Design Studies

Lecture 3 CPSC 533C, Fall 2004
Mon Sep 20 2004
Tamara Munzner

Papers Covered

Cluster and Calendar based Visualization of Time Series Data.
Jarke J. van Wijk and Edward R. van Selow, pp 4–9

Using Multilevel Call Matrices in Large Software Projects.
Frank van Ham,

Constellation: Linguistic Semantic Networks
Tamara Munzner,
Interactive Visualization of Large Graphs and Networks (PhD thesis) Chapter 5, Stanford University, 2000, pp 87–122

Design Study

describe task

justify solution

refine until satisfied

Design Study Definition

Design study papers explore the choices made when applying infosvis techniques in an application area, for example relating the visual encodings and interaction techniques to the requirements of the target task. Although a limited amount of application domain background information can be useful to provide a framing context in which to discuss the specifics of the target task, the primary focus of the case study must be the infosvis content. Describing new techniques and algorithms developed to solve the target problem will strengthen a design study paper, but the requirements for novelty are less stringent than in a Technique paper.

InfoVis 03 CFP: [infvis.org/infovis2005/CFP]

Time-series Data Analysis

data: N pairs of (value, time)
   - N large: 50K

tasks
   - find standard day patterns
   - find how patterns distributed over year, week, season
   - find outliers from standard daily patterns
   - want overview first, then detail on demand

possibilities
   - predictive mathematical models
   - details lost, multiscale not addressed
   - scale-space approaches (wavelet, fourier, fractal)
      - hard to interpret, known scales lost
   - 3D mountain: x hours, y value, z days

excellent example, emulate for project writeups!

3D Time-series Data

3D extrusion pretty but not useful
   - daily, weekly patterns hard to see

**Hierarchical Clustering**

- start with all M day patterns
  - compute mutual differences, merge most similar: M-1
  - continue up to 1 root cluster
- result: binary hierarchy of clusters
- choice of distance metrics
dendrogram display common
  - but shows structure of hierarchy, not time distribution

**Power Consumption**

![Graph of power consumption data]

**van Wijk Lessons**

- derived space: clusters
- visual representation of time: calendar
  - linked display
  - interactive exploration
- clear task analysis guided choices
  - reject standard 3D extrusion
  - reject standard dendrogram
- critique
  - color choice not so discriminable
  - especially legend

**Multilevel Call Matrices, van Ham**

- large software project, implementation vs. spec
- link matrix vs. node network

| matrix | force-directed | layered subset |

**Matrices**

- uniform, recursive, stable
- subdivide by:
  - total component count
  - visible subcomponent count
Zooming

abstraction levels

linear interpolation plus crossfade trajectories; will read van Wijk 03 in week 6

Additional Encoding

color: call allowed by spec
color: local region closest red
transparency: call density

histograms: size distribution

Tasks Successfully Supported

visual categorization
  · i.e. libraries with mostly incoming calls

previous summary shown to be incomplete

spotting unwanted calls

determining component dependencies

Linguistic Networks, Munzner

data: MindNet query results

definition graph
  · dictionary entry sentence
  · nodes: word senses
  · links: relation types

Semantic Network

definition graphs used as building blocks

unify shared words

large network
  · millions of nodes
  · grammar checking now, translation future
  · global structure known: dense

probes return local info

Path Query

best N paths between two words

words on path itself

definition graphs used in computation
Task: Plausibility Checking
paths ordered by computed plausibility
researcher hand-checks results
  • high-ranking paths believable?
  • believable paths high-ranked?
  • are stop words all filtered out?

Top 10 Paths Kangaroo→Tail

Goal
create a unified view of relationships between paths and definition graphs
  • shared words are key
  • thousands of words (not millions)
special purpose algorithm debugging tools
  • not understand structure of English

Constellation Video

Traditional Layout
avoid crossings
reason: avoid false attachments

ambiguity artifact salience

Information Visualization Approach
spatial position is strongest perceptual cue
  • encode domain specific attribute
  • plausibility gradient
**Constellation Semantic Layout**

- novel layout algorithm
  - paths as backbone, definition graphs attached
  - curvilinear grid
  - iterative design for maximum semantics with reasonable information density
  - allow crossings for long-distance proxy links

**Selective Emphasis**

- highlight sets of boxes and edges
  - interaction
  - additional perceptual channels
  - avoid perception of false attachments

**Hidden State**

- avoid hidden state
  - change salience instead of toggle drawing

why? closed world assumption

- implicit assumption: if not visible, doesn't exist
- easy to forget previous actions
- draw false negative conclusions

**Information Density**

- early prototype: poor

**Information Density**

- design tradeoff with visual salience
Information Density

grid adjustment

Task–oriented design

previous methods

Task–oriented design

task–specific methods