Conference or Workshop Item:
Beecham, R orcid.org/0000-0001-8563-7251, Dykes, J, Slingsby, A et al. (1 more author) (2015) Supporting crime analysis through visual design. In: IEEE VIS 2015, 25-30 Oct 2015, Chicago, IL, USA.

Supporting crime analysis through visual design

Roger Beecham*
City University London

Jason Dykes
City University London

Aidan Slingsby
City University London

Cagatay Turkay
City University London

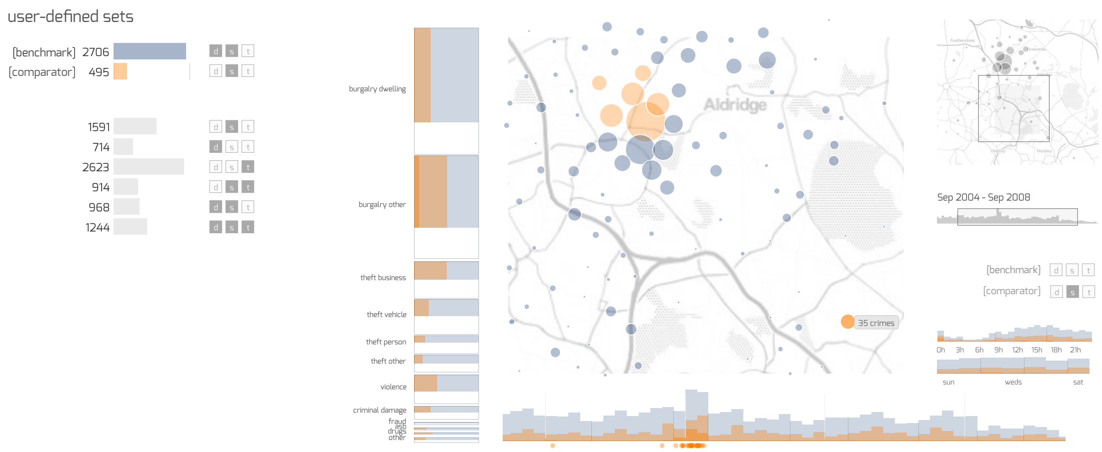

Figure 1: Prototype *volume crime explorer*. Map: density-based clustering is used to summarise occluded point-patterns. Timeline (below): bars represent numbers of crimes in continuous time (month-on-month) and above right cyclic time (day of week and hour of day). Description (immediate left): spine plots currently summarising high-level crime type. Sets (far left): user-defined sets and queries on which they are based. Contextual views (right top): inset timeline and map represent the spatial and temporal extents of the current map and timeline and the current selection relative to the global extent of the dataset.

Abstract

We describe and discuss a visual analysis prototype to support *volume crime analysis*, a form of exploratory data analysis that aims to identify and describe patterns of criminality using historical and recent crime reports. Analysis requirements are relatively familiar: analysts wish to identify, define and compare sets of crime reports across multiple attributes (space, time and description). A challenge particular to the domain, identified through workshops with Police analysts in Belgium and the UK, is in developing exploratory data analysis software that offers some sophistication in data selection, aggregation and comparison, but with interaction techniques and representations that can be easily understood, navigated and communicated. In light of ongoing discussion with Police analysts, we propose four visual design and interaction maxims that relate to this challenge and discuss an early visual analysis prototype that we hope conforms to these maxims.

1 Introduction

Volume crime analysis aims to identify and describe patterns of criminality from mining historical crime reports: structured text providing descriptions of *how* a crime was committed, where it was committed and an estimate as to when it was committed. The outcome of a volume crime analysis might be statistical and descriptive information that supports the deployment of police resources; or it might be the definition and description of a crime series – that is, a set of crimes that are assumed to be linked due to a discriminating spatiotemporal and descriptive signature.

*e-mail:roger.beecham.1@city.ac.uk

Volume crime analysis is relatively speculative and exploratory. Analysts wish to identify emerging sets of crimes that are of interest and usually these sets are defined through comparison. For example, there might be a concentration of crimes of a particular type that do not follow an expected seasonal, weekly or daily pattern; or there might be an unexpected spatial concentration of crimes that, when investigated further, have a particular descriptive characteristic.

When designing exploratory analysis software for supporting *volume crime analysis*, many analytic and design challenges are thus familiar. However, in addition to the more general analytic challenges are difficulties described in Roth et al.[4] and also apparent in recent workshops held with Police analysts, related to the crime analysis domain. Perhaps the most substantial is around misinterpretations that might result from introducing new software and workflows: new data aggregations, data filters and selections and new visual representations. Roth et al. argue that attention should not focus purely on novelty and sophistication in analytic techniques, but on "*investigating how to make* [...] *tools and techniques transparently usable*" [4, p.238] .

Recognising these concerns, we identify four Design Maxims (DM) to which our exploratory analysis prototype for *volume crime analysis* should adhere:

- **DM1**. Techniques for filtering, selection and defining sets should be coherent and consistent across different views of differing data types.
- **DM2**. Representations should support understanding of both the size and relative distribution of a phenomenon.
- **DM3**. Visual cues should be provided to help orient analysts.
- **DM4**. Data aggregation and abstractions should not be vulnerable to misinterpretation.

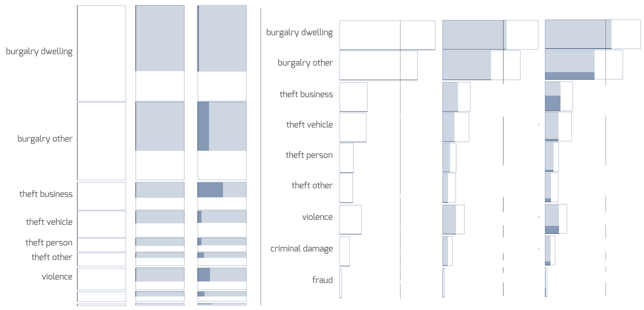

Figure 2: Summaries of all (outline) navigated (light blue) and selected crime reports (dark blue). When selections are made, the orientation is flipped, enabling relative comparison across crime types.

2 Exposition

2.1 Filtering, selection and set definition (DM1)

An initial prototype is shown in Figure 1. Each of the three summary views (description, space and time) are by default linked. By **navigating** (zooming and panning) to a spatial area or time period, a filter is applied on each of the corresponding views: only data visible within the navigated spatiotemporal extent are displayed. It is possible to enable and disable any of these filters by clicking on the *[s]*, *[t]* and *[d]* buttons that appear in the middle right of Figure 1. Temporary **selections** can be performed by hovering or brushing an element of a view; for example, a circle representing an aggregated set of records or a set of bars representing a block of time. To support comparison, temporary selections are represented in the associated spatial, temporal or descriptive views. A description of this encoding appears in Section 2.2. By clicking the hovered or brushed views, these temporary selections can form **sets**: collections of crime reports that are held, can be labelled and returned to. Sets can then be labelled with text.

2.2 Representing relative and absolute distributions (DM2)

That comparisons are supported when data are **navigated**, temporarily **selected** or held as **sets**, is an important requirement. Figure 2 demonstrates how our prototype supports such comparisons. The outlined bars show the complete dataset of crime reports, with heights representing absolute numbers. The blue fill represents the **navigated** crime records and heights again absolute numbers scaled to the total number of crimes of that type in the full dataset. The relative amount of blue versus white space is thus an indicator of crime type prevalence within a given **navigation**. When a **selection** is made or a set defined, the relative size of that selection or set is shown by orienting bars horizontally (as in spine plots [2]); there are more thefts and violent crimes for the **selected** reports than would be expected given the prevalence of those crime types in the **navigated** dataset. We speculate that providing analysts with the ability to switch between vertical and horizontal orientations for absolute and relative sizes might support interpretation of the absolute and relative relationships between sets and subsets.

2.3 Persistent orientation through visual cues (DM3)

In *volume crime analysis*, a common goal is to find and then define crime series with some detail: a set of crimes of a particular type or description that occur within a particular spatial area and time period. Since sets can be so flexibly defined in our prototype, we attempt to provide some visual summary of the spatial *[s]*, temporal *[t]* and descriptive *[d]* queries underlying those sets (far left of Figure 1). In later design iterations more expressive summaries will be developed. Additionally, since spatial filters are automatically applied when **navigating** the map or timeline views, the cues in the middle right help clarify the data types that are being queried within a current **navigation**.

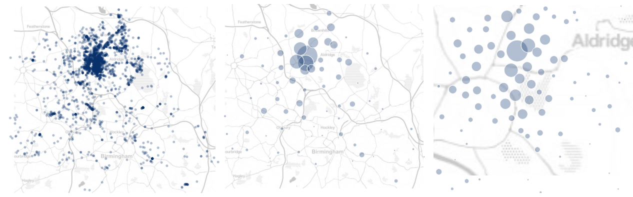

Figure 3: Left: dot-density map suffering from occlusion. Middle: density-based clustering applied. Right: recomputed density-based clustering after navigation.

2.4 Dealing with spatial density: an alternative to 'hotspot' maps (DM4)

Crimes are often heavily spatially concentrated and using semi-transparent ellipses to represent individual crimes is problematic. This problem of occlusion exists even for the 3,000 crime reports to which we have access (Figure 3); it is conceivable that hundreds of thousands of crime reports might be interrogated in a volume crime analysis. A technique currently used by crime analysts is 'hotspot' mapping. Here, a continuous surface of spatial densities is created using *kernel density estimation* [3]. Problems around interpreting and communicating hotspot maps are well-documented and discussed in Roth et al. [4]. We instead use density-based clustering to generate collections of spatially proximate point locations, draw ellipses at the spatial centres of these clusters and, to prevent occlusion, draw clusters in descending size order. Clusters are recalculated as the map is **navigated** (zoomed and panned). This clustering procedure is described in Andrienko & Andrienko [1].

3 Conclusion

We introduce an early data prototype for supporting *volume crime analysis*. The prototype offers some sophistication in data selection, aggregation and comparison, but informed by four stated Design Maxims, attempts to provide interaction techniques and representations that can be easily understood and communicated. We continue to develop specifications around design and interaction for the prototype and we hope this poster will provide interesting discussion at Vis 2015. The work forms part of a long-term project that aims to develop visual analytics software to support crime analysis.

Acknowledgements

This research has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) to Project VALCRI under the EC Grant Agreement No. FP7-IP-608142 awarded to Middlesex University and partners.

References

[1] N. Andrienko and G. Andrienko. Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, 17:205–219, 2011.

[2] J. Hummel. Linked bar charts: Analysing categorical data graphically. *Computational Statistics*, 11:23–33, 2007.

[3] D. O'Sullivan and D. Unwin. *Geographic Information Analysis*. John Wiley & Sons, New Jersey, USA, 2002.

[4] R. E. Roth, K. S. Ross, B. G. Finch, W. Luo, and A. M. MacEachren. Spatiotemporal crime analysis in US. law enforcement agencies: Current practices and unmet needs. *Government Information Quarterly*, 30(3):226–240, July 2013.