Time Series Data

CS 4460 – Intro. to Information Visualization
November 27, 2017
John Stasko

Learning Objectives

• Identify different types of temporal data
  – discrete, interval, linear, cyclic, continuous, ordinal, branching
• List potential tasks for temporal data analysis
• Familiarity with basic temporal representations
  – Line graph, stacked graph, stream graph, bubble tracks, connected scatterplot
• Familiarity with specific temporal representation techniques and systems
  – Lifelines 1-2 & EventFlow, ThemeRiver, Cluster/calendar view, MieLog, LiveRAC,
• Discuss the benefits & limitations of all the techniques
• Be able to apply learned knowledge and examples to the design of visualizations for new data and problems
Time Series Data

- Fundamental chronological component to the data set

Data Sets

- Each data case is likely an event of some kind
- One of the variables can be the date and time of the event
- Examples:
  - sunspot activity
  - baseball games
  - medicines taken
  - cities visited
  - stock prices
  - How about events with a duration?
  - Discrete vs. Interval
Data Mining

- Data mining domain has techniques for algorithmically examining time series data, looking for patterns, etc.

- Good when objective is known a priori
- But what if not?
  - Which questions should I be asking?
  - InfoVis better for that

Tasks

- What kinds of questions do people ask about time series data?
Time Series Tasks

- Is there an order that things occur in?
- How long do events last?
- Is there a cycle?
- How does an event change over time?
- How often does an event happen?

Time Series User Tasks

- Examples
  - When was something greatest/least?
  - Is there a pattern?
  - Are two series similar?
  - Do any of the series match a pattern?
  - Provide simpler, faster access to the series
Other Tasks

- Does data element exist at time $t$?
- When does a data element exist?
- How long does a data element exist?
- How often does a data element occur?
- How fast are data elements changing?
- In what order do data elements appear?
- Do data elements exist together?

Muller & Schumann '03 citing MacEachern '95

Taxonomy

- Discrete points vs. interval points
- Linear time vs. cyclic time
- Ordinal time vs. continuous time
- Ordered time vs. branching time vs. time with multiple perspectives

Muller & Schumann '03 citing Frank '98
Classical Presentation

• What is the tried and true, most common way of representing time series data?
  – Focus here is measuring some value over time

Line Graph
A look back at market history shows that the U.S. stock market, represented here by Standard & Poor’s 500 Composite Index, demonstrated strength after big declines. Even after three steep drops, the S&P 500 still provided an average 10-year annualized return of nearly 11% as of December 31, 2012. However, it’s important to note that past results aren’t predictive of the future.
Fun One

What If Everybody in Canada Flushed At Once?

What are these presenting?
- One continuous quantitative value over time (time on x, variable on y)

What if there are multiple values to track?
Multiple Lines

Proportions of Total

Alternative?
**Data?**

- What if the different values don't comprise a whole?
  - Don’t add up to 100%
**ThemeRiver**

Havre et al
InfoVis '00

![ThemeRiver Diagram]

**Streamgraph**

Byron & Wattenberg
TVCG ’08

![Streamgraph Diagram]

Design Issues

- Curve shape
  - Wiggle, symmetry, balance
  - Definitely some interesting math to do it
- Color choice
- Labeling
- Layer ordering

- Paper provides very nice discussion of this

http://researchweb.watson.ibm.com/history/

Document Edits
Brightness indicates text age
Registered authors color-coded
Anonymous authors in white

Spacing by revision #
Spacing by time

Different Data

- Nominally-typed events occurring over time with durations

- Do days/weeks/months/years matter?
  - If yes, then...
Calendar View

More Context

• How do we see more context/overview?
DateLens

**Fisheye approach**

Bederson et al  
*ACM ToCHI '04*

Alternative

**Spiral Calendar**

Mackinlay, Robertson & DeLine  
*UIST '94*
Alternative

Uses projected shadows on walls

Empty spots on back wall show good times

Time Lattice

Different Data

- Nominally-typed events occurring over time with durations

- Do days/weeks/months/years matter?
  - If no, then...

Revisit
Gantt Chart

Potential tasks:
- Put together complete story
- Garner information for decision-making
- Notice trends
- Gain an overview of the events to grasp the big picture

Lifelines Project

Visualize personal history in some domain

Plaisant et al
CHI ’96
Challenges

- Scalability (could be thousands of tests)
- Can multiple records be visualized in parallel (well)?
  - Comparisons
    - What trends do you see in the last 8 EKGs?
    - Compare the 8 people who all seem to have the same problem
- Support (reg-ex text) queries
- Support alignment, rank, and filter
- Medical application:
  - Look for temporal coincidence of two events
    - First pneumonia and asthma attack
  - Medical professionals don’t want to fool with zooming and panning

Follow-on
LifeLines2: Focus on alignment along events

Wang et al
CHI ’08
More Follow-on

EventFlow

Transform

Smart aggregations to show overviews of large collections of events

Monroe et al

*TVCG* '13

http://www.cs.umd.edu/hcil/temporalviz

HCIL Projects

http://www.cs.umd.edu/hcil/temporalviz
Data?

- What if you want to show two continuous variables over time?
  - And not just use two lines

Bubblechart Animation

Strengths?

Weaknesses?
Alternative

• How do we address weaknesses?
  – How to get rid of time slider?

Trace View

“Traces” in Gapminder-style visualization

Robertson et al
TVCG (InfoVis) ’08
Connected Scatterplot

- Showing two variables over time
  - Use standard scatterplot
  - Plot the two values at different points in time
  - Connect those points, in order, with a line
  - Label key times (e.g., years)
Janet L. Yellen, on the Economy’s Twists and Turns

Inflation and unemployment
The Federal Reserve is said to have a “dual mandate” keeping inflation in check and the unemployment rate low. These macroeconomic goals are, however, not always in line with each other, as the Federal Reserve has to balance the risks of inflation and unemployment.

In speeches and in meetings, Ms. Yellen, the nominee for the next Fed leader, has commented on the Fed’s actions during significant periods, providing a window into her views and priorities.


http://www.dundas.com/blog-post/in-praise-of-connected-scatter-plots/

Nice Article
Fundamental Tradeoff

• Is the visualization time-dependent, ie, changing over time (beyond just being interactive)?
  – Static
    Shows history, multiple perspectives, allows comparison
  – Dynamic (animation)
    Gives feel for process & changes over time, has more space to work with

InfoVis to the Rescue

• What about some more unique data sets?
• Can we come up with good individual solutions?
Case Study 1

- Understand patterns of presence/resource usage/events over time
- Show this large amount of data in an easily understandable and query-able manner
- Scenarios:
  - Workers punch in and punch out of a factory
    Want to understand the presence patterns over a calendar year
  - Power plant electricity usage over a year

Ideas

- Any ideas on what we could do here?
One Idea

![Graph showing energy consumption](image)

**Good**
- Typical daily pattern
- Seasonal trends

**Bad**
- Weekly pattern
- Details

---

**Approach Taken**

- Cluster analysis
  - Find two most similar days, make into one new composite
  - Keep repeating until some preset number left or some condition met

- How can this be visualized?
  - Ideas?
Display

Characteristics

- Unique types of days (individual or cluster) get their own color
- Contextually placed in calendar and line graph for it is shown
- Stop clustering when a threshold met or at a predetermined number of clusters
- Interactions
  - Click on day, see its graph
  - Select a day, see similar ones
  - Add/remove clusters
Insights

- Traditional office hours followed
- Most employees present in late morning
- Fewer people are present on summer Fridays
- Just a few people work holidays
- When the holidays occurred
  - School vacations occurred May 3-11, Oct 11-19, Dec 21-31
- Many people take off day after holiday
- Many people leave at 4pm on December 5
  - Special day in Netherlands, St. Nicholas’ Eve

Case Study 2

- Computer system logs
- Potentially huge amount of data
  - Tedious to examine the text
- Looking for unusual circumstances, patterns, etc.
**System View**

**What kind of display (technique)?**

**MieLog**

- **Tag area**: Block for each unique tag, with color representing frequency (blue-high, red-low)
- **Click to filter on that tag**

- **Time area**: Days, hours, & frequency histogram (grayscale, white-high)
  - Can filter by day

- **Message area**: Actual log messages (red—predefined keywords, blue—low frequency words)
  - Can filter on specific words

**Outline area**: Pixel per character, can filter on length

**Takada & Koike**

LISA '02

**Case Study 3**

- **Domain**: Computer systems management
- **Very large scale temporal log data**
  - Many processes, machines
- **Show more context of what else was going on at that time**
  - Likely have to abstract some then
  - Allow several different levels of detail at once
- **Allow drill-down for details**
Case Study 4

- How about events in time and place?
  - Many applications of this problem
GeoTime

- Represent place by 2D plane (or maybe 3D topography)
- Use 3rd dimension to encode time
- Object types:
  - Entities (people or things)
  - Locations (geospatial or conceptual)
  - Events (occurrences or discovered facts)

Example

Source: http://www.oculusinfo.com/
Example

Nice overview

Visual Methods for Analyzing Time-Oriented Data

Wolfgang Aigner, Silvia Miksch, Wolfgang Müller, Heidrun Schumann, and Christian Tominski

Abstract—Providing appropriate methods to facilitate the analysis of time-oriented data is a key issue in many application domains. In this paper, we focus on the unique role of the parameter time in the context of visually driven data analysis. We will discuss three major aspects: visualization, analysis, and the user. It will be illustrated that it is necessary to consider the characteristics of time when generating visual representations. For that purpose, we take a look at different types of time and present visual examples. Integrating visual and analytical methods has become an increasingly important issue. Therefore, we present our experience in temporal data abstraction, principal component analysis, and clustering of large volumes of time-oriented data. The third main aspect we discuss is supporting user-centered visualization. We describe event-based visualization as a promising means to adapt the visualization pipeline to needs and tasks of users.

Index Terms—Time-oriented data, visualization, analysis, user.

1 INTRODUCTION AND MOTIVATION

Considering the characteristics of data is vital when designing visual representations. A salient characteristic is whether or not data are related to time. That time is an outstanding dimension is reflected by Shneiderman's 'Three V's' data type taxonomy [1], where temporal data are identified as one of seven basic data types. Nowadays, time-oriented data are ubiquitous in many application domains, for example, in business, medicine, history, planning, or project management. For a long time, visual methods have been successfully applied to analyze such data. A wide repertoire of interactive techniques for visualizing data sets with temporal dependencies is available. However, many current

- visualization,
- analysis, and
- user.

In Section 2, we focus on visualization methods for time-oriented data. We will show that the term time-oriented data comprises several types of data with different meanings and applications. Designing or applying visual representations can only be successful if one is aware of these different types. This will be demonstrated with several examples of visualization techniques that stem from our own work or are available in the literature. Usually, time-oriented data are large—not only in terms of the number of data items but also in terms of the number
Bigger overview

http://www.timeviz.net/

Useful Widgets

http://simile-widgets.org/

Simile project
Conclusions

- Think about the data
  - What characteristics?
- Can InfoVis help?
  - Maybe not needed
- Think about the visualization techniques
  - Which technique(s) work best for your problem?
Learning Objectives

• Identify different types of temporal data
  – discrete, interval, linear, cyclic, continuous, ordinal, branching
• List potential tasks for temporal data analysis
• Familiarity with basic temporal representations
  – Line graph, stacked graph, stream graph, bubble tracks, connected scatterplot
• Familiarity with specific temporal representation techniques and systems
  – Lifelines 1-2 & EventFlow, ThemeRiver, Cluster/calendar view, MieLog, LiveRAC,
• Discuss the benefits & limitations of all the techniques
• Be able to apply learned knowledge and examples to the design of visualizations for new data and problems

Upcoming

• Visual Analytics
  – Prep: VisMaster video
• Lab: Maps and geo-data
• Review
References

• Spence and CMS books
• All referred to articles