

GMap: Drawing Graphs as Maps

Emden R. Gansner, Yifan Hu, and Stephen G. Kobourov

AT&T Labs - Research, Florham Park, NJ USA
{erg,yifanhu,skobourov}@research.att.com

1 Introduction

In graph drawing, vertices are typically represented as points in two or three dimensional space and edges are represented as lines between the corresponding vertices. Other representations have also been considered. For example, treemaps use a recursive space filling approach to represent trees. There is also a large body of work on representing planar graphs as contact graphs of geometrical objects. GMap is an algorithm that represents general graphs as maps [2]. Our overall goal is to create a representation which makes the underlying data easy to understand and visually appealing. Our map representation is especially effective when the underlying graph contains structural information such as clusters and/or hierarchy. The traditional point-and-line representation of graphs can require considerable effort to comprehend, and often puts off general users. On the other hand, a map representation is more intuitive, as people are very familiar with maps and even enjoy carefully examining maps.

2 The Mapping and Coloring Algorithm

The first step in our GMap algorithm is to embed the graph in the plane. In our implementation we use a scalable force directed algorithm [3]. The second step is a cluster analysis of the underlying graph or the embedded pointset. Here we use modularity based clustering [4] as it is a good fit [5] for the force directed algorithm we employ. In the third step the embedding and the clustering are used to create the map. A Voronoi diagram of the vertices is generated. To create “European-style” borders we use the vertex sets in each cluster together with some random points and generate “form fitting” outer boundaries. Vertex weights are used to determine the font size of the vertex label, and the size of the label is used to create the area in the map that corresponds to the vertex. We then merge Voronoi cells that belong to the same cluster, thus forming regions of complicated shapes. The overall algorithm has complexity $O(|V| \log |V|)$ and easily scales to graphs with tens of thousands of vertices.

Because countries in GMap are not necessarily contiguous, we need as many colors as the total number of countries, in order to make each country uniquely identifiable by its color. We use a two-step heuristic color assignment algorithm to ensure that neighboring countries are colored with as distinctive colors as possible. In the first step we apply a spectral algorithm to the *country graph* that

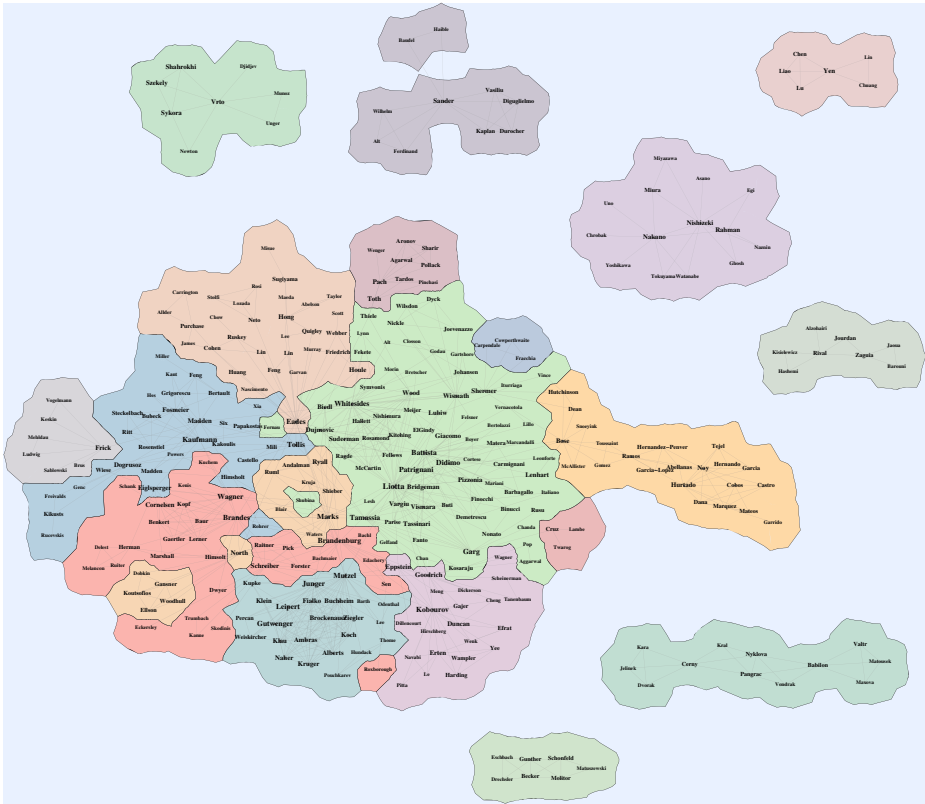


Fig. 1. Author collaboration map for the GD conference, 1994-2004

captures the neighboring structure of countries. In the second step we apply a greedy color-swapping algorithm, with the objective of maximizing the minimal color difference between any two neighboring countries. For further details see [2].

3 GMap Example

We consider the collaboration graph of authors with publications in the Symposium on Graph Drawing in the period 1994 to 2004. Authors are vertices and there is an edge between two authors if they have collaborated on at least one paper in this period. The font size of the vertices is proportional to the number of papers by that author; see Fig. 1. It is easy to see that European authors dominate the main continent. Several well-defined German groups can be seen on the west and southwest coasts. A largely Italian cluster occupies the center, with an adjacent Spanish peninsula in the east. The northwest contains a mostly Australasian cluster. Two North American clusters are to be found in the southeast and in the southwest, the latter one made up of three distinct components.

A combinatorial geometry cluster forms the northernmost point of the main continent. Most Canadian researchers can be found in the central Italian cluster and the Spanish peninsula. Northeast of the mainland lies a large Japanese island and southeast of the mainland there is a large Czech island. Northwest of the mainland is *Crossing Number* island.

GMap has been used to visualize data from several different areas [2, 1], including similarities between musicians (last.fm data) and books (Amazon.com data).

References

1. <http://www.research.att.com/~yifanhu/gmap/>
2. Gansner, E.R., Hu, Y.F., Kobourov, S.G.: Gmap: Drawing graphs as maps (2009), <http://arxiv1.library.cornell.edu/abs/0907.2585v1>
3. Hu, Y.F.: Efficient and high quality force-directed graph drawing. *Mathematica Journal* 10, 37–71 (2005)
4. Newman, M.E.J.: Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* 103, 8577–8582 (2006)
5. Noack, A.: Modularity clustering is force-directed layout. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* 79 (2009)