

# Time Series Data

---



CS 7450 - Information Visualization  
April 7, 2011  
John Stasko

# Time Series Data

---



- Fundamental chronological component to the data set

75 % of 4000 samples of graphics from newspapers and magazines ('74-'80) were time-series data!

Tufte Vol. 1



# 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

Spring 2011

CS 7450

3

# Meta Level



- Consider multiple stocks being examined
- Is each stock a data case, or is a price on a particular day a case with the stock name as one of the other variables?
- Confusion between data entity and data cases

Spring 2011

CS 7450

4

# 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

Spring 2011

CS 7450

5

# Tasks



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

Spring 2011

CS 7450

6

# 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

Spring 2011

CS 7450

7

# 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?

Spring 2011

CS 7450

Muller & Schumann '03  
citing  
MacEachern '95

8

# 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

Spring 2011

CS 7450

Muller & Schumann '03  
citing  
Frank '98

9

# 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

Spring 2011

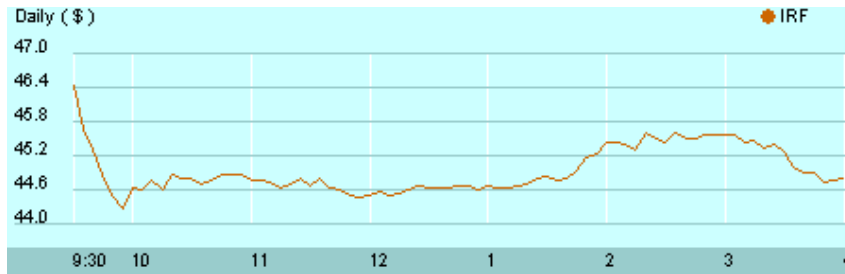
CS 7450

10

# Standard Presentation



- Present time data as a 2D line graph with time on x-axis and some other variable on y-axis

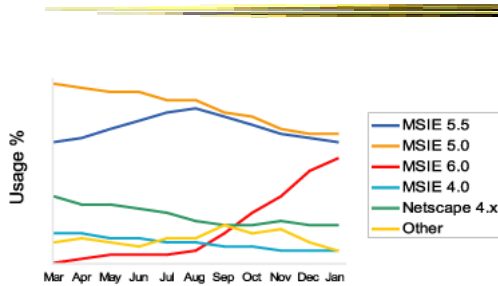


Spring 2011

CS 7450

11

# Classic Views



Spring 2011

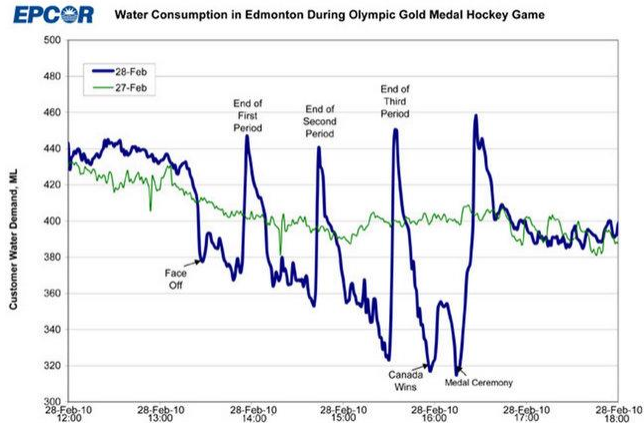
CS 7450

12

# Fun One



## What If Everybody in Canada Flushed At Once?



[http://www.patpapers.com/blog/item/what\\_if\\_everybody\\_flushed\\_at\\_once\\_Edmonton\\_water\\_gold\\_medal\\_hockey\\_game/](http://www.patpapers.com/blog/item/what_if_everybody_flushed_at_once_Edmonton_water_gold_medal_hockey_game/)

Spring 2011

CS 7450

13

## Today's Focus



- Examination of a number of case studies
- Learn from some of the different visualization ideas that have been created
- Can you generalize these techniques into classes or categories?

Spring 2011

CS 7450

14





# Tasks



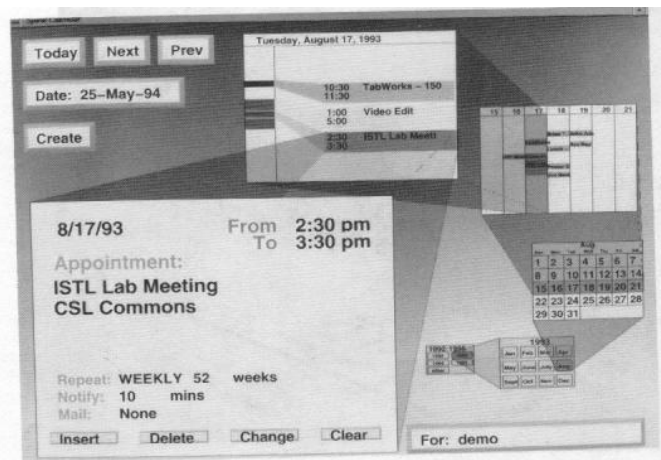
- See commonly available times for group of people
- Show both details and broader context

Spring 2011

CS 7450

17

# One Solution



Spiral Calendar

Mackinlay, Robertson & DeLine  
UIST '94

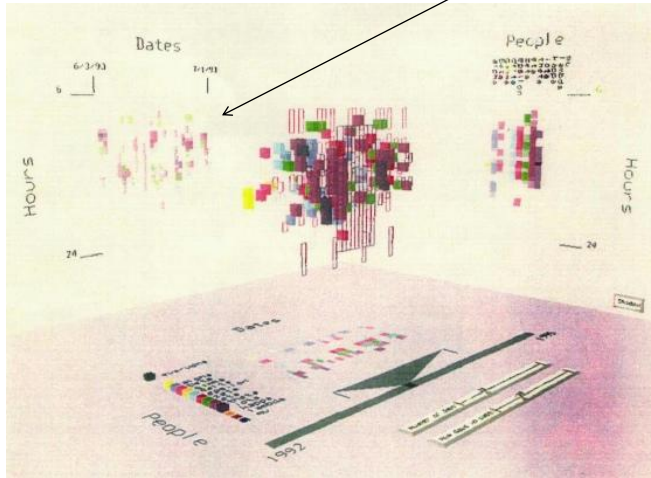
Spring 2011

CS 7450

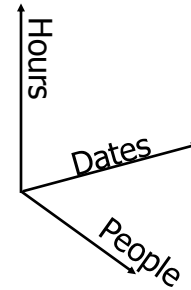
18

## Another View

Empty spots on back wall show good times



Time Lattice



Uses projected shadows on walls

Spring 2011

CS 7450

19

## Example 2

- Personal histories
  - Consider a chronological series of events in someone's life
  - Present an overview of the events
  - Examples
    - Medical history
    - Educational background
    - Criminal history

Spring 2011

CS 7450

20

# Tasks



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

Spring 2011

CS 7450

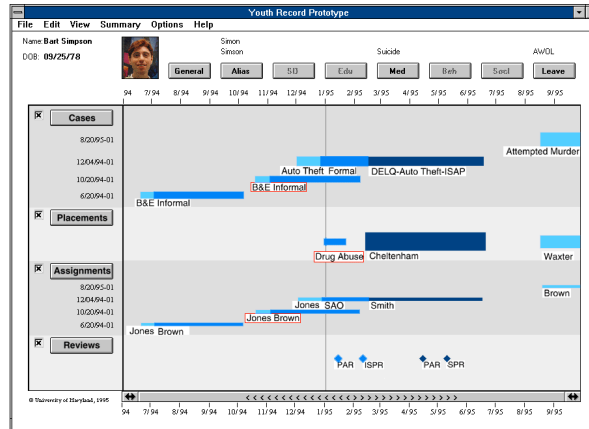
21

# Lifelines Project



Visualize personal history in some domain

Video Demo



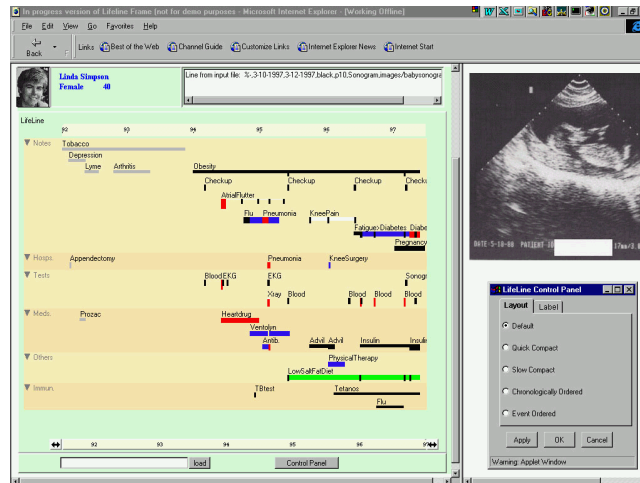
Spring 2011

CS 7450

Plaisant et al  
CHI '96

22

# Medical Display



Spring 2011

CS 7450

23

## Features



- Different colors for different event types
- Line thickness can correspond to another variable
- Interaction: Clicking on an event produces more details
- Certainly could also incorporate some Spotfire-like dynamic query capabilities

Spring 2011

CS 7450

24

## Benefits



- Reduce chances of missing information
- Facilitate spotting trends or anomalies
- Streamline access to details
- Remain simple and tailorable to various applications

## Challenges



- Scalability (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.

# New Work



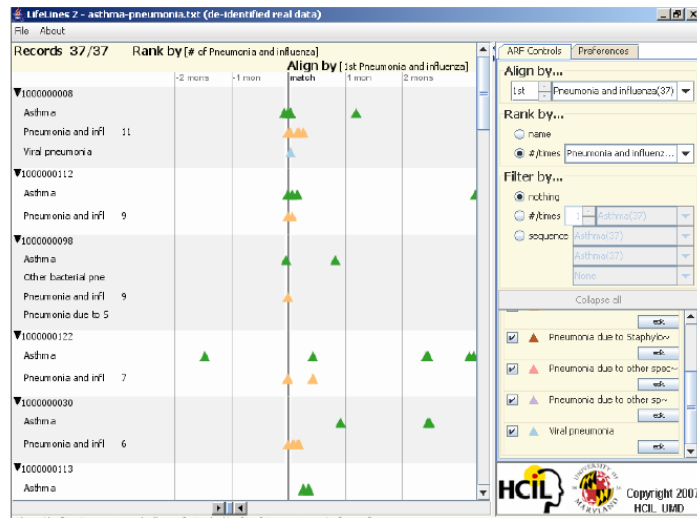
- Work with query results
- Need to align, rank, and filter
- Medical application:
  - Look for temporal coincidence of two events  
First pneumonia and asthma attack
  - Medical professionals don't want to fool with zooming and panning

Spring 2011

CS 7450

27

## LifeLines2: Focus on alignment along events



Video

Spring 2011

CS 7450

Wang et al  
CHI '08 28

## Example 3



- Understand patterns of presence/events over time
- Focus: People's presence/movements in some space
- Situation:
  - Workers punch in and punch out of a factory
  - Want to understand the presence patterns over a calendar year
- Alternate: Power plant electricity usage over a year

## Particulars



- Who is user? – Factory boss/manager
- Problem – Show this large amount of data in an easily understandable and query-able manner
- Data – Punch in/out times for workers

# Ideas



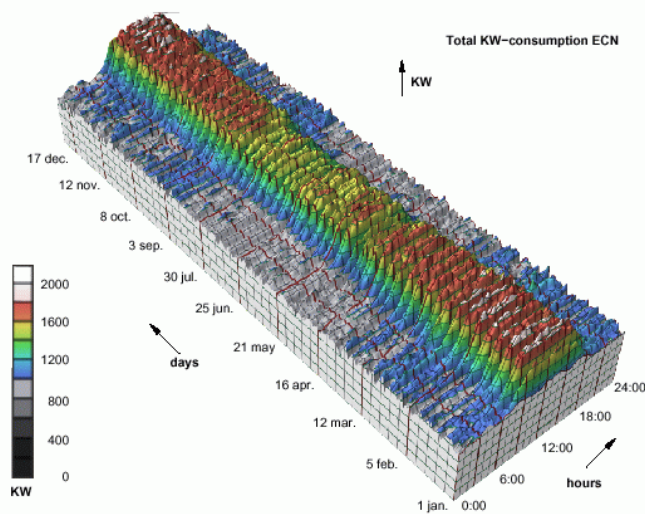
- Any ideas on what we could do here?

Spring 2011

CS 7450

31

# One Idea



Good  
Typical daily pattern  
Seasonal trends

Bad  
Weekly pattern  
Details

Spring 2011

CS 7450

32



# Approach Taken



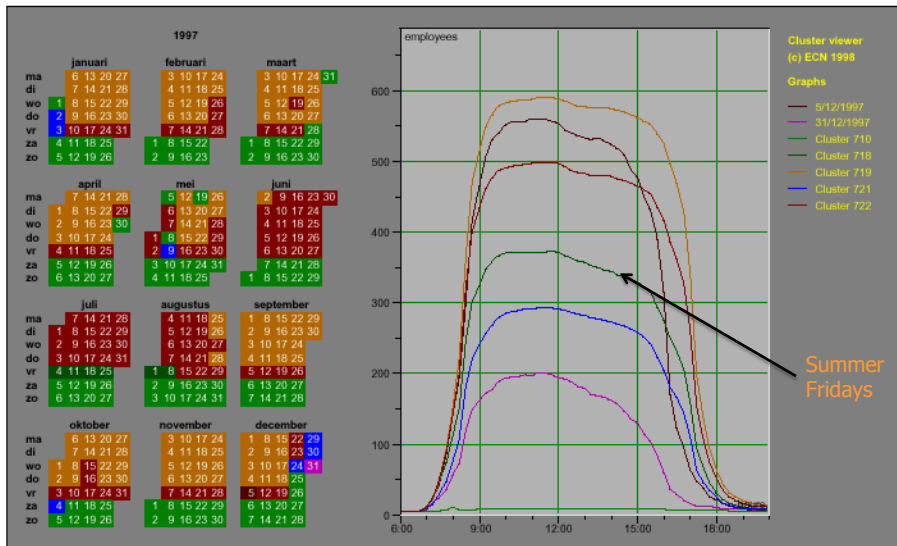
- Cluster analysis
  - Find two most similar days, make into one new composite
  - Keep repeating until some preset number left or some condition met
- How can this be visualized?
  - Ideas?

Spring 2011

CS 7450

33

# Display



Spring 2011

CS 7450

34

## 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

Spring 2011

CS 7450

35

## Interaction



- Click on day, see its graph
- Select a day, see similar ones
- Add/remove clusters

Spring 2011

CS 7450

36

# 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

Spring 2011

CS 7450

37

# Example 4



- Consider a set of speeches or documents over time
- Can you represent the flow of ideas and concepts in such a collection?

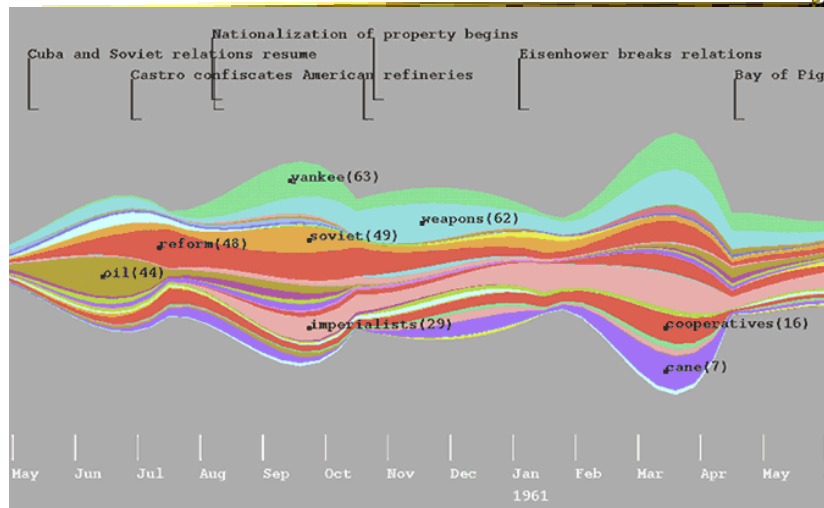
Spring 2011

CS 7450

38

# ThemeRiver

Havre et al  
InfoVis '00



Spring 2011

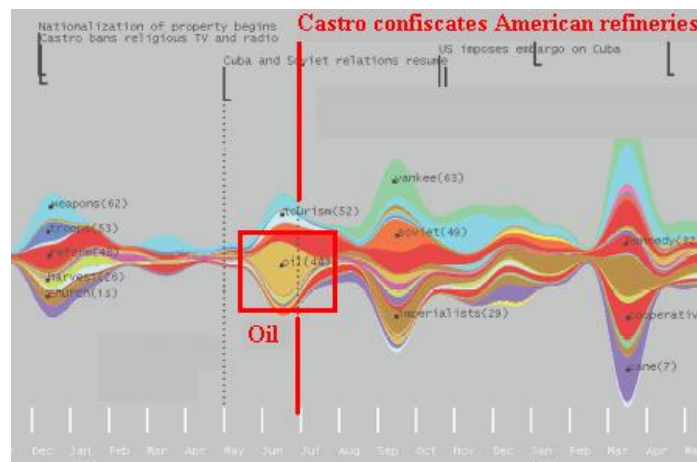
CS 7450

We saw earlier

39

# Mapping

- River height (thickness) encodes relative frequency of themes
- Key events overlaid



Spring 2011

## Example 5



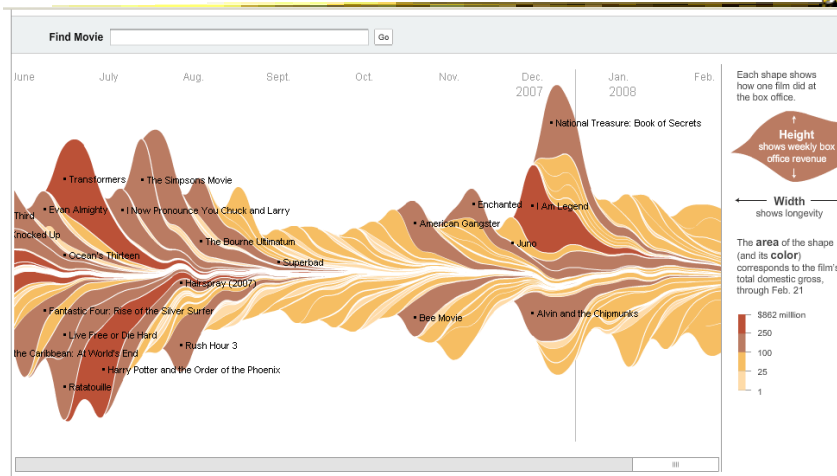
- Similar idea – Stacked graph
- Created new technique called Streamgraph
- Goals:
  - Show multiple time series
  - Be able to see sum
  - Make labels legible
  - Be able to distinguish different layers
  - Make it aesthetically pleasing

Spring 2011

CS 7450

41

## Interactive Application



[http://www.nytimes.com/interactive/2008/02/23/movies/20080223\\_REVENUE\\_GRAPHIC.html](http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html)

Spring 2011

CS 7450

42

# 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

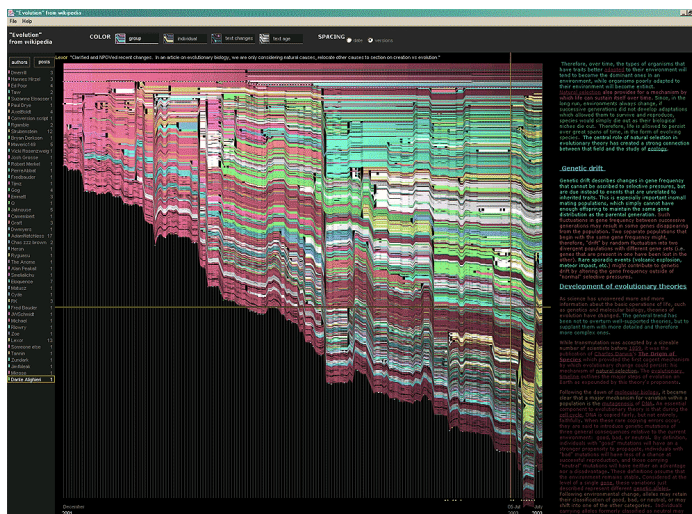
Spring 2011

CS 7450

43

## Example 6

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



Flow of changes across electronic documents

Spring 2011

CS 7450

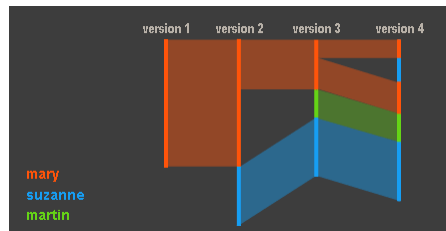
44

# Technique



Length – how much text

Time



Make connections

Spring 2011

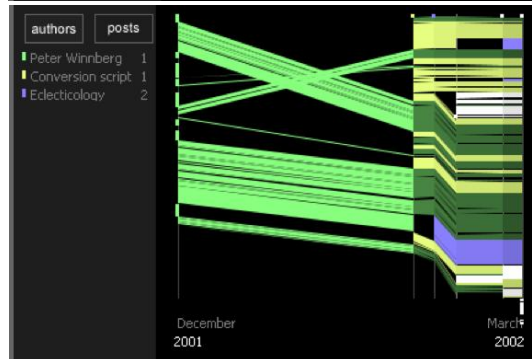
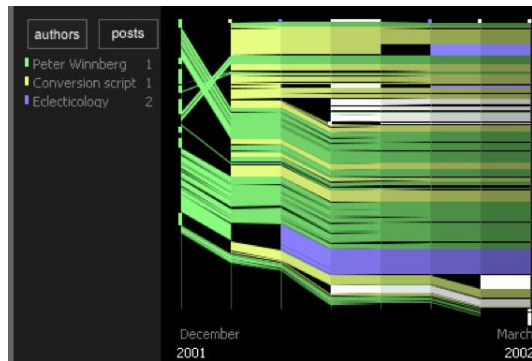
CS 7450

45

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

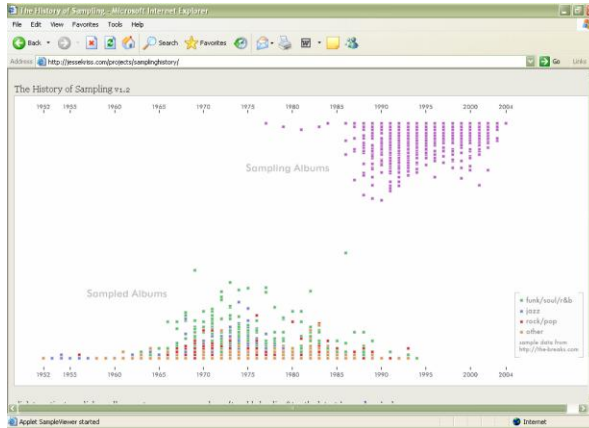
Spacing by revision #

Spacing by time



Spring 2011

# Example 7



<http://jessekriss.com/projects/samplinghistory/>

Spring 2011

CS 7450

47

# Interaction



- Note key role interaction plays in previous two examples
- Common theme in time-series visualization

Spring 2011

CS 7450

48



## Example 8



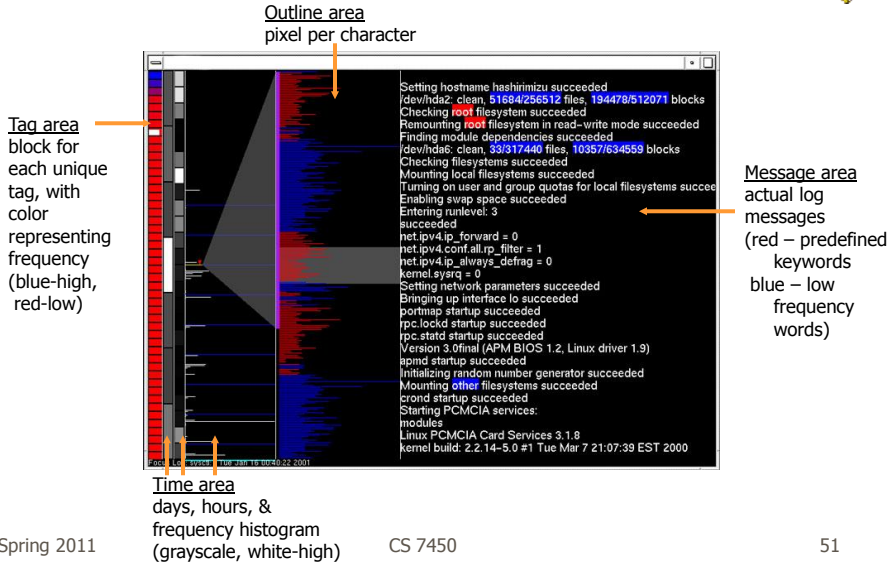
- Computer system logs
- Potentially huge amount of data
  - Tedious to examine the text
- Looking for unusual circumstances, patterns, etc.

## MieLog

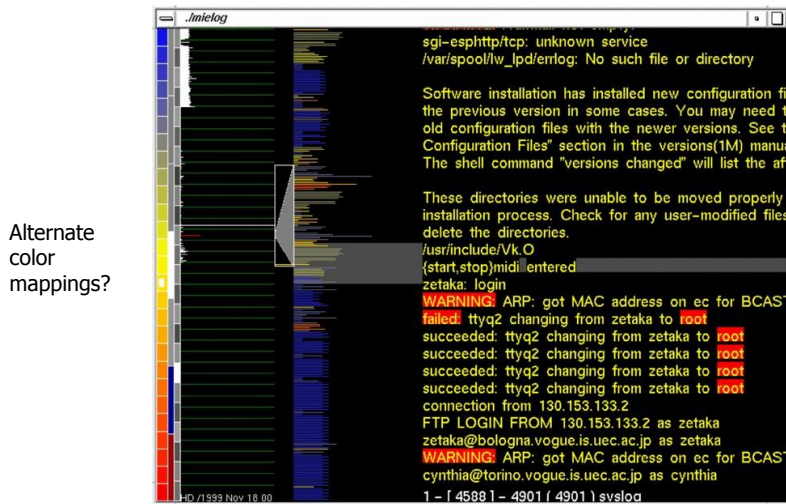


- System to help computer systems administrators examine log files
- Interesting characteristics
  - Discuss

# System View



# Another View



Spring 2011

CS 7450

52

# Interactions



- Tag area
  - Click on tag shows only those messages
- Time area
  - Click on tiles to show those times
  - Can put line on histogram to filter on values above/below
- Outline area
  - Can filter based on message length
  - Just highlight messages to show them in text
- Message area
  - Can filter on specific words

# Thoughts



- Strengths/weaknesses?
- Other domains in which a similar system could be used?

# Example 9



- Very large scale temporal log data
- 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
- Domain: Computer systems management

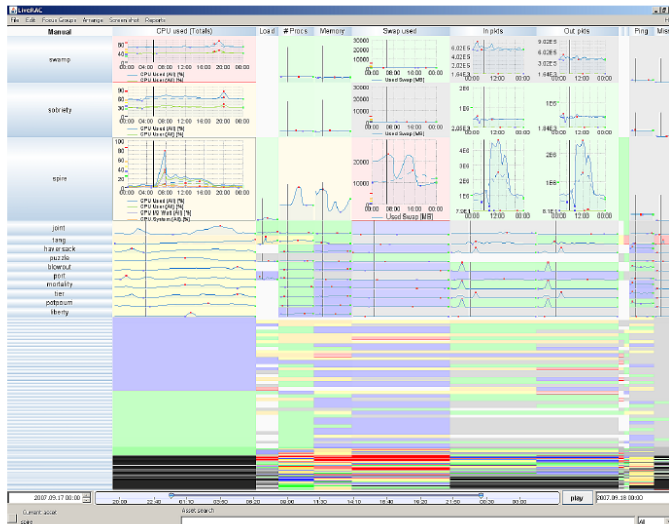
McLachlan et al  
CHI '08

Spring 2011

CS 7450

55

## LiveRAC: Computer system management data



Heavy interaction  
Semantic zooming

Spring 2011

CS 7450

56

# Interaction is Vital

Video

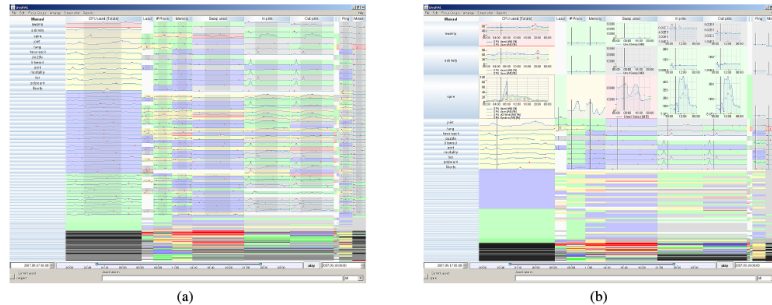


Figure 3. LiveRAC shows a full day of system management time-series data using a reorderable matrix of area-aware charts. Over 4000 devices are shown in rows, with 11 columns representing groups of monitored parameters. (a): The user has sorted by the maximum value in the *CPU* column. The first several dozen rows have been stretched to show sparklines for the devices, with the top 13 enlarged enough to display text labels. The time period of business hours has been selected, showing the increase in the *In pkts* parameter for many devices. (b): The top three rows have been further enlarged to show fully detailed charts in the *CPU* column and partially detailed ones in *Swap* and two other columns. The time marker (vertical black line on each chart) indicates the start of anomalous activity in several of *spire*'s parameters. Below the labeled rows, we see many blocks at the lowest semantic zoom level, and further below we see a compressed region of highly saturated blocks that aggregate information from many charts.

Spring 2011

CS 7450

57

## Design Principles



- Show familiar visual representations whenever possible
- Provide side-by-side comparisons of small multiple views
- Spatial position is strongest visual cue
- Multiple views are more effective when coordinated through explicit linking

Spring 2011

CS 7450

58

# Design Principles



- Follow Shneiderman's mantra
- Avoid abrupt visual change
- User actions should receive immediate visual feedback
  
- Assertion: Showing several levels of detail simultaneously provides useful high information density in context

## Example 10

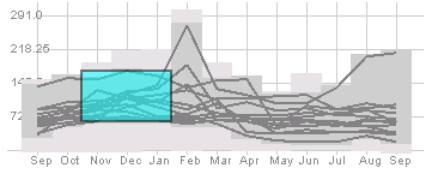
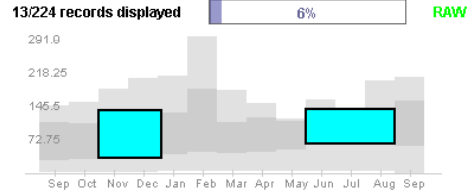


- Most systems focus on visualization and navigation of time series data
- How about *querying*?

# TimeFinder



Can create rectangles that function as matching regions



Light gray is all data's extent

Darker grayed region is data envelope that shows extreme values of queries matching criteria

Multiple boxes are "anded"

Hochheiser & Shneiderman  
Proc. Discovery Science '01  
*Info Vis '04*

Spring 2011

CS 7450

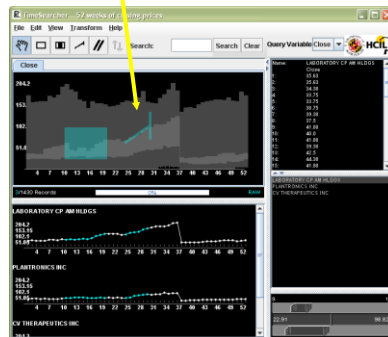
61

# Capabilities



Angular queries

Leaders and laggards



Spring 2011

CS 7450

62

# Limitations



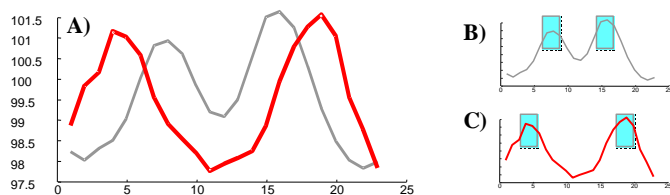
- Can you think of a fundamental limitation of such an approach?

Spring 2011

CS 7450

63

# Problematic Example



Hodgkins patients exhibit double spike in temperature...  
But that can be with differing amounts of time in-between

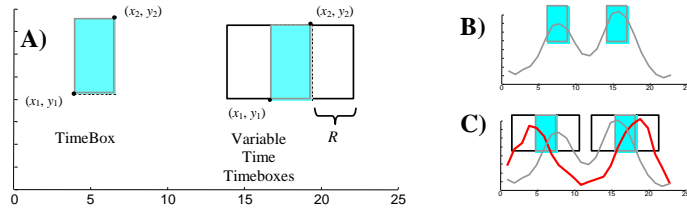
Spring 2011

CS 7450

64



# Solution



Allow time boxes with deltas on each side

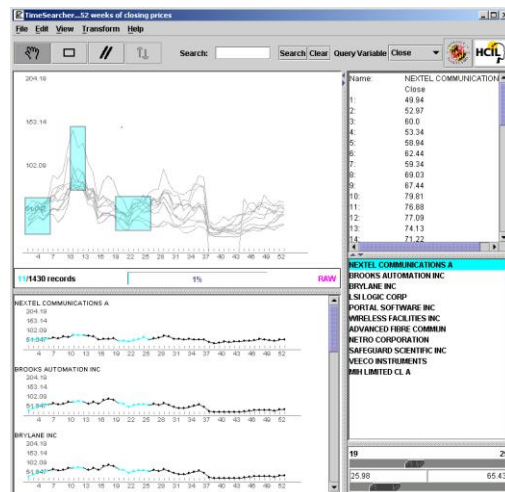
TimeSearcher

Spring 2011

CS 7450

65

# TimeSearcher Interface



Demo

Spring 2011

CS 7450

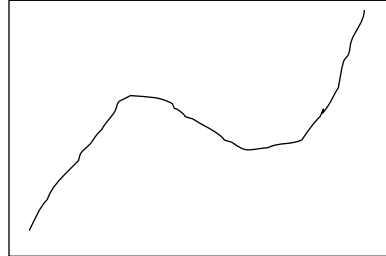
66

# Drawing Queries



## Query Sketch

You specify a timeline query by drawing a rough pattern for it, the system brings back near matches



User-drawn query



Responses

M. Wattenberg  
CHI '01

Spring 2011

CS 7450

67

## Example 11



- Serial, periodic data
- Data with chronological aspect, but repeats and follows a pattern over time
  - Hinted at in last case study
- How might one visualize that?

Spring 2011

CS 7450

68

# Using Spirals



- Standard x-y timeline or tabular display is problematic for periodic data
  - It has endpoints
- Use spiral to help display data
  - One loop corresponds to one period

Carlis & Konstan  
UIST '98

Spring 2011

CS 7450

69

# Basic Spiral Display



One year per loop  
Same month on radial bars  
Quantity represented by size of blob

Is it as easy to see serial data as periodic data?

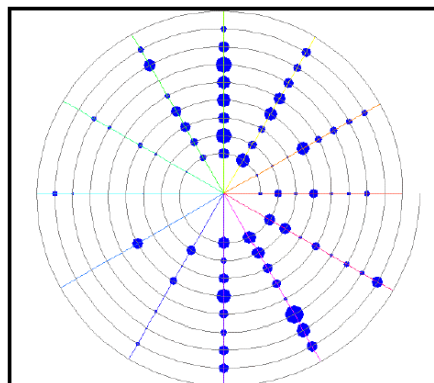


Figure 2. An indented spiral, with spokes, showing monthly consumption percentages for *Baphia Cappardifolia* during the period 1980 – 1988.

Spring 2011

CS 7450

70

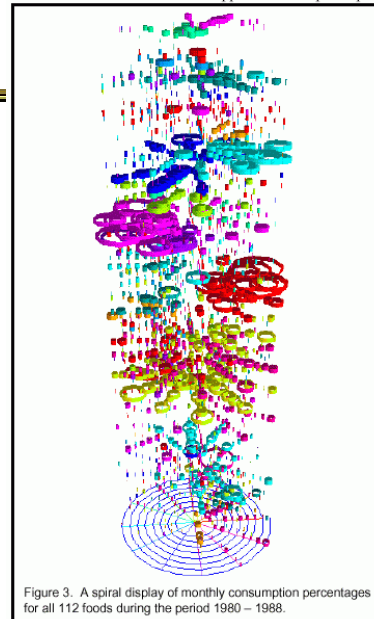
# Advanced Spiral

Same mapping as previous one

Different foods represented by different colors and drawn at different heights

Can you still see serial and periodic attributes?

As with all 3-D, requires navigation



Spring 2011

CS 7450

71

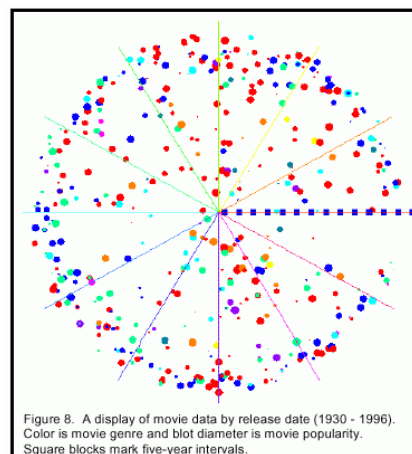
# Compare with Spotfire



Another standard spiral display

Color mapped to movie type

+/- compared to Spotfire?

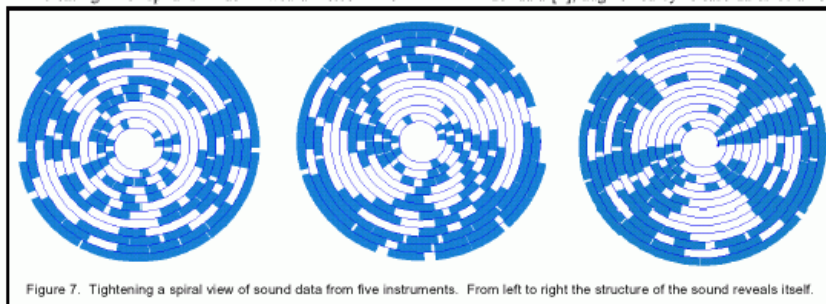


Spring 2011

CS 7450

72

# Unknown Periods



What if a data set doesn't have a regular temporal period?  
Must do some juggling to align periods

Spring 2011

CS 7450

73

## Example 12



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

Spring 2011

CS 7450

74

# GeoTime



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

Kapler & Wright  
InfoVis '04

Spring 2011

CS 7450

75

# Overview



- Objective: visualize spatial interconnectedness of information over time and geography with interactive 3-D view

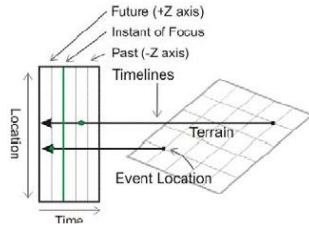
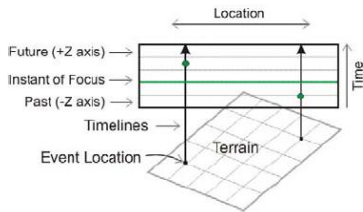


Source: <http://www.oculusinfo.com/>  
Spring 2011

CS 7450

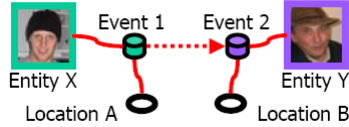
76

# Design Characteristics

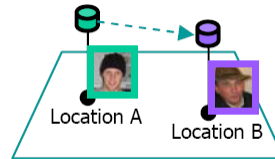


## Dimension usage

Spring 2011



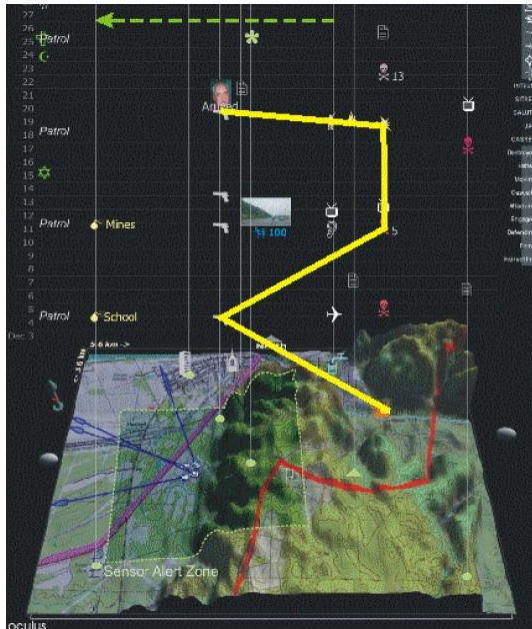
Vector Group with 1+ Actors  
*(Phone Call, email, money transfer)*  
 5 Associations: Same as above plus...  
 • Entity X present at Event 1  
 • Entity Y present at Event 2



## View objects

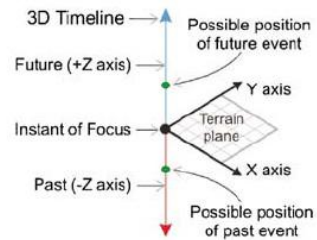
CS 7450

77



## Spatial timelines

- 3-D Z-axis
- 3-D viewer facing
- Linked time chart



Spring 2011

CS 7450

78

# Sample View



Figure 6: Screenshot of GeoTime with time slider at bottom and moveable time scale at right. The green line traces one entity's movement in time and geography.

Spring 2011

CS 7450

79

# Move Time Forward

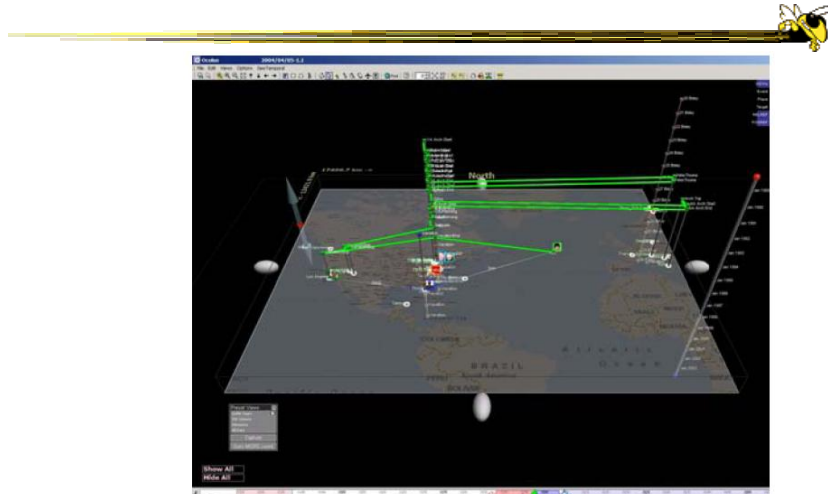


Figure 7: Screenshot of GeoTime with overhead view and time slider advanced forward in time from Figure 6.

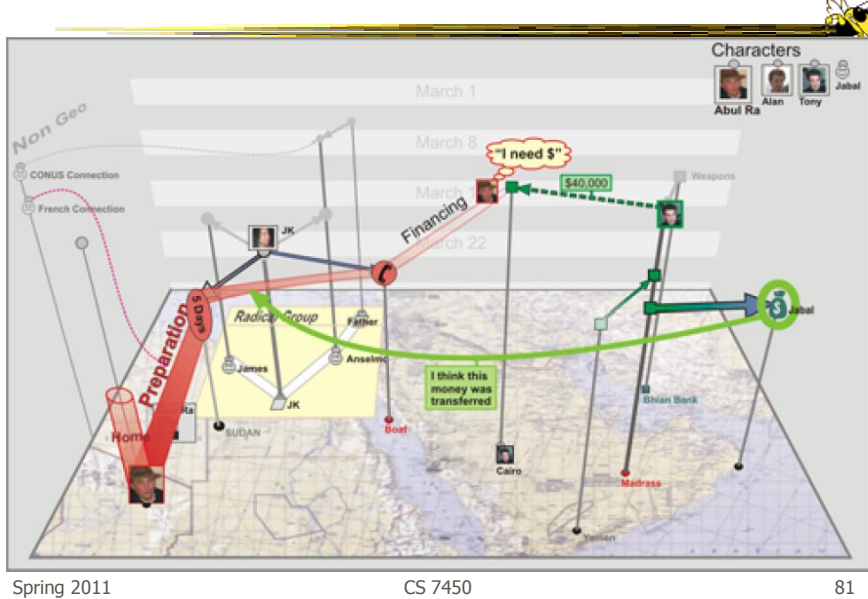
Spring 2011

CS 7450

80



# Telling a Story



# 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?

# 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

Spring 2011

CS 7450

Muller & Schumann '03  
citing  
Frank '98

83

# Upcoming



- Big data
  - Reading  
Yang et al
- Evaluation
  - Reading  
Plaisant

Spring 2011

CS 7450

84

# References



- Spence and CMS books
- All referred to articles
- Jim Foley & Chris Plaue's take on these slides