Introduction to GeoVisualization
{and Visual Analytics}

Special focus: spatio-temporal and movement data

Gennady Andrienko
Natalia Andrienko

http://geoanalytics.net
{Geo} Information Visualization, basic principles

- Interactive Maps and Multiple Coordinated Views
Maps: not only for orientation!

People live in geographical space. Most of people’s decisions and actions depend on
- where the things are;
- how are their locations related.

Maps allow people to perceive the space beyond the directly observable extent.

A map serves as a model of reality and helps to detect patterns existing in the reality
Example: Dr. John Snow’s discovery

Map of locations of deaths from cholera
London, September 1854
Interactive maps

Interactive maps can change in response to user’s actions

Many interactive maps are available on the Web, e.g. street maps, tourist maps, election maps, ...

Interaction techniques are used to

- compensate for the display deficiencies, e.g. limited size (zoom and pan, showing additional information related to mouse position, …)
- increase the display expressiveness
- enable more sophisticated analyses
Typical interactive operations

Select information layers, e.g. on a tourist map: accommodation, museums, restaurants, nightlife

Select time moments or intervals in displays of time-related information, e.g. election year

Change the spatial scale, e.g. states or counties

Change the theme, e.g. president elections or governor elections, absolute values or differences in comparison to the previous time

Choose the visualization method, e.g. area painting or proportional symbols
Examples of analytical interactions
Removing Outliers (1)

**Outlier**: a very high (or very low) value, far apart from others.

**Interactive outlier removal**

**Effect of outlier removal**: maximal darkness corresponds now to a lower value.

Maximal darkness corresponds now to a lower value.
Removing Outliers (2)

After the removal of two outliers, the differences are better seen.
Object Comparison

The diverging colour scale allows us to compare an object with all others:

Click!

After clicking on Beja:
- lower values than in Beja → blue
- higher values than in Beja → brown
Pattern Investigation (1)

click here to transform the colour scale from sequential to diverging

move the slider and observe how the map changes
Pattern Investigation (2)

By moving the slider, we see more patterns and gain more understanding of value distribution.

- West-to-east increase
- Clusters of low values around Porto and Lisboa
- One more cluster of low values
- Coast-inland contrast
- Clusters of high values in central-east
Focusing and Visual Comparison on Other Map Types

Outlier

Maximum represented value

Value to compare with

Minimum value
Piechart Map

Applicable to several attributes that together give some meaningful whole

“Pie” size is proportional to the total (sum of the attribute values)

The division into slices shows proportion of each attribute in the total

Here the population is very small in comparison to the large cities. Therefore, the pies are too small to be seen

However, the map often looks like this:
Piechart Map: Focusing

Move this delimiter to the left

The largest pies are gradually removed (replaced by hollow circles)

The remaining pies become larger

Now the maximum pie size corresponds to this value
In districts with much population people work in industry (magenta) and services (cyan).

Northwest: more industry

Centre-west: more services

At this stage, the agricultural part (green) becomes visible

In districts with little population considerable proportion of people works in agriculture, but services still prevail.
Why to Use Multiple Views?

**Dot plot**: Distribution of attribute values within a range

**Stacked dot plot**: Distribution of attribute values within a range

**Frequency histogram**: Distribution of attribute values within a range

**Scatter plot**: shows how two attributes are related

**Parallel coordinates**: object characteristics profiles; relationships between attributes (look at line slopes)
Display Linking by Highlighting

An object pointed on the map with the mouse is simultaneously highlighted here, and here, but not here: this is an aggregated view that does not show individual objects.
Display Linking by Selection

Selection (durable highlighting) does not disappear after the mouse is moved away. One or more objects may be selected e.g. by mouse-clicking on them.

We have clicked on each of these 3 objects

These black lines correspond to the selected objects

These black dots correspond to the selected objects
Let us examine characteristics of districts in this area.

The characteristics in terms of the upper 4 attributes are rather coherent.

Two distinct clusters in the value space of these two attributes.

Enclose the area in a frame.

The values of this attribute are split in two groups with a gap between.

The values of this attribute greatly vary.

The districts fit in the left half of the histogram, mostly in bars 1 and 4.
Let us look at the districts with high % employed in industry:

The districts form 3 spatial clusters.

The districts have average or high proportions of children and young people.

Low proportions of agricultural workers and people without primary school education.

Population change: mostly between -0.1% and 12.4%.
Using Display Linking (3)

Let us look at the districts with the highest population growth:

Click on the rightmost bars in the histogram

The districts form some spatial clusters

The districts have average or high proportions of children and young and low proportions of old people

The proportions of agricultural workers and people without primary school education are mostly low, but there is an outlier
Dynamic Query

Dynamic query allows us to set constraints on attribute values

Limits the maximum value

The maximum limit can be also explicitly given

Limits the minimum value

The minimum limit can be also explicitly given

Statistics of constraint satisfaction
Dynamic Query in Action (1)

Query condition: % 0-14 years must be below 20

Query result: the objects that do not satisfy the condition has been removed from all displays
Dynamic Query in Action (2)

A second query condition added: % 25-64 years must be between 45 and 65

177 objects (64%) satisfy the 1st condition
228 objects (83%) satisfy the 2nd condition
163 objects (59%) satisfy both conditions

Query result changed: the objects that do not satisfy both conditions has been removed from all displays
Dynamic Query in Action (3)

One more query condition added: % pop. change from 1981 to 1991 must be positive

... 98 objects (36%) satisfy the 3rd condition
44 objects (16%) satisfy all 3 conditions

Now the objects that do not satisfy all 3 conditions has been removed from all displays
Propagation of Object Classes

Object classes (e.g. defined on a map) can be propagated to other displays.
Propagation of Object Classes: Use Example

Lower class prevails among districts with low % unemployed, average or high % 0-14 years and % 15-24 years, and low % 25-64 years.

Upper class prevails where % 25-64 years is high and % 65 or more years is low. It occupies mostly the middle part of the axis % 15-24 years.

Characteristics of the middle class are highly variant.

Upper class co-occurs with low % employed in agriculture and average to high % employed in services.

Lower class co-occurs with low % employed in services.

In districts with high % employed in agriculture the proportion of people with high school education is low (mostly below 3.5%). For % employed in services we see the opposite relationship.
Table View and Table Lens (1)

Table cell shading shows the relative position of the values between the minimum and maximum values of the respective attributes.
The same information can be represented in a “condensed” form. We do not see the details about particular objects but get an overall impression about value variation and relationships between attributes.

High proportions of people having high school education often co-occur with high population density.

Surprisingly, the districts with the lowest proportions of people having high school education in 1991 had much higher proportion of such people in 1981.
Class Propagation to Table View (1)

The table rows are grouped by the classes and sorted within each group.

These linked views show us, for example, that the general educational level tends to be higher in districts with high proportion of people employed in services.
Class Propagation to Table View (2)

It may also be useful to switch the grouping off.

We see that the red rows occur mostly at the top of the table and blue ones at the bottom. Note that the rows are sorted according to % people with high school education in 1991.
Summary

This lecture was supposed to

- introduce the concept of analytical interactive maps
- stress the importance of exploring various aspects of data using multiple views
- demonstrate some types of non-cartographic displays useful in analysis of geodata
- demonstrate various techniques of display linking
- show how to use this in data analysis
See also

- Natalia and Gennady Andrienko

  *Exploratory Analysis of Spatial and Temporal Data*
  A Systematic Approach
  Springer-Verlag, December 2005

  section 4.8, pp.428-449
Data structure

So far:

• Id, x, y, attribute(s)

Let’s add one more special component: time

• Id, x, y, t, attribute(s)
Overview

Types of Analysis Tasks on Spatio-Temporal Data

Three primary task foci (target information)
- Focus on objects
- Focus on space
- Focus on time

Two types of task subject
- Characteristics
- Relations

Two levels of analysis
- Elementary: focus on one or more elements of a set
- Synoptic: focus on a set as a whole, disregard individual elements
- Task may be elementary w.r.t. one subject and synoptic w.r.t. another
Data Types and Transformations

Methods & Techniques for Different Spatio-Temporal Data

- Spatial time series
- Events
- Trajectories
Data Types and Transformations

Methods & Techniques for Different Spatio-Temporal Data

Representation forms offer different conceptual views on ST data
Convert data into a form suitable for the task
Spatial Time Series

Data structure

Spatial references: states of the USA

Temporal references: years from 1960 till 2000 (41)

Attributes: population + various crime rates
High level analysis questions

Spatial Time Series (STS)

How are attribute values distributed over the territory at a given time moment?
How do the attribute values at a given place vary over time?
How does the overall spatial pattern of value distribution evolve over time?
How are different behaviour patterns distributed over the territory? Are there spatial clusters of similar behaviours?
Spatial Time Series

Methods & Techniques for Different Spatio-Temporal Data I

Visualization methods

- Animated maps
- "Layman techniques": (animated) charts embedded in maps (bar charts, pie charts, …)
- “Small multiples” map displays
- Time Graphs and their transformations
Spatial Time Series

Methods & Techniques for Different Spatio-Temporal Data I

Visualization methods

■ Animated maps

■ “Small multiples” map displays

■ „Layman techniques”: (animated) charts embedded in maps (bar charts, pie charts, …)

■ Time Graphs and their transformations
Time Map

Spatial Time Series: Basic Visualization Methods I

Time-dependent data may be represented on a time map, which is manipulated through time controls and, in particular, allows animation.

For a time map, one can use any representation method suitable for static data. Choropleth maps are good for exploring spatial distribution patterns.
When time map is useful

Spatial Time Series: Basic Visualization Methods I

How did the spatial distribution pattern develop over time?
At moment $t$, how were the values distributed over the whole space?
At what time moment were the values distributed over the space in the given manner?
To compare the spatial distributions of attribute values at two or more time moments, we need to see these distributions simultaneously. Best of all is to use multiple maps displayed in a common panel and manipulated through a common set of controls.
Exploring the Distribution of Changes

Spatial Time Series: Data Transformations

Instead of original attribute values, a time map or map series can represent changes, that is, differences or ratios to the previous moment or to any selected moment.

Here the maps correspond to years 1990, 1991, and 1992 and represent differences to the previous years. Positive differences (i.e. increased values) are shown in brown and negative differences (i.e. decreased values) in blue.
Temporal smoothing and computing of residuals

For each time moment computes differences or ratios to a particular object or value or to the mean or median among all objects at this moment

Value scale transformation, e.g. logarithmic
Comparison to country’s median

Spatial Time Series: Data Transformations

Build map series with transformed data:
relative difference to median value

Spatial distribution patterns of attribute values may become more vivid
Value and patterns evolution over time can be seen more easily
Spatial Time Series

Methods & Techniques for Different Spatio-Temporal Data I

Visualization methods

- Animated maps
- „Layman techniques”: (animated) charts embedded in maps (bar charts, pie charts, …)
- “Small multiples” map displays
- Time Graphs and their transformations
The value flow symbols show us the evolution of attribute values (*temporal behavior*) at each location.

Unfortunately, symbol overlapping creates significant inconveniences, and zooming does not always help.
Value Flow Map (2)

Spatial Time Series: Basic Visualization Methods II

Temporal smoothing allows to disregard small fluctuations and see overall trends.

Here the values for each year have been replaced by 5-year means. You can compare to the previous variant and see the effect of the smoothing.
This appears to be a spatial cluster of similar behaviors.

Around the Great Lakes, the theft rates were high, but tended to decrease in recent years.

The theft rates along the western coast are, in general, higher than inland.

There are also some unusual behaviors ("behavioral outliers").

The theft rates are relatively moderate, but tend to grow.
Temporal Behavior Exploration

Spatial Time Series: Data Transformations

As with time maps, various transformations can be applied to value flow maps.

Here: comparison to each country’s mean:

- For every time moment, each state’s values is replaced by its difference to the country’s overall mean value at that moment.
- Yellow color corresponds to positive differences, and blue – negative.
Value Flow Map Disadvantages

Spatial Time Series: Basic Visualization Methods II

The diagrams are perceived as separate entities → the map must be scanned and cannot be grasped as a single image

Absence of ordering complicates seeking for specific behaviour patterns

Diagram overlapping is a serious problem
Spatial Time Series

Methods & Techniques for Different Spatio-Temporal Data I

Visualization methods

- Animated maps
- „Layman techniques”: (animated) charts embedded in maps (bar charts, pie charts, …)
- “Small multiples” map displays
- Time Graphs and their transformations
Comparison to country’s median

Spatial Time Series: Data Transformations
Comparison to country’s median

Spatial Time Series: Data Transformations

A constant cluster of low values (since 1965)
Comparison to country’s median

Spatial Time Series: Data Transformations

A “belt” of high values NW-SE

The shape perseveres till 1970
Comparison to country’s median

Spatial Time Series: Data Transformations

Utah falls out of the “belt” in 1975 and keeps lower values till 2000
Comparison to country’s median

Spatial Time Series: Data Transformations

New Mexico occasionally destroys the “belt” in 1980
Comparison to country’s median

Spatial Time Series: Data Transformations

The “belt” spreads to the east (starting from 1987)
Comparison to country’s median

Spatial Time Series: Data Transformations

The original NW-SE “belt” transforms into a cluster of high values on the south.

Comparison to country’s median

Spatial Time Series: Data Transformations
Spatial Time Series

Methods & Techniques for Different Spatio-Temporal Data I

Visualization methods

- Animated maps
- „Layman techniques”: (animated) charts embedded in maps (bar charts, pie charts, …)
- “Small multiples” map displays
- Time Graphs and their transformations
Time Graph

Spatial Time Series: Basic Visualization Methods III

Putting all behaviors together makes their comparison more convenient

- Numeric details on-demand by pointing on an object’s line

Plus, to better understand general development trends:

- “mean behavior line” connecting means of each year
- “median behavior” connecting the year medians
Time Graph: Multiple View Comparisons

Spatial Time Series: Basic Visualization Methods III

Juxtaposed time graphs are suitable for comparing trends and temporal variations of two or more attributes.

This example shows that attributes “Motor vehicle theft rate” and “Burglary rate” have quite different trends of general development.
Time Graph: Multiple View Comparisons

Spatial Time Series: Basic Visualization Methods III

Additionally employ **smoothing** (value averaging over intervals)
Mitigates small fluctuations
Exposes trends more clearly
Time Graph: Level of Detail

Spatial Time Series: Basic Visualization Methods III

Mean and median lines only give very coarse picture of the general value variation properties.

For a finer analysis, may also look at the quartiles or even smaller percentiles.
Linked Views: Map + Time Graph

Spatial Time Series: Basic Visualization Methods IV

At place $L_1$, how did the values behave over the entire time period?

Compare the temporal behavior at places $L_1$ and $L_2$
Time Graph++: Time Histogram

Spatial Time Series: Basic Visualization Methods V

Divide value range of the attribute into intervals

Choose specific color/shade for each interval

Size of colored segments encodes relative frequencies of values from corresponding interval, for each time moment

This example:
- Each bar: one year
- Shows increase of the crime rates over the country in 70’s & early 80’s
- Note two peaks in 1975 and 1981-82, followed by gradual decrease
Time Graph++: Time Histogram

Spatial Time Series: Basic Visualization Methods V

Time histograms facilitate comparison of 2 (or more) attributes

- Compare temporal trends *despite differences in value ranges*

- Here: attributes “Burglary rate” and “Motor vehicle theft rate”
Time Histogram of Changes

Spatial Time Series: Basic Visualization Methods VI

Transform attribute values into relative differences
Apply aggregation by value intervals to the transformed values
Encode values using partitioned color scale

- “Burglary rate” values, transformed into relative difference to previous year
- Three color segments: red for increase (< +5%, +5-10%, > +10%), yellow for no change (±1%), green for decrease (< -5%, -5-10%, > -10%)
An extraordinarily bad year (1974): increase in all but one states, in 46 states over 10%

1980 was not much better than 1974

High increase in most states in 1961-1964 and 1966-1970

An occasional good year (1976) among bad ones

3 good years in a row (1982-1984) but not a long-term tendency

A good tendency (prevailing decrease of the burglary rates) since 1992
See also

• Natalia and Gennady Andrienko
  **Exploratory Analysis of Spatial and Temporal Data**
  A Systematic Approach
  Springer-Verlag, December 2005

  Chapter 4