A Design, Analysis and Evaluation Model to Support the Visualization Designer-User

Iain Dillingham, Jason Dykes and Jo Wood – giCentre, School of Informatics, City University London

Existing visualization design and evaluation frameworks rest on a distinction between the designer and the user. However, there is little explicit guidance on design, analysis and evaluation when the designer *is* the user. A simple solution to this problem is for the researcher (who combines the designer and user roles) to be clear about which activity they are conducting at which point in time. To support the researcher, we propose a design, analysis and evaluation model. This model complements existing visualization design and evaluation frameworks. We have adopted this model in our ongoing research into uncertainty in crowdsourced crisis information.

Formulate working hypotheses

Write a scenario

John wants to explore the distribution of this cluster of incidents in space and time so he brushes the map. Starting this action returns all dots and bars to their unhighlighted state. Completing this action re-subsets the incidents: Bars on both the histogram and the bar chart transition to reflect the new subset. The brushed area is represented by a rectangle on the map whilst brushing is underway. This rectangle remains on the map when brushing has completed.

John selects the standard ellipse function and a standard ellipse is drawn inside the brushed area. John then brushes a bar on the histogram (further subsetting the subset of incidents that he created when he brushed the map) and the standard ellipse is redrawn. John moves the temporal selection forwards and backwards through time: Upon each 'move', the standard ellipse is redrawn and the bars representing the temporal selection are highlighted. Having explored the cluster of incidents around Port-au-Prince, John resets the map, histogram and bar chart to their initial state.

Identify and justify features

Brushing the histogram highlights histogram bars and map dots, and transitions bar chart bars.

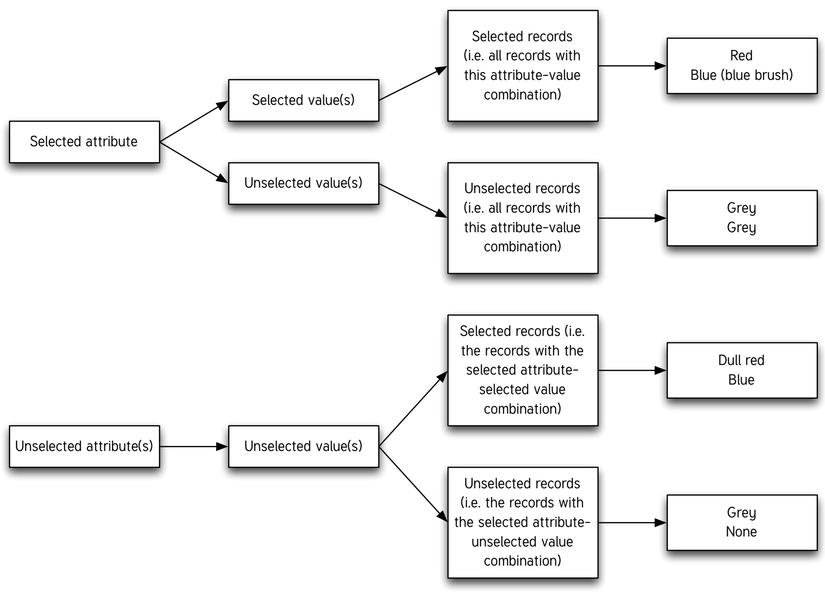
Brushing the histogram results in two instances of a type of interaction that Crampton (2002) terms 'interaction with the data': Histogram bars and map dots are highlighted and bar chart bars are filtered. These instances address the 'suppress' and 'extract' tasks, which are common in geovisualization environments (Crampton, 2002). In this way, brushing the histogram removes information from the bar chart but does not change the amount of information on the histogram or the map.

Standard ellipse function.

A standard ellipse is centred at the mean centre; the long axis extends in the direction of the maximum dispersion; the short axis extends in the direction of the minimum dispersion; the axes are always orthogonal to each other; and each axis is one standard deviation in length, either side of the mean centre (Kitchin and Tate, 2000). Standard ellipses are useful for making spatio-temporal comparisons (O'Sullivan and Unwin, 2003).

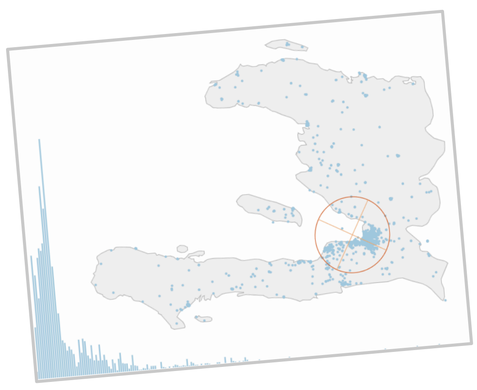
In scenario-based design, a scenario helps the designer identify features of a system; the designer can then make claims about these features, where a claim establishes a causal relationship between a feature and its psychological consequences.

How do Place Survey [1] and Crossfilter [2] encode the filtering process?
[1] http://www.tsr-online.org/placesurvey.html [2] http://www.tsr-online.org/placesurvey.html



Formulate a development plan

Develop a design



Document design decisions

Which base map data should be used?

Two sources of base map data were identified: GADM and Natural Earth. In both cases, the data are available for academic use. Natural Earth data are generalised to 1:10m, 1:50m and 1:110m scales. Furthermore, as well as countries, a range of data are available for mapping populated places, urban areas and transport networks. Consequently, Natural Earth
data were used at the 1:10m scale.

Which map projection should be used?

Snyder's hierarchical selection guideline (Snyder, 1987 cited in Slocum et al., 2009) suggests two projections would be appropriate for Haiti: One conformal (angle-preserving), the stereographic projection; and one equivalent (area-preserving), the Lambert azimuthal equivalent projection. However, neither projection is available in LandSerf, QGIS or OGR.

Undertake analysis

Document findings

Write a research report

It should be possible to:

- simultaneously display standard ellipses for different spatio-temporal selections;
- compare standard ellipses, for example by 'fading-out' standard ellipses, so when the temporal selection is moved forwards or backwards each previous standard ellipse remains visible for a short period;
- summarise the change in location of the mean centre over time
- obtain 'details on demand' (Shneiderman, 1996) such as 'Which day does this histogram bar represent?';
- explore the location attribute, possibly by filtering based on a free-text search.

International Workshop on Visual Analytics (2012)
K. Matkovic and G. Santucci (Editors)

Exploring Patterns of Uncertainty in Crowdsourced Crisis Information

Iain Dillingham, Jason Dykes and Jo Wood[1]
[1]giCentre, School of Informatics
City University London

The humanitarian community is reluctant to use reports from social media when responding to a crisis event, as it fears the costs of untrustworthy and inaccurate information [TBJ*11]. Organisations such as Ushahidi use crowdsourcing to identify 'reliable' reports, but this introduces further uncertainty. By combining human and computational approaches, VA offers a powerful means of tackling complex socio-technical issues such as these. We present an ongoing research programme that combines interactive visual representations with spatial statistical functions to explore patterns of uncertainty in crowdsourced crisis information.

Our prototype software allows us to explore the spatial, temporal and thematic distribution of crisis event reports through linked views that incorporate spatial statistical functions. For example, we can construct a standard ellipse to summarise the spatial distribution of reports from a given time period (Figure 1). We can then step through time period-by-period to see how this distribution changes. We plan to incorporate multiple selections to allow for comparisons between different areas at the same time periods, or the same area at different time periods.

Crisis event reports have a 'locality description' that volunteers generate to produce geographic coordinates. In previous work, we found that most locality descriptions refer to features, such as named places, rather than to distances and directions from features [DJW12]. Potentially, uncertainty varies between types of locality descriptions, so we could, for example, expect features to be located more accurately than features combined with distances or directions. By incorporating text mining techniques, we hope to test this hypothesis and investigate the relationship between locality descriptions and geographic coordinates, and explore how this relationship varies over space and time.

We have applied our software to a dataset relating to the 2010 earthquake in Haiti and note the even spatial distribution of reports for the duration of the dataset (Figure 1). This

Figure 1: *Summary of the spatial distribution of reports for the duration of the Haiti dataset.*

is unexpected because the underlying characteristics, such as population and topography, are clearly non-uniform. However, even this even spatial distribution hides variation at finer temporal resolutions. For example, each of the first seven days of the dataset are characterised by greater dispersion in an approximate NE–SW direction. We plan to extend our analysis to related datasets and are developing a methodology for iterative, analysis-guided VA software development.

References

[DJW12] DILLINGHAM I., DYKES J., WOOD J.: Characterising locality descriptions in crowdsourced crisis information. In *Proceedings of the GISRUK 2012 Annual Conference* (2012). 1

[TBJ*11] TAPIA A. H., BAJPAI K., JANSEN J., YEN J., GILES L.: Seeking the trustworthy tweet: Can microblogged data fit the information needs of disaster response and humanitarian relief organizations. In *Proceedings of the 8th International Conference on ISCRAM* (2011). 1

Hypothesise → Plan action → Take action → Specify learning

We situate scenarios, features and justifications within a cyclical action research process of hypothesising (diagnosing), planning action, taking action, evaluating action and specifying learning. We begin a research cycle by hypothesising; that is, we formulate working hypotheses (an analysis activity). We then plan action, where we write a scenario and identify, and justify, features (design activities). We also formulate a development plan. We then take action, where we develop a design and document design decisions (design activities), and undertake analysis and document findings (analysis activities). Finally we evaluate action and specify learning (an evaluation activity); we write a research report where we reflect on the research cycle.

iain@dillingham.me.uk
dillingham.me.uk

CITY UNIVERSITY LONDON

gicentre.org