

Quality Metrics for Tile Grid Maps

Krist Wongsuphasawat*

Twitter, Inc.

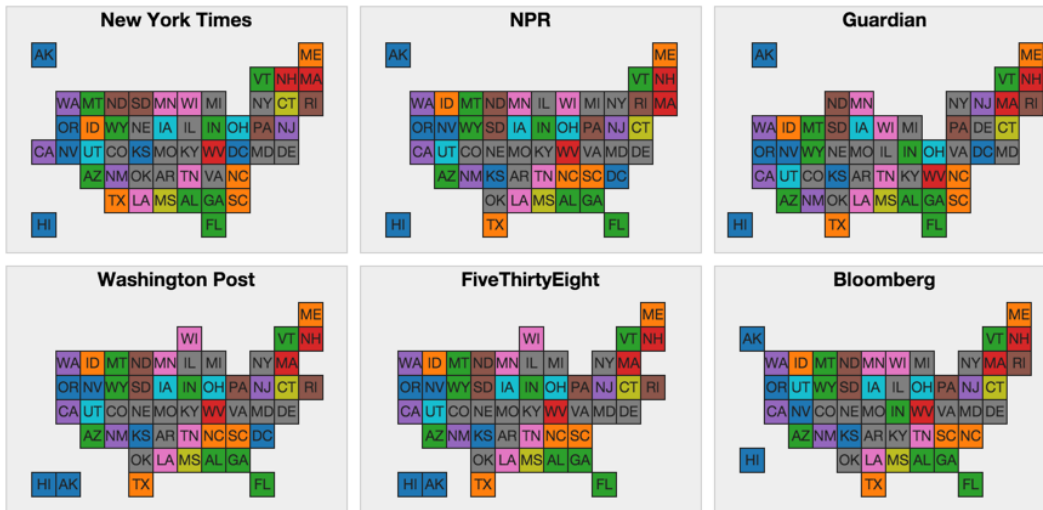


Figure 1: Different US map layouts from six publishers

ABSTRACT

Tile Grid Map has been adopted for creating Choropleth map. However, multiple publishers had already developed different map layouts. To compare these layouts, I define a set of quality metrics that can be used to evaluate and compare them. A case study comparing six map layouts was also included.

Index Terms: H.5.m [Information Systems]: Information Interfaces and Presentation (e.g. HCI)—Miscellaneous

1 INTRODUCTION

Cartograms combine statistical and geographical information in thematic maps, where areas of geographical regions are encoded in proportion to some statistic [5]. *Tile Grid Map* [1, 6] has become a quite popular choice lately [7] for creating Choropleth map [3, 4]. In Tile Grid Map, each region is represented as a tile of the same shape and equal size, therefore does not introduce biases due to different sizes of the regions. The tiles are placed to fit within a grid at approximate positions close to the actual geographic position.

However, due to approximate positionings, multiple publishers had already developed slightly different map layouts. How should one decide which layout to use? It is also not difficult to create another layout [2]. What would be an effective way to check that if the new layout is better? For these purposes, I have defined a set of metrics for evaluating the quality of a Tile Grid Map layout.

2 QUALITY METRICS & INTUITIONS BEHIND THEM

These following metrics can be used as guidelines for selecting a layout to use, or optimization goals for new layouts:

*e-mail: kristw@twitter.com

1. **Lookalike (Boolean)** : The overall shape looks similar to the geographic map. I define this as boolean rather than a quantitative value based on the assumption that the resemblance only helps with first impressions for viewers to recognize that it is a map of a certain geographical region, but does not affect the interpretation afterwards.
2. **Topology (%)** : Regions that are neighbors in reality should appear as neighbors in the Tile Grid Map. For example, North Dakota and South Dakota should be adjacent. A pair of regions that appear as neighbors in the map are considered valid if the two regions are really neighbors in a geographic map. A pair is considered invalid otherwise. Topology is computed from this formula:
$$\text{No. of valid neighbor pairs} / \text{No. of geographic neighbor pairs}$$
3. **Inaccuracy (%)** : Regions that are not neighbors should not appear as neighbors in the Tile Grid Map. For example, California and Florida should not be next to each other.
$$\text{No. of invalid neighbor pairs} / \text{No. of neighbor pairs}$$
4. **Misdirection (%)** : The relative positions between neighbor regions should be close to reality. For example, North Dakota is on the north of South Dakota, so it should be in the north of South Dakota in the Tile Grid Map. To compute this metric, an angle between the two regions in each valid neighbor pair is computed and compared against an angle between their centroids on a geographic map. If the difference is greater than 45 degrees, it is considered a misdirection.
$$\text{No. of misdirections} / \text{No. of valid neighbor pairs}$$
5. **Area**: A good Tile Grid Map should be compact. This is simply calculated from No. of Rows \times No. of Columns

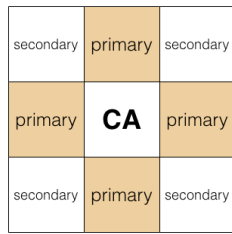


Figure 2: Neighbors – primary and secondary

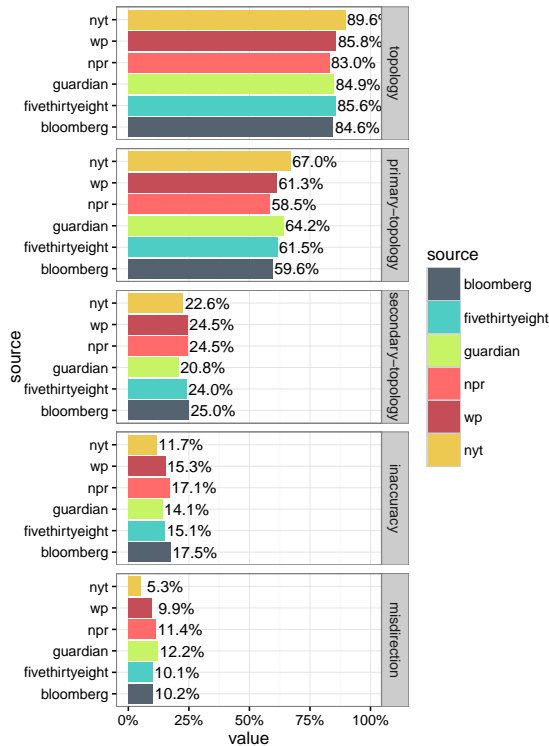


Figure 3: Topology, Inaccuracy and Misdirection

3 CASE STUDY

Six square US Tile Grid Map layouts from *New York Times*, *NPR*, *Guardian*, *Washington Post*, *FiveThirtyEight* and *Bloomberg* were selected for the study. (See Figure 1.) These were mentioned by the NPR Blog [1]. Of these six, *FiveThirtyEight* and *Bloomberg* excluded District of Columbia (DC) from their maps.

In this study, two square tiles are considered *primary neighbors* if they share one side of the squares. They are considered *secondary neighbors* if their corners touch, as illustrated in Figure 2. Both primary and secondary neighbors are included for *topology*. However, only primary neighbors are counted for *inaccuracy*.

4 RESULTS

All of them, although look slightly different, resemble the US map. And as mentioned above, there is no sense trying to compare which one looks more similar to the US. One could conduct this study with Mechanical Turks to confirm if the mass audience agrees. Let's look at other metrics.

From Figure 4, *New York Times* (nyt) receives the highest topology (89.6%). Meaning that it can capture 89.6% of the neighbor relationships represented by a geographic map. Its primary topol-

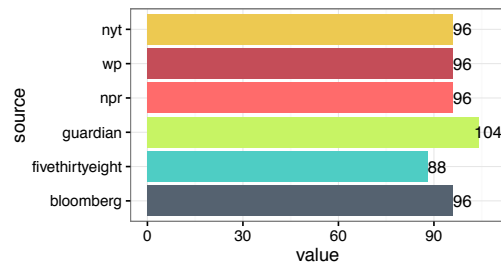


Figure 4: Area

ogy is also the highest at 67%, showing that this layout makes best use of adjacent cells. It also causes the least amount of inaccuracy (11.7%). There are 5 misdirections, which is also the lowest of the six layouts (5.3%). Most notable misdirections are Virginia on the west (180) of North Carolina instead of north(272) and South Dakota on the east (0) of North Dakota instead of south (89).

On the contrary, NPR receives the lowest topology (82.1%), which is probably related to the fact that they also use the adjacent cells less effective than the others (58.5% primary topology). Guardian gets the highest misdirection (12.2%) while Bloomberg puts states that are not neighbors next to each other the most (17.5% inaccuracy).

In term of area, most of the layouts use 8x12 grid and yield an area of 96. The exceptions are Guardian and FiveThirtyEight who use 8x13 and 8x11, respectively.

5 CONCLUSION

Based on the results above, the *New York Times* version seems to perform best as it can capture most relationships while also providing the least misleading visuals. More importantly, I hope the readers will find these metrics useful and believe that they should be applicable to other types of Tile Grid Map, such as hexagon tile. It would be interesting to see how they perform compare to the square ones. The full results, data and code for computing metrics are available at <https://github.com/kristw/gridmap-layout-usa>.

ACKNOWLEDGEMENTS

The author wishes to thank the readers of the original blog post that led to this manuscript for their feedback and thoughtful comments from the anonymous reviewers.

REFERENCES

- [1] D. DeBelius. Let's tessellate: Hexagons for tile grid maps. NPR Blog blog.apps.npr.org/2015/05/11/hex-tile-maps.html, May 2015. [Online; accessed 23-Jun-2016].
- [2] C. Dempsey. How to make a tile grid map using excel. GIS Lounge www.gislounge.com/how-to-make-a-tile-grid-map-using-excel, Nov 2015. [Online; accessed 23-Jun-2016].
- [3] N. Gale and W. C. Halperin. A case for better graphics: The unclassed choropleth map. *The American Statistician*, 36(4):330–336, 1982.
- [4] M.-J. Kraak and F. Ormeling. *Cartography: visualization of spatial data*. Guilford Press, 2011.
- [5] S. Nusrat and S. G. Kobourov. The State of the Art in Cartograms. In *Computer Graphics Forum, special issue on EUROVIS'16*, 2016.
- [6] Z. Romano. Data visualization strategies using tile grid maps. GIS Lounge www.gislounge.com/data-visualization-strategies-using-tile-grid-maps, Nov 2015. [Online; accessed 23-Jun-2016].
- [7] N. Yau. The great grid map debate of 2015. Flowing Data flowingdata.com/2015/05/12/the-great-grid-map-debate-of-2015, May 2015. [Online; accessed 23-Jun-2016].