Poster Abstracts at Eurographics / IEEE Symposium on Visualization 2011 (EuroVis 2011)
J. Parulek (Poster Chair)

Hierarchical Data Explorer (HiDE) for Exploring Data and Tweeting Insights

A. Slingsby, J. Dykes and J. Wood

giCentre, Department of Information Science, City University London, London, UK

Abstract

The Hierarchical Data Explorer (HiDE) software implements an approach to visual analysis in which datasets are explored by reconfiguring hierarchical graphical layouts of the data in the pursuit of research questions. The HiVE language can describe these layouts and can be used to help record the data exploration process and allow these graphics to be shared through blogging infrastructures. We demonstrate *this approach of visual analysis using HiDE,* give examples *how HiDE is being used and* reflect *on the extent to which HiDE and HiVE facilitates collaborative visual analysis.*

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1. Introduction

Visual representations of data often involve conditioning data into discrete subsets and then depicting numerical summaries of these subsets using visual variables. For example, a bar-chart conditions data into subsets, using length (size), colour and/or order to convey characteristics of these.

This concept can apply recursively – subsets can be further subset and summary statistics of these can be depicted in the same way. This is the case with hierarchical representations such as mosiac plots [Fri94] and treemaps [Shn92], but it also applies to trellis displays [BCS96], small multiples [Tuf83], scatterplot matrices [CCKT83] and thematic maps with embedded charts. The latter examples are usually two-level hierarchies, but can be generalised to more levels.

Datasets can be explored by navigating the design space of hierarchical layouts – by reconfiguring the mappings between data variables and visual variables and by modifying the conditioning hierarchy in response to research questions [SDW09]. These layouts can be described with the HiVE language [giC11] which can describe the *state* of a graphic and provides a set of *operators* for reconfiguring the graphic as part of the data exploration process. Here, we demonstrate this approach using HiDE and consider its role for documenting and sharing the visual exploration process and for collaborative visual analysis.

A number of web-based collaborative visualisation tools allow the capture, storage and sharing of graphical states (e.g. [Hee07]). These are produced and commented upon by from members of a web-based social community for collaborative analysis. These web-based tools typically provide both the infrastructure an integrated interface. HiDE takes a different approach in that it takes advantage of existing blogging infrastructures and use a higher-level human-readable language for describing the graphics that emphases how visual variables map to data variables.

2. Hierarchical Data Explorer (HiDE)

Hierarchical Data Explorer (HiDE; `http://gicentre.org/hide/`) is free software that implements this approach. It provides a GUI that lets users build highly configurable hierarchical graphics of their data resembling various types of commonly-used graphics including small-multiples, spine-plots, mosaic plots, treemaps, rectangular cartograms and scatterplots. Fig. 1 (left) shows how the HiDE GUI is used to construct a graphic of the US Presidential elections results by state, year and party.

HiDE can import and export HiVE representations of

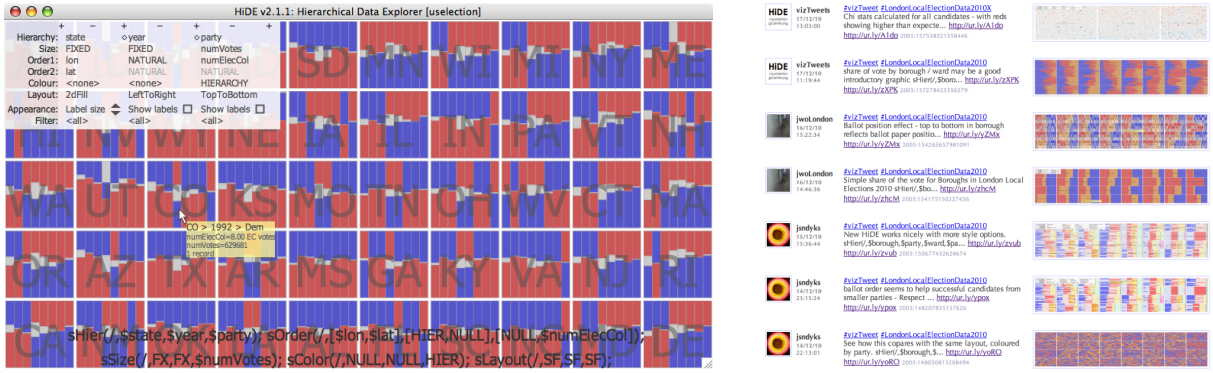

Figure 1: Left: *Screenshot of HiDE being used to construct a graphic of US Presidential elections (*`#uselection`*) by state, year and party. The corresponding HiVE shown at the bottom. States are arranged geographically; years (within each state) are ordered temporally from left-right; party is sized by the number of votes and ordered such that the winning party is at the bottom.* Right: *Graphical views shared on Twitter. All tweets sent from HiDE contain a searchable* `#viztweet` *hashtag and one that identifies the dataset (*`#uselection` *in the case of that on the left).*

these graphics and share them via Twitter. These are searchable with the `#viztweet` hashtag and the hashtag used to identify the dataset. HiDE can also interpret HiVE encoded as a QR code (2D barcode) through a webcam, allowing graphics to be shared and reproduced through printed media such as posters.

3. Uses of HiDE

HiDE has been used in a wide range of different contexts including: tutorials at VisWeek 2009 and the Visual Analytics Summer School [giC10] for teaching the effects of visualisation design decisions on visual analysis; analysing UK voting data in which we identified significant bias; and analysing European Energy data at the OKF Eurostat Hackday in London [Ope10]. Recently, we distributed HiDE with data on socio-economic impacts of natural disasters to users of the Earthquake-report.com website and asked them to explore the data, tweet their findings and provide feedback on whether and how HiDE supported their data exploration process [SDDW11].

4. HiDE demonstration

We will demonstrate HiDE in the various contexts it is being used and will use our experiences to discuss its potential for exploring data, recording the exploratory process, sharing insights, collaborating and making more datasets more accessible to more people. A demonstration video is available at `http://vimeo.com/20526866`.

5. Conclusion

HiDE is open source software that implements a recently published approach to visual analysis [SDW09]. Here, we demonstrate its application, reflect upon its usage in various contexts and reflect on the contribution and potential for HiDE and the approaches that it embodies to collaborative visual analysis.

References

[BCS96] BECKER R. A., CLEVELAND W. S., SHYU M.: The visual design and control of trellis display. *Journal of Computational and Statistical Graphics 5* (1996), 123–155.

[CCKT83] CHAMBERS J., CLEVELAND W. S., KLEINER B., TUKEY P.: *Graphical methods for data analysis*. Duxbury Press, 1983.

[Fri94] FRIENDLY M.: Mosaic displays for Multi-Way contingency tables. *Journal of the American Statistical Association 89*, 425 (Mar. 1994), 190–200.

[giC10] GICENTRE: Exploring design decisions for effective information visualization. http://www.gicentre.org/vass10/, 2010.

[giC11] GICENTRE: Hierarchical Visualisation Expression (HiVE). http://gicentre.org/hive, 2011.

[Hee07] HEER J.: Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of the CHI'07* (2007), pp. 1029–1038.

[Ope10] OPEN KNOWLEDGE FOUNDATION: Exploring european energy data. http://bit.ly/hA9vCr, 2010.

[SDDW11] SLINGSBY A., DANIELL J., DYKES J., WOOD J.: Visualising natural disaster losses from catdat in hide. http://ur.ly/FEop, 2011.

[SDW09] SLINGSBY A., DYKES J., WOOD J.: Configuring hierarchical questions to address research questions. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 977–984.

[Shn92] SHNEIDERMAN B.: Tree visualization with tree-maps: 2D space-filling approach. *ACM Trans. Graph. 11*, 1 (1992), 92–99.

[Tuf83] TUFTE E.: *The Visual Display of Quantitative Information*. Graphics Press, 1983.