Revealing the Presence of Absence in Geovisual Analytics

Anthony C. Robinson

GeoVISTA Center, Department of Geography, The Pennsylvania State University, University Park, USA
arobinson@psu.edu

Abstract

The emergence of big data in Geography presents new opportunities to see and understand more about our world than ever before. While this era provides a broad array of new spatial analytical affordances through new types of observations and measurements, it should also encourage us to consider how analysts recognize and account for missing information and attributes of absence. In fact, we stand to learn a great deal from the silence that may be present in our maps. Here we present ideas for new GIScience research to tackle the challenge of visual representation and analytical reasoning with missing data and data that have attributes of absence. We highlight the potential analytical affordances that arise from explicitly visualizing what is missing, and show how this problem area can be approached from representational as well as analytical reasoning viewpoints.

1998 ACM Subject Classification H.2.8 Database Applications - Spatial databases and GIS

Keywords and phrases absence, missing data, geovisual analytics, cartography

Digital Object Identifier 10.4230/LIPIcs.GVIZ.2018.1

1 Introduction

Users have increasingly high expectations for the coverage, timeliness, and relevance of spatial data sources. The development of streaming sensors, rapidly updating satellite sources, and new forms of spatially-oriented social media contributes to these increased expectations. We argue that the increasing availability of spatial data should urge us to explore and emphasize where we *do not* have data, or where our data indicate attributes of absence. For example, a great deal of emphasis is placed on social media as a proxy for broad social patterns in cities, but in reality it is only a small proportion of the population who actually use such services, and even fewer who enable geolocation for their posts [9]. In addition to mapping these new and interesting data sources, we argue that we should also focus on mapping where we do not have information, as understanding where we do not have data becomes a powerful analytical clue in its own right.

In the sections that follow we begin by characterizing the problem domain for revealing the presence of absence in geovisual analytics. Next we propose visual approaches for representing the presence of absence. Finally, we describe key research challenges related to the future development of geovisual analytics in order to support analytical reasoning with absence.

2 The presence of absence

In this work we refer to both missing data and data with attributes of absence as the challenge of representing *the presence of absence* in geovisual analytics. We build here on a rich body of prior work to develop approaches for calculating and visualizing missing data. In statistics,

it is quite common to calculate missing values [8], and there have been multiple efforts to explicitly visualize missing values [10, 3]. These approaches however have treated the visual representation of missing data as simply another qualitative attribute, with the most common approach utilizing a specific hue (for example, bright red) to represent missing data. In the present era there are many situations in which spatial data may have missing values. For example, we may have household demographic survey data for which some respondents have chosen not to respond to certain questions. We could also have sensor networks spread throughout a large region which monitor environmental attributes that periodically fail or go offline.

We also focus here on data that include attributes of absence. These data are not missing, but rather have qualities that indicate a form of absence. For example, property data in a city may include information that suggests a given structure is vacant. It becomes therefore a cartographic design decision to represent these data as a simple attribute, perhaps using a dedicated hue as is common in the visualization of missing data. On many maps however we note that cartographers frequently choose to not represent attributes of absence at all. And in that case, the user only has blank spaces on a map to interpret. This is of course problematic because knowing where blank spaces indicate an attribute of absence that we have measured versus their indication of a place where we have no data at all is essentially impossible.

One could argue that missingness and absence in geographic data is simply a specific type of uncertainty, and we should focus then on applying what has been learned already from many previous studies on representing and reasoning with uncertainty in GIScience. We hypothesize however that visualizing the presence of absence provides a unique challenge in that the phenomena it hopes to represent has natural visual affordances in the real world. For example, we may notice when things are physically missing from their place. Some physical artifacts leave behind a residue when they are taken away (e.g. a glass of water may leave behind a ring on a table). And we understand already that visual perception of missingness is quite challenging for most people. For example, studies of visual search asymmetry suggest that it is much easier for people to find the presence of attributes on multiple visual targets than it is for them to find the absence of attributes on multiple visual targets [11]. We also know that detecting change in animated maps can be quite problematic, which suggests that it may not be simple make visualizing absence an interactive feature in geovisual analytics systems [4].

3 Visual approaches

It should be possible to apply a wide range of visual approaches for representing the presence of absence in geovisual analytics. In broad terms, we propose three major categories of visual methods; static methods, dynamic methods, and interactive methods.

Static methods can build on the set of basic visual variables that have been proposed and refined over multiple decades of graphic design and research [2, 6], texture, and opacity, we also highlight the fact that a very common method for representing absence is to simply leave those areas blank. Manipulating darkness would also appear to be a variation of color saturation that would have an apparent connection to the affordances of missingness that we encounter in the real world (e.g. shadows and residues).

Dynamic methods for representing the presence of absence could make use of short animations that transition between visual effects to show where data are missing or have attributes of absence. Such methods could also leverage creative techniques used in illustration

and movies, for example to have data vanish in a puff of smoke, or to have the appearance of going down a drain.

Finally, both static and dynamic approaches could be manipulated interactively by users to control the intensity of their effects, the onset of transitions, and other design aspects. We note here the popular use of swipe maps to compare two different satellite images, and the lack of significant scholarship so far regarding how well that interactive paradigm compares to other means for visual comparison.

4 Research challenges for absence in geovisual analytics

In order to begin addressing the representational and reasoning challenges associated with the presence of absence in geovisual analytics, we suggest several key questions that should be the focus of future research.

4.1 How can we visually represent absence in geovisual analytics systems?

We have proposed several basic visual approaches above, but we do not assume we have exhausted the potential for creative visual methods for expressing the presence of absence in a geovisual analytics environment. For example, there may be combinations of the individual methods we propose in section 3 that work particularly well for making absence more salient in geovisual analytics systems. Additionally, we must carefully potential evaluate approaches for visual representation of absence in order to understand their relative performance and the extent to which users have preferences for some methods over others. Situating these studies in real contexts of usage will be critically important [7], as the potential impact of the presence of absence may be reasonably expected to vary a great deal depending on particular analytical scenarios.

4.2 What are the affordances of absence provided by computational methods in geovisual analytics systems?

A core element of any geovisual analytics environment is the connection between computational methods for extracting and analyzing patterns with interactive map-oriented interfaces [1]. We propose that new work should focus on exploring how the computational methods we leverage in geovisual analytics may provide affordances from which we can derive information about missing data or data with attributes of absence. Rather than simply avoiding incompleteness or hiding it from our users, we could potentially capture, store, and represent that information explicitly. A starting point for this work could be to focus on computational methods that have elements of imputation, to calculate and represent the extent to which imputation is occurring.

4.3 How do users reason with representations of absence in geovisual analytics systems?

The core goal of any geovisual analytics system is to support human reasoning with spatial information in order to support problem solving in complex contexts. While a great deal of progress has been made to date in understanding how people reason with uncertainty in geovisualization [5], we argue for new work that focuses on the extent to which missing data

or data with attributes of absence may intervene in analytical reasoning with geographic information.

5 Conclusion

We have argued here for attention to the problem of visualizing what is normally left unrepresented in contemporary geovisual analytics systems. The presence of spatial data will of course remain very interesting, but we must also pay close attention to where we do not have coverage, or where our data convey attributes of absence. To do so we need to develop new visual approaches that make absence more salient, and we need to better understand the ways in which users could interact with absence through geovisualization, as well as to understand how users incorporate knowledge of absence in their sensemaking.

References

1 G. Andrienko, N. Andrienko, P. Jankowski, D. Keim, M-J. Kraak, A. M. MacEachren, and S. Wrobel. Geovisual analytics for spatial decision support: setting the research agenda. *International Journal of Geographical Information Science*, 21(8):839–857, 2007.

2 J. Bertin. *Semiology of graphics: Diagrams, networks, maps*. University of Wisconsin Press, Madison, WI, 1967.

3 C. Eaton, C. Plaisant, and T. Drizd. *Visualizing Missing Data: Graph Interpretation User Study*, pages 861–872. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. URL: http://dx.doi.org/10.1007/11555261_68, doi:10.1007/11555261_68.

4 C. Fish, K.P. Goldsberry, and S. Battersby. Change blindness in animated choropleth maps: an empircal study. *Cartographic and Geographic Information Sciences*, 38(4):350–362, 2011.

5 C. Kinkeldey, A. M. MacEachren, M. Riveiro, and J. Schiewe. Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science*, pages 1–21, 2015. URL: http://dx.doi.org/10.1080/15230406.2015.1089792, doi:10.1080/15230406.2015.1089792.

6 A. M. MacEachren, R. E. Roth, J. O'Brien, B. Li, D. Swingley, and M. Gahegan. Visual semiotics amp; uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2496–2505, Dec 2012. doi:10.1109/TVCG.2012.279.

7 R.E. Roth, A. Çöltekin, L. Delazari, H. F Filho, A. Griffin, A. Hall, J. Korpi, I. Lokka, A. Mendonça, K. Ooms, and C.J.P.M van Elzakker. User studies in cartography: opportunities for empirical research on interactive maps and visualizations. *International Journal of Cartography*, 3(sup1):61–89, 2017. doi:10.1080/23729333.2017.1288534.

8 D. B. Rubin. Inference and missing data. *Biometrika*, 63(3):581–592, 1976. URL: http://dx.doi.org/10.1093/biomet/63.3.581, doi:10.1093/biomet/63.3.581.

9 Luke Sloan and Jeffrey Morgan. Who tweets with their location? understanding the relationship between demographic characteristics and the use of geoservices and geotagging on twitter. *PLOS ONE*, 10(11):1–15, 11 2015. URL: https://doi.org/10.1371/journal.pone.0142209, doi:10.1371/journal.pone.0142209.

10 D.F. Swayne and A. Buja. Missing data in interactive high-dimensional data visualization. *Computational Statistics*, 13(1):15–26, 1998.

11 A. Treisman and J. Souther. Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology*, 114(3):285–310, 1985.