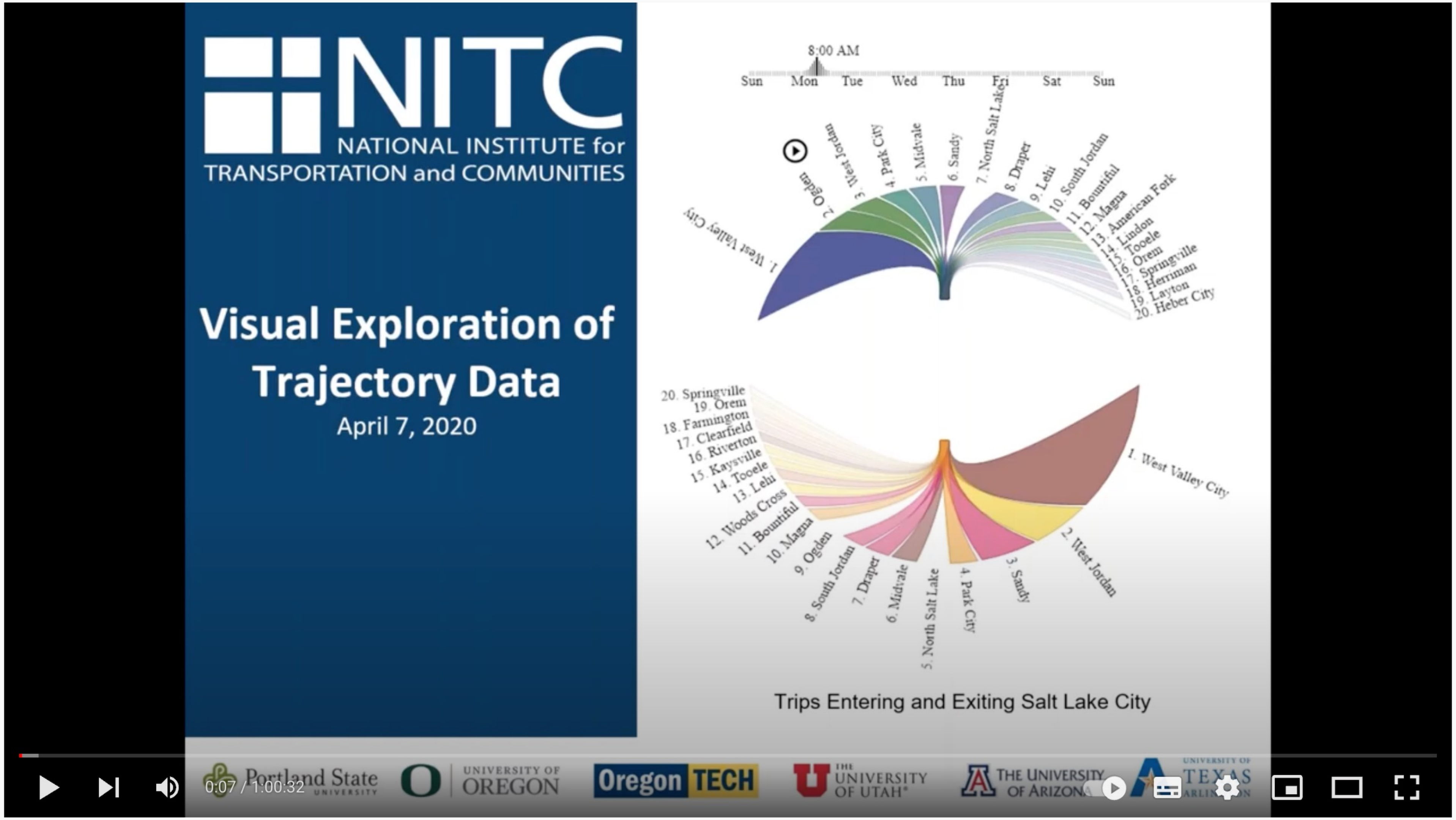
- Reviewed the trajectory data characteristics
- Introduced the traditional trajectory pattern analysis
- Summarized the trajectory visualization techniques
- Introduced three applications

6.2 Discussion

- Many work originally come from KDD or GeoVis
  - Their models and application scenario are both strong and clear.
  - Visual methods or display are some intuitive ways.
- Some work demonstrated professional visualization
  - The vessel visualization is a good example of a successful combined KDD method with professional visualization (color, illumination)

6.3 Future Work

- Incorporate data mining with interactive visual analysis system
  - To help visualizing trajectory dataset in a large scale
  - To help analyzing the fuzzy patterns and complex correlations
- Design novel trajectory visualization for further exploration
  - Trajectory visualization with ICON design
  - Trajectory visualization with geographic-based attributes projection methods



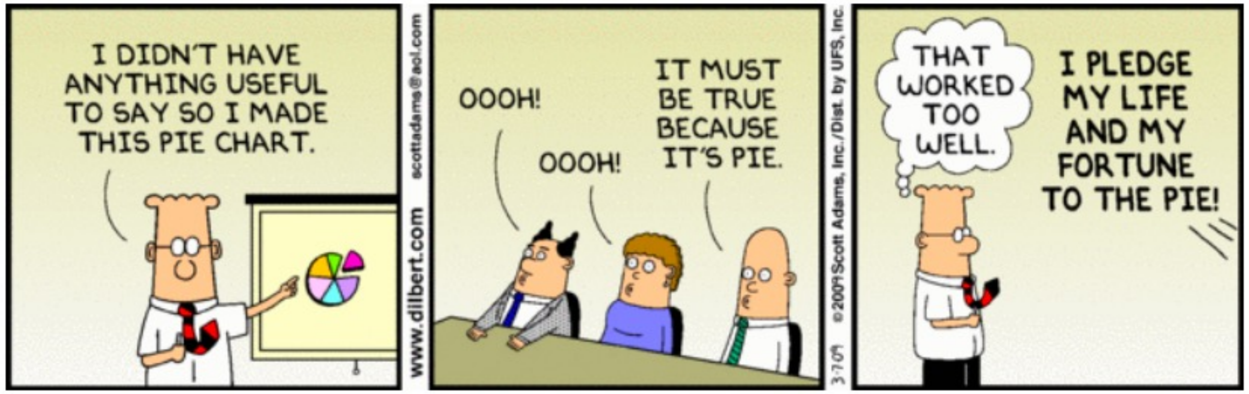
Webinar: Visual Exploration of Trajectory Data

https://www.youtube.com/watch?v=xY6g0T7MWWA

More than Design: Evaluation

Evaluation: Why?

- We have seen many complex, novel and fun/cool visualizations over the semester
  - Design e.g., river, glyph, animation, fisheye, etc.
  - Layout, e.g., radial, 3D, etc.
- But How do we determine if a visualization is effective?
  - Does the visualization actually help users?
  - Is it an improvement over
  - What aspect(s) of a visualization work/do not work?



Evaluation: When?

- Pre-design
- Design
- Prototype
- Deployment
- Redesign

Evaluation: when and what?

- Earlier stages:
  - Observe and interview target users (needs assessment)
  - Design data abstraction/operation (data types, transformation, operations)
  - Justify encoding/interaction design (design heuristics, perception research)
  - Informal analysis/qualitative analysis of prototypes (task-based)
  - Algorithm complexity analysis/evaluation
- Mid- and later stages:
  - Qualitative analysis of system (task-based)
  - Algorithm performance analysis
  - Lab or crowdsourced user study
  - Field study of the deployed system

Evaluation: What?

- Moscow Prioritization
  - Must
  - Should
  - Could
  - Won't

- Severity Rating
  - Not a real problem
  - Cosmetic
  - Minor issue
  - Major issue
  - Critical issue

Types of Evaluation Methods

- Inspection or Principled Rationale
  - Apply design heuristics, perceptual principles
  - Expert interviews
- Formal User Study
  - Controlled experiments
  - Interviews
  - Surveys
  - Usability study
- Field Deployment or Case Studies
  - Have people use visualization, observe and interview
  - Document effects on work practices
- Theoretical Analysis and Benchmarking
  - Time and space complexity
  - Performance

How to Choose a Proper Method?

- Different forms
  - Qualitative vs. quantitative, objective vs. subjective, …
  - What Metrics to use?
  - What Data to collect?

- Expert Inspection vs. User Study?
- Controlled Experiment vs. Subjective Assessment?
- Evaluating Visualization or UI?
- Usability vs. Utility?

Expert Inspection vs. User Study?

- Expert Interview
  - Participants:
    - Design experts, visualization experts, usability experts, or domain experts
  - Types
    - Heuristic evaluation, guidelines review consistency inspection, cognitive walkthrough, metaphors of human thinking, formal usability inspection, accessibility inspection, etc.
  - Pros: fast and affordable

- User Study
  - Participants: target users
  - Pros:
    - Less reliance on expert knowledge
    - Reflect real user experience
    - Less subjective

Controlled Experiment vs. Subjective Assessment

- Controlled Experiment
  - Good for measuring performance or comparing multiple techniques
  - Control confounding variables
  - Often quantitative in nature
    - Performance, time, errors, retention, subjective satisfaction, etc.

- Subjective Assessment
  - Often observational with interview
  - Learn people's subjective views on tool
    - Enjoyment, confusion, trust, etc.
    - Personal judgment strongly influence use and adoption, sometimes even overcoming performance deficits

Evaluating Visualization or UI?

- Evaluating visualization design
  - Very difficult to compare "apples to apples"; different tools were built to address different user tasks
  - Controlled experiment or subjective assessment?
- Evaluating user interface (UI) design
  - UI can heavily influence usability, utility (usefulness) and value of visualization technique
  - General principles:
    - Strive for consistency
    - Cater for universal usability
    - Offer informative feedback
    - Design dialogs to yield closure
    - Prevent errors
    - Permit easy reversal of actions
    - Support internal locus of control
    - Reduce short-term memory load

Usability vs. Utility?

- Big difference
  - Usability is not the same as utility, which seems to be a key factor for visualization
  - A visualization can be very usable but not useful or helpful
- How to define and measure success of a visualization?
  - Impact on community as a whole, influential ideas
  - Assistance to people in the tasks they care about
- Challenges in evaluation
  - More domain knowledge and situated use is required
  - Sometimes the chain of influence can be long and drawn out
    - System X influences System Y influences System Z which is incorporated into a practical tool that is of true value to people

Evaluation Decision Process

- Characteristics of a desirable evaluation method
  - Generalizability
  - Precision
  - Realism

- Four specific aspects of visualization evaluation
  - Choosing a focus
  - Picking suitable scenarios, goals, and questions
  - Considering applicable approaches
  - Creating evaluation design and planned analyses

Evaluation Scenarios

- Understanding data analysis
  - Understanding environments and work practices (UWP)
  - Evaluating visual data analysis and reasoning (VDAR)
  - Evaluating communication through visualization (CTV)
  - Evaluating collaborative data analysis (CDA)

- Understanding visualizations
  - Evaluating user performance (UP)
  - Evaluating user experience (UE)
  - Evaluating visualization algorithms (VA)

Example Scenario 1: UWP

- Understanding environments and work practices
- Goals and outputs
  - Goals: understand the work, analysis, or info processing practices by given users with or without software in use
  - Outputs: Design implications based on a holistic understanding of workflows and practices, the conditions of working environments, and potentially current tools in use
- Questions
  - Context, daily activities, tasks supported, user characteristics, data used, challenges and user barriers, etc.
- Methods
  - Field observation, interviews, laboratory observation

Common Quantitative Method Logistics

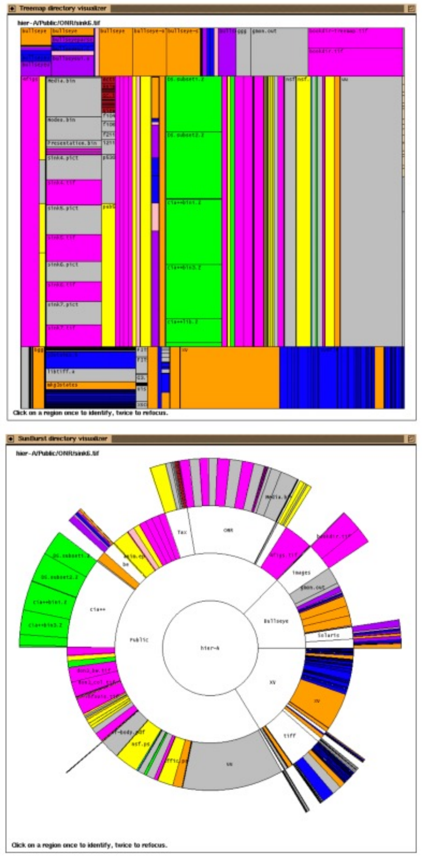
- Visualization with test data loaded
- Consent form (if required)
- Task list
- Protocol (study procedures and debrief questions)
- Surveys/interviews and any additional data-collection instruments e.g., mouse tracking/navigation tracking
- Audio or video recorded, notepad
  - Video of users
  - Screen capture of user actions
  - Audio of entire session

Common Visualization-Supported Tasks

- Basic Insights:
  - Read a value
  - Identify extrema
  - Characterize distribution
  - Describe correlation

- Comparative insights:
  - Compare values
  - Compare extrema
  - Compare distribution
  - Compare correlation

Example of Controlled Experiment (1)



A

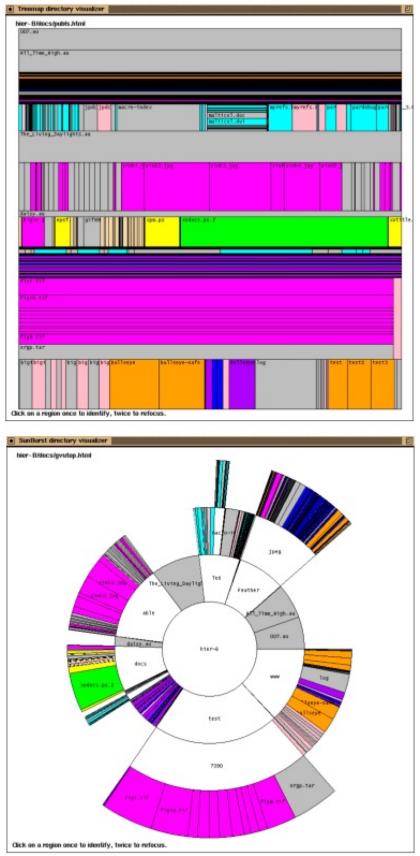
Four Conditions

Small Hierarchy
(~500 files)

A B

Large Hierarchy
(~3000 files)

A B



B

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. International journal of human-computer studies, 53(5), 663-694.

上海科技大学 ShanghaiTech University

Example of Controlled Experiment (1)

- Treemap vs. Sunburst (space-filling hierarchical views)
- Tasks
  - Perform typical file/directory-related tasks
    - Find largest file/directory
    - Locate a file
    - Identify a directory of certain depth or contain certain file
    - Compare file and directory
- Evaluation
  - 60 participants, balanced condition orders
  - Task performance on both correctness and time
  - Notes, interview responses, survey responses, etc.
- Hypotheses
  - Treemap (TM) better for comparing file size
  - Sunburst (SB) better for searching & understanding structure

Example of Controlled Experiment (1)

Results

Small Hierarchy

Hierarchy A

Tool	Phase	Correct	
TM ($n = 8$)	1	9.88	(3.23)
SB ($n = 8$)	1	12.88	(1.96)
TM ($n = 8$)	2	12.25	(1.75)
SB ($n = 8$)	2	12.63	(2.00)
TM (collapsed across phase)		11.06	(2.79)
SB (collapsed across phase)		12.75	(1.91)

Hierarchy B

Tool	Phase	Correct	
TM ($n = 8$)	1	11.50	(2.14)
SB ($n = 8$)	1	10.38	(1.69)
TM ($n = 8$)	2	10.75	(2.77)
SB ($n = 8$)	2	11.50	(2.00)
TM (collapsed across phase)		11.13	(2.42)
SB (collapsed across phase)		10.94	(1.88)

Large Hierarchy

Hierarchy A

Tool	Phase	Correct	
TM ($n = 8$)	1	8.71	(1.60)
SB ($n = 8$)	1	11.43	(1.27)
TM ($n = 8$)	2	11.57	(1.27)
SB ($n = 8$)	2	11.00	(2.16)
TM (collapsed across phase)		10.14	(2.03)
SB (collapsed across phase)		11.21	(1.72)

Hierarchy B

Tool	Phase	Correct	
TM ($n = 8$)	1	8.29	(2.14)
SB ($n = 8$)	1	11.14	(2.67)
TM ($n = 8$)	2	10.86	(1.57)
SB ($n = 8$)	2	11.00	(2.00)
TM (collapsed across phase)		9.57	(2.24)
SB (collapsed across phase)		11.07	(2.27)

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. International journal of human-computer studies, 53(5), 663-694.

Example of Controlled Experiment (1)

Performance was relatively mixed, trends favored Sunburst, but not clear-cut

Possible order effect

Most users preferred Sunburst (51/60)

TM seemed better for size-related tasks

SB seemed better for determining directory relations

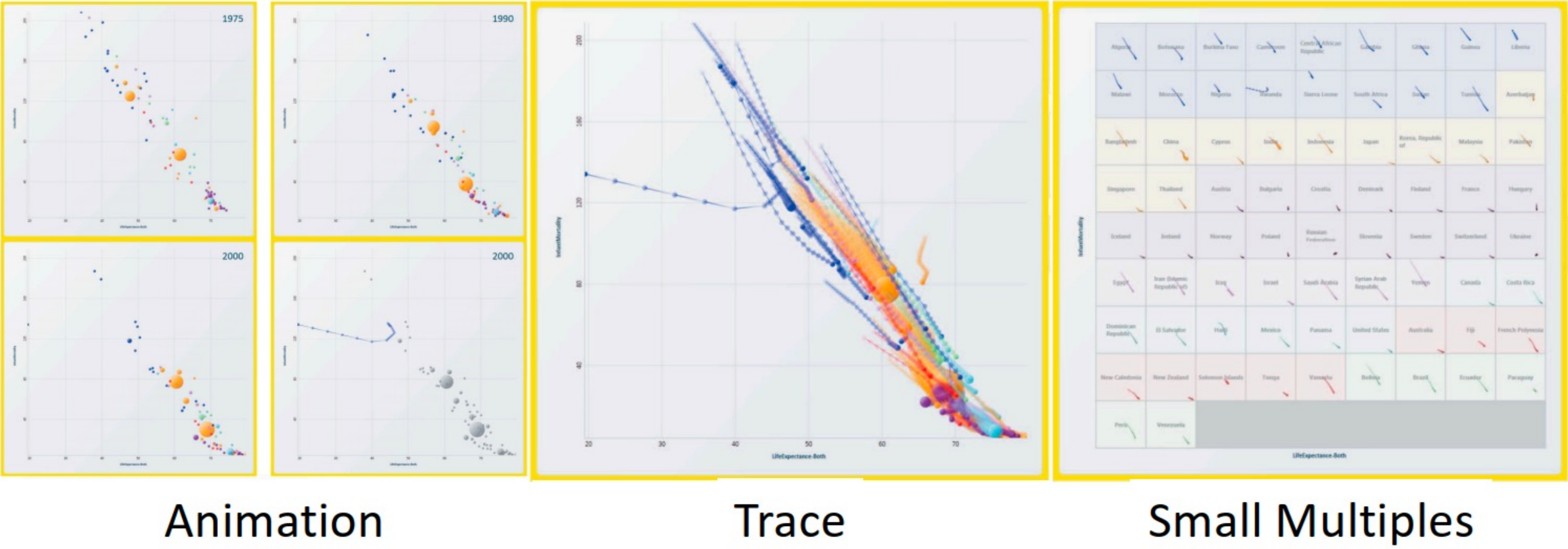
Users' search strategies mattered

Jump out to total view, start looking

Go level by level

Example of Controlled Experiment (2)

- Are animated bubble charts (a la Rosling and GapMinder) beneficial for analysis and presentation?



Conditions

Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of animation in trend visualization. IEEE transactions on visualization and computer graphics, 14(6), 1325-1332.

Example of Controlled Experiment (2)

- Experiment Design
  - 3 (animation types) * 2 (data size: small & large) * 2 (presentation vs. analysis)
    - Presentation vs analysis: between subjects
    - Others: within subjects
- Data: UN data about countries
- Tasks:
  - Analysis: select countries based on given criteria (24)
  - Presentation: describe a trend relevant to the tasks
- Measurements:
  - Quantitative: accuracy and speed
  - Qualitative: perception and preference

Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of animation in trend visualization. IEEE transactions on visualization and computer graphics, 14(6), 1325-1332.

Example of Controlled Experiment (2)

上海科技大学 ShanghaiTech University

- Accuracy: Small Multiples >> Animation
- Speed:
  - Analysis: Animation << small multiple & traces
  - Presentation: Animation >> small multiples & traces
- Perception:
  - Animation more fun, small multiples more effective
  - As data grows, accuracy drops (size & visual clutter)
- Preference

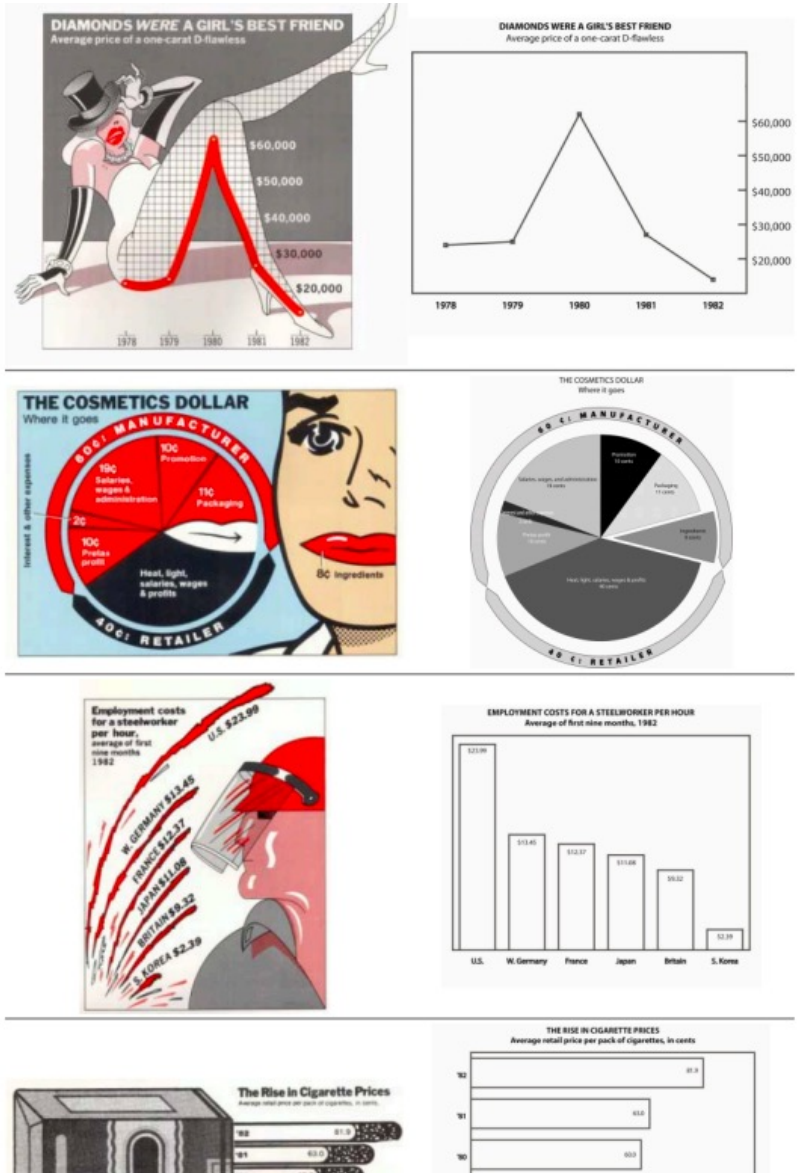
Presentation, small:	Animation (9)	>	SM (6)	>	Traces (3)
Presentation, large:	Traces (8)	>	SM (6)	>	Animation (4)
Analysis, small:	Animation (7)	>	SM (6)	>	Traces (5)
Analysis, large:	Animation (8)	>	SM (6)	>	Traces (4)

Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of animation in trend visualization. IEEE transactions on visualization and computer graphics, 14(6), 1325-1332.

立志成才报国裕民

Example of Controlled Experiment (3)

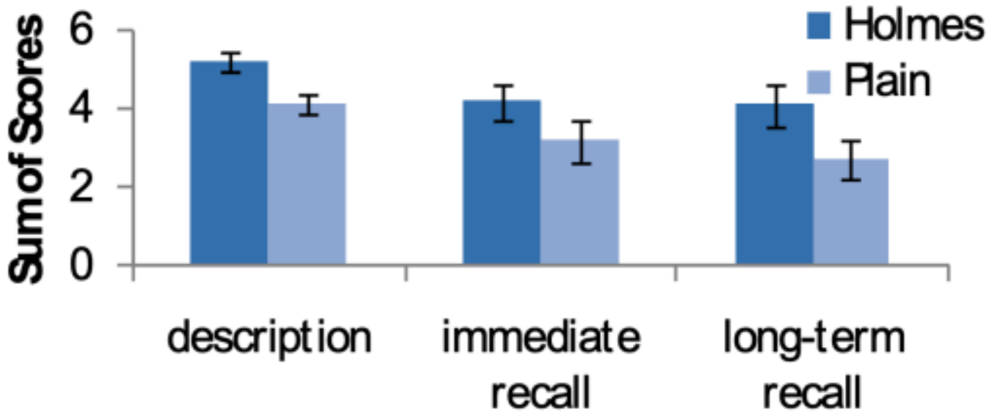
- Is chart junk no good?
- Experiment Design
  - Asked immediate interpretation accuracy questions
  - Asked similar questions again 5 minutes or 2-3 weeks later
- Twenty Participants
- Tasks: Answer Questions
  - Subject, value, trend, message
- Measurements:
  - Answer score, preference, gaze



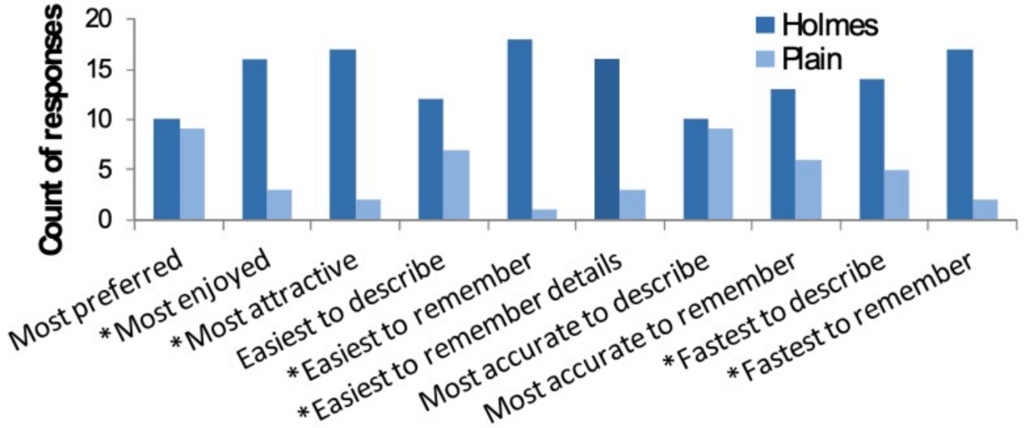
Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010, April). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2573-2582)

Example of Controlled Experiment (3)

- No significant difference in interpretation accuracy
- Better recall for chart junk graphs



- Participants found the chart junk graphs more attractive, enjoyable, and easy and fast to remember



Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010, April). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2573-2582)

Example of Controlled Experiment (3)

上海科技大学 ShanghaiTech University

- Limitation of Findings
  - Small datasets
  - “Normal” charts were really plain
  - No interaction
  - How about other added interpretations from the flowery visuals?

- Be careful reading too much into this

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010, April). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2573-2582)

Example of Controlled Experiment (4)

- Scenario: Plan learning path on online question pool

- PeerLens System
  - Find peers for a specific learning scenario
  - Compare with peers' performance
  - Offer flexible learning path suggestions

- Consider investigative analysis tasks
  - Sensemaking, awareness, and understanding

Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., ... & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12)

PeerLens: Peer-inspired Interactive Learning Path Planning in Online Question Pool

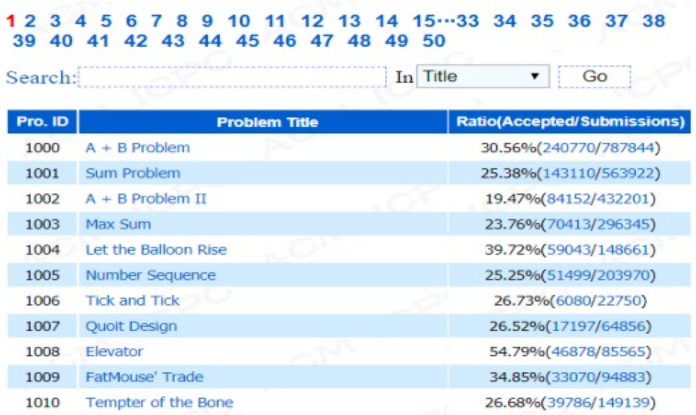
Meng Xia, Mingfei Sun, Huan Wei, Qing Chen,

Yong Wang, Lei Shi, Huamin Qu, Xiaojuan Ma

香港科技大學
THE HONG KONG
UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Example of Controlled Experiment (4)

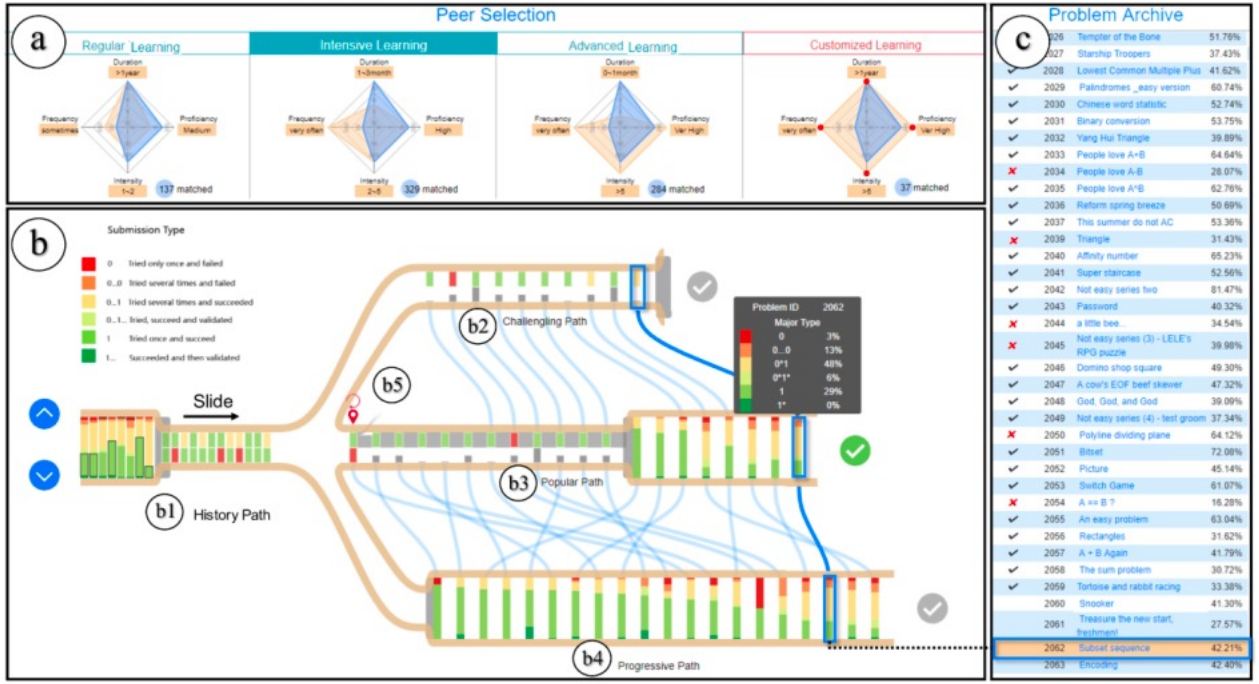
Baseline system



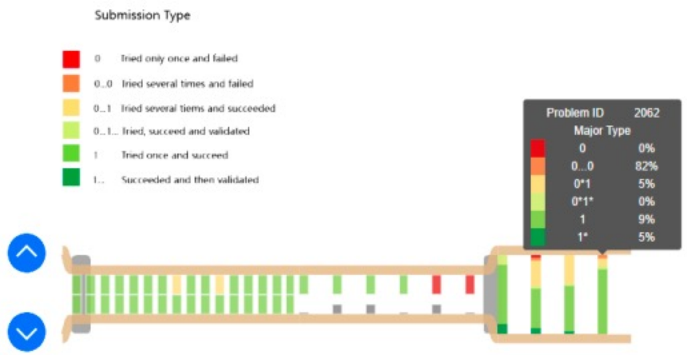
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15…33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

Search: In Title Go

Pro. ID	Problem Title	Ratio(Accepted/Submissions)
1000	A + B Problem	30.56%(240770/787844)
1001	Sum Problem	25.38%(143110/563922)
1002	A + B Problem II	19.47%(84152/432201)
1003	Max Sum	23.76%(70413/296345)
1004	Let the Balloon Rise	39.72%(59043/148661)
1005	Number Sequence	25.25%(51499/203970)
1006	Tick and Tick	26.73%(6080/22750)
1007	Quoit Design	26.52%(17197/64856)
1008	Elevator	54.79%(46878/85565)
1009	FatMouse' Trade	34.85%(33070/94883)
1010	Tempter of the Bone	26.68%(39786/149139)



Primitive PeerLens



Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., ... & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12).

Example of Controlled Experiment (4)

- Dataset: A popular programming question pool
  - ~4.6M submission records ~54K learners ~5K programming questions
- Participants: 18 CS students; within-subjects
- Conditions: Full PeerLens, Primitive PeerLens, Baseline
- Tasks:
  - Determine the starting question given a learning scenario
  - Find the next question given a historical learning path
- Learning scenarios:
  - Basic programming practice
  - Coding qualification test for IT company interviews
  - International Programming Contest

Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., … & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12)

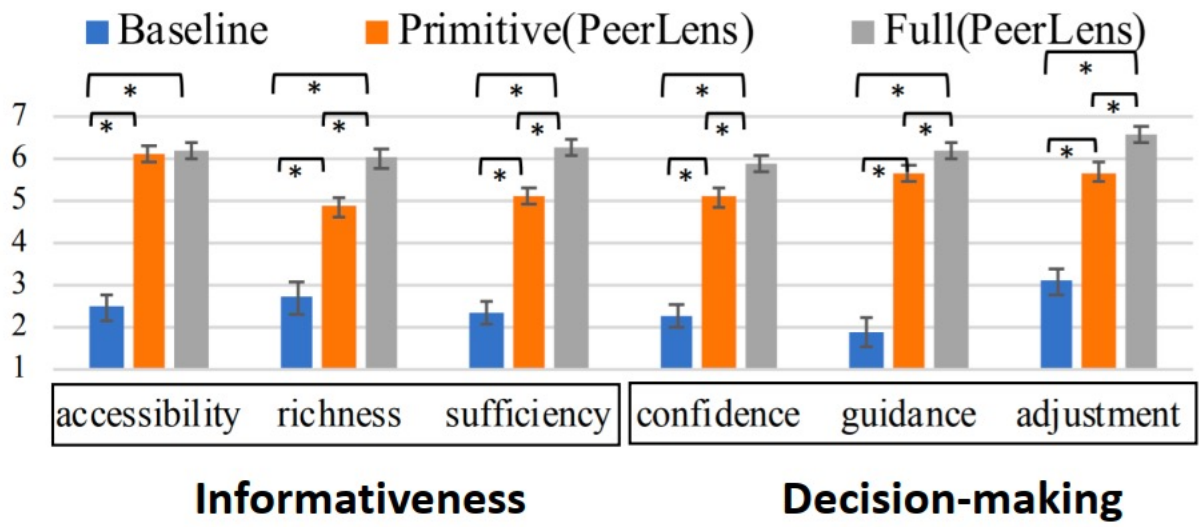
Example of Controlled Experiment (4)

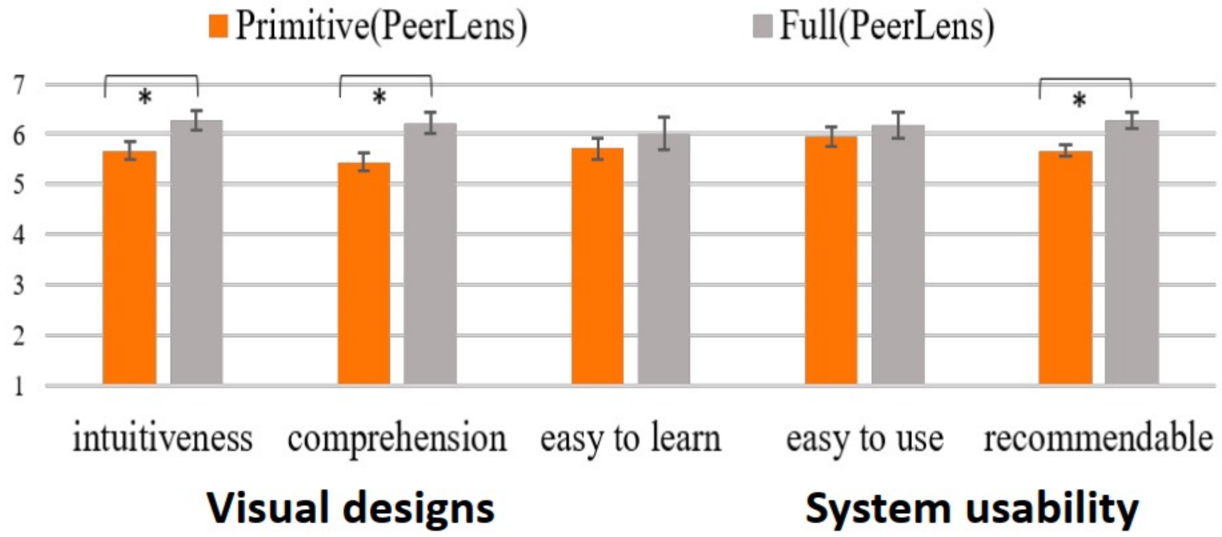
Informativeness	Q1	The information needed to plan a learning path is easy to access.
	Q2	The information needed to plan a learning path is rich.
	Q3	The information is sufficient to plan a learning path.
Decision making	Q4	The system was helpful for me to find a proper learning path for a specific learning scenario.
	Q5	I am confident that I find a suitable learning path for the learning scenario.
	Q6	The system helps make adjustment according to previous performance.
Visual design	Q7	The learning path design is intuitive.
	Q8	The learning path design helps me understand the suggested path.
System Usability	Q9	It was easy to learn the system.
	Q10	It was easy to use the system.
	Q11	I would like to recommend this system to others.

Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., … & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12)

Example of Controlled Experiment (4)

上海科技大学 ShanghaiTech University





Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., ... & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12)

立志成才报国裕民

Example of Controlled Experiment (4)

- The educational recommendation system should provide richer information with more options
- Visual design for presenting information is more important than truncating the data
- Intuitive explanations help reduce information load

Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., ... & Ma, X. (2019, May). Peerlens: Peer-inspired interactive learning path planning in online question pool. In Proceedings of the 2019 CHI conference on human factors in computing systems (pp. 1-12)

Challenges of Quantitative Evaluation

Table 1. Type I and Type II Errors

		Reality	
		H_0 TRUE	H_0 FALSE
Experimental decision	H_0 TRUE	ok	Type II
	H_0 FALSE	Type I	ok

- Conclusive validity:
  - Is there a relationship?
- Internal validity:
  - Is the relationship causal?
- Construct validity:
  - Can we generalize to the constructs (ideas) the study is based on?
- External validity:
  - Can we generalize the study results to other people/places/times?
- Ecological validity:
  - Does the experimental situation reflect the type of environment in which the results will be applied

Types of Qualitative Methods

- Nested Qualitative Methods
  - Experimenter observations
  - Think-aloud protocol
  - Collecting participant opinions
- Inspection evaluation methods
  - Usability Heuristics
  - Collaboration Heuristics
- Information Visualization Heuristics
  - In situ observational studies
  - Participatory observation
  - Laboratory observational studies
  - Contextual interviews

Challenges for Qualitative Evaluation

- Sample Sizes
  - Often lower than required for quantitative research
  - Often determined during the study
- Subjectivity
  - Reliability and validity
  - Transferability rather than generalizability
- Analyzing qualitative data
  - Qualitative, quantitative, or a combination of both methods
  - Qualitative analysis methods:
    - Data-driven, motivated from previous research
    - Theory-driven, each with respectively decreasing levels of sensitivity to the data

Summary

- Why do evaluation of information systems?
  - We need to be sure that new techniques are really better than old ones
  - We need to know the strengths and weakness of each tool; know when to use which tool
- Challenge
  - There are no standard benchmark tests or methodologies to help guide researchers
    - Moreover, there's simply no one correct way to evaluate
  - Defining the tasks is crucial
    - Would be nice to have a good task taxonomy
    - Data sets used might influence results
  - What about individual differences?
    - Can you measure abilities (cognitive, visual, etc.) of participants?

Conclusion

- Insight is important
  - Great idea, but difficult to measure
  - Design and analyze visualization techniques in context of real-world use
- Utility is a real key
  - Usability matters, but some powerful systems may be difficult to learn and use
  - Time/error analyses can be insightful, but they do not provide a complete picture
  - Performance measures may be more suited to serious analysis than casual use
- Exploration
  - Visualization most useful in exploratory scenarios when you don't know what task or goal (so how to measure that?)

Questions?
Thank you ☺

Quan Li

liquan@shanghaitech.edu.cn