

Flow Maps - Automatic Generation and Visualization in GIS

Birgit Pieke, Antonio Krüger

Institute for Geoinformatics, University of Muenster, Germany
{b.pieke, antonio.krueger}@uni-muenster.de

Abstract. Flow maps are used in order to indicate movements by persons or objects in a given time period. The major problem concerning the production of flow maps is the appropriate placement of movement lines in the map. This paper presents a method of placing lines automatically, based on a modification of an existing algorithm. The algorithm has been implemented as an ArcMap extension and was subsequently tested with a data set representing pupil school transitions.

1 BASIC PRINCIPLES

Flow maps are of use in many applications: whenever movements by people or objects take place, they can be shown on a flow map. The data visualized in these maps exhibit a high complexity, due to the high number of transition edges present in such a map. Drawing a flow map is difficult, because every flowline has to be placed very carefully depending on the other edges and nodes. In order to arrange the map as effectively as possible, a special emphasis in its generation has to be put on a clever placement of the edges within the map.

A study on the cross-linking of schools (in terms of transition of students) within a defined region served (us) as a use case to analyse the problem of flow maps.

The effectiveness of a map is at its highest if the designer and the map user use the same concepts for the organization of the representation, see MacEachren (1995). Following MacEachren (1982), the effectiveness of a map is determined by a variety of aspects, of which map complexity and symbolization are the most important ones. The symbolization specifies how data are represented on a map, whereas the complexity specifies the degree to which the symbols on a map appear intricate (MacEachren 1982). In this use case the important factors are:

1. The complexity, because of its negative impact on effectiveness. Thus it is important to reduce the number of objects depicted/visualized in the map to such an extent that no visual clutters are arising.

2. The symbolization, because the producer can control it. Linear signatures are a common method to represent movement in maps, wherein the width of the line can be used to show different quantities.

The difficulty in automatically generated flow maps lies in the inherent complexity of the data. In order to correctly represent the information contained in a flow from one location to another, each flow between two points has to be placed very carefully in dependence of its environment. The more extensive the movements, the more lines are included into the map, and the more difficult the representation gets, especially keeping track of a single transition.

2 RELATED WORK

Phan et. al. (2005) developed an automatic flow map layout algorithm. They reduce visual clutter by merging edges which share the same destination. Furthermore they use intelligent distortion of locations and edge routing. The algorithm produces a single layer for each node. The resulting layers are then overlaid on the map. Since all layers are handled/treated independently from each other, intersections will occur. The final map usually suffers from heavy visual clutter which decreases its effectiveness.

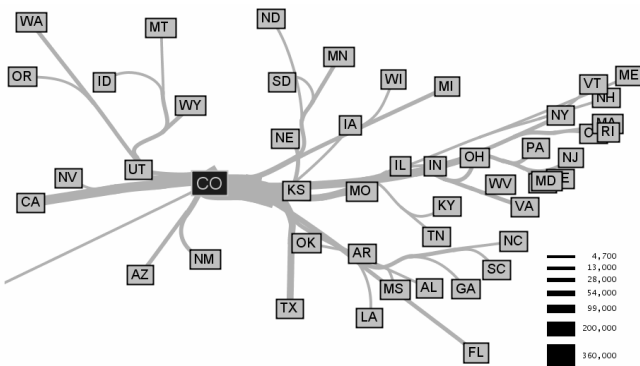


Figure 1: Outgoing migration map from Colorado from 1995-2000, generated by the algorithm without layout adjustment or edge routing. (Phan et. al. 2005)

Glennon et al. (2005) developed a GIS Flow Data Model based on the Flow Mapper by Waldo Tobler (1987). The tool has no cartographic requirements and creates only straight-line flow maps. The straight lines can overlay each other which leads to a loss of information. Dobkin et. al. (1998)

present an edge-routing algorithm to layout a graph on basis of shortest-path calculations, resulting in the same disadvantages as the method of Phan. There are quite a lot of examples of maps representing circular arcs in the public press or other literature (e.g. flight plans). These are well known samples for many users. The advantage of this representation is the consistency and simple geometry of the lines.

3 AUTOMATIC FLOW MAP LAYOUT

The most important characteristic all introduced concepts have in common is that they eliminate intersections between edges and non-incident nodes within the map.

A simple complexity measurement is defined to prove the effectiveness of our approach. As it is most important to reduce the number of intersections between edges and nodes to zero, an objective complexity measurement (OM) is introduced based on the count of intersections of edges and nodes, because it is the most important to reduce this number to zero.

Our approach follows a two-step strategy: First we try to eliminate the intersections between edges and non-incident nodes. In a second step the edges are examined for intersections with other edges.

3.1 Metaobject-Placement

The solution is to combine the advantages of the presented methods: If many start- and endnodes exist (as in the use case), the straight line maps seem to be a better representation, because the lines are easy to pursue with the eye. But the disadvantages of crossing edges must be avoided as far as possible.

The algorithm starts to generate edges as straight-lines. In a second step it tries to find arrows on basis of circular arcs for each edge which intersects a node directly. In this case, the radius of the arc is modified until there are no more intersections with other nodes, or until the stop criterion is reached. Currently the radius will be reduced in five predefined steps until the new geometry forms a semi-circle (stop criterion). In best case the OM then is zero: there are no more intersections between edges and nodes.

The algorithm tries to find a new geometry for the edge depending on the count of nodes in the neighbourhood which is formed by a circle with a diameter of the straight-line length. The advantage of this approach is that the probability of the new edge geometry having an intersection again is minimized because the side with fewer nodes is chosen.

REFERENCES

- Dobkin, D.P., E.R. Gansner, E. Koutsofios, S.C. North (1998): Implementing a General-Purpose Edge-Router. In: Di Battista, Giuseppe, Eds. *Proceeding Graph Drawing*, pp. 262-271, Rome.
- Glennon, A., M. Goodchild (2004): A GIS Flow Data Model. Unpublished paper.
- MacEachren, M.A. (1982): Map Complexity: Comparison and Measurement. *The American Cartographer*, Vol. 9, No. 1, pp. 31-46.
- MacEachren, M.A. (1995): *How maps work*. Guilford Press, New York.
- Phan, D, L. Xiao, R. Yeh, P. Hanrahan, T. Winograd (2005): Flow Map Layout. *IEEE Symposium on Information Visualization*. Second Edition, Pearson Education Prentice Hall, New York.
- Tobler, W. (1987): Experiments in Migration Mapping by Computer. *The American Cartographer*, 14, 2, pp. 155-163.

