Project Update:
Law Enforcement Resource Allocation (LERA) Visualization System

Michael Welsman-Dinelle

April Webster

Motivation:

- It is difficult to assess the real impact of different policy decisions and management programs on crime rates
  - E.g., Do anti-drug programs help to reduce youth crime rates?

- Searching for relationships between different variables in a large dataset can be time consuming and frustrating

- In many cases, crime analysts perform this type of analysis using a statistics program (e.g., SAS, SPSS, R) or a data analysis program (e.g., Excel)

Our goal:

- To enable crime analysts to answer these types of questions by bringing together both crime data and crime enforcement policies into a single INTERACTIVE visualization system that supports correlation/regression analysis

The Data:

- We have 2 types of data sets for about 800 US law enforcement agencies for the year 2000:

  - Crime report data
    - Violent (e.g., murder, robbery, rape, etc) and non-violent (e.g., burglary, motor vehicle theft, larceny-theft,) crime rates for an agencies jurisdiction

  - Law Enforcement Management data
    - Specialized units operated by an agency (e.g., juvenile crime unit, drug education in schools, etc)
    - Investment in technology, training, budgets

Supported tasks:

- We aim to support three different types of tasks required of a crime analyst:
  - How does a program impact a crime rate
    - E.g., How does field training impact violent crime?
  - How does a program impact different crime rates
    - E.g., Do drug education programs have an impact on motor vehicle theft rates? How about larceny-theft?
  - How do different programs impact a crime rate
    - E.g., Which programs have been most successful in reducing violent crime rate?

Solutions considered:

- We considered 4 different solutions for the task of interactively visualizing correlation:

  - Parallel Coordinates
    - Repeating an axis for a program; not many dimensions used

  - Table Lens
    - Interested in trends and patterns, not detailed numerical info

  - General Graph Drawing Techniques
    - No compelling info for connecting local agencies by edges

  - Scatterplots
    - Tool commonly used by crime analysts

Our solution:

- An interactive scatterplot visualization system
- Implementation:
  - Java
  - Prefuse Java toolkit
    - Support for scatterplots, tables, SQL queries
    - And for display issues such as mapping from field values to axes, colour, shape, etc
  - Statistical toolkit
    - We have located a Java class that contains formulae for calculating different types of regression curves (linear, quadratic, exponential, etc)

Our solution:

- Specific features:
  - Outlier removal - ability to easily remove outliers, manually and automatically
  - Regression curves
  - Ordering of small multiples – using some scagnostic (e.g., correlation based one?)
  - Aggregation – a focus and context feature
  - Marking – simultaneous, interactive on multiple scatterplots
  - Use of filtering to select one or more states

Anticipated Challenges:

- Finding a Java statistics toolkit with support for outlier detection and regression curve generation

- Determining good orderings of scatterplot small multiples

- Finding a domain expert to use our tool and assess its usability

Progress:

- **Phase 0 – completed**
  - ✓ Downloaded and cleaned sample data
  - ✓ Using Prefuse toolkit for scatterplots
  - ✓ Found Java code for regression curve generation
  - ✗ Unable to find Java code for outlier detection

- **Phase 1 – completed**
  - ✓ Single scatterplot has been implemented
  - ✓ Domain expert has been contacted for usability study; waiting for confirmation of participation

Progress:

- **Phase 2 Part 1 – in progress**
  - Implementation of statistical methods: regression curves, manual outlier removal
  - Plan evaluation component

- **Phase 2 Part 2 – to begin Nov 24**
  - Implementation of small multiples

- **Phase 3 – to begin Dec 1**
  - System evaluation
  - Implementation of optional features – marking
  - Draft report