MODELLING SPATIAL INFORMATION

Franz-Benjamin Mocnik
Department of Geodesy and Geoinformation

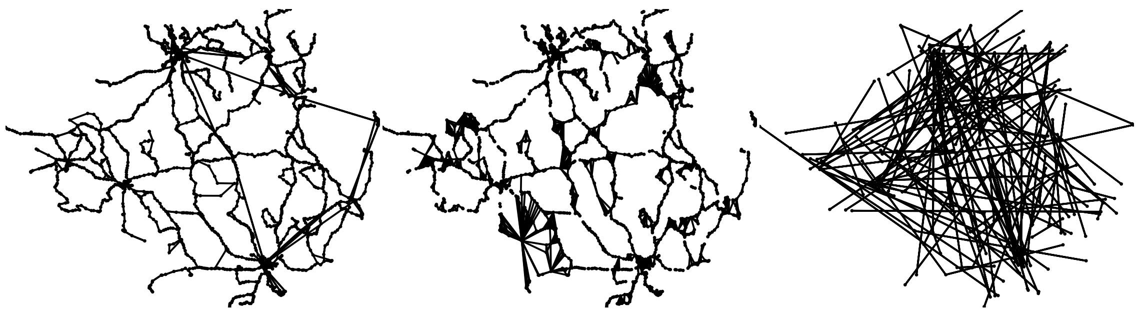
Tobler's law claims that nearer things are more related than distant ones. Based on this principle, a scale-invariant spatial graph model for spatial data is developed. Comparing data with this model allows a software to decide if it has a spatial structure or not.

INTRODUCTION

Spatial information is ubiquitous. It can be crucial to understand the spatial structure of data to solve problems. Yet, there exists no widely used model of spatial information in general (and not just of space).

THE MODEL

Things in the same neighbourhood (a core concept of space) are related much likely:
Tobler's first law of geography. [1] *Everything is related to everything else, but near things are more related than distant things.*
Based on this law, I introduce a spatial graph model that incorporates some central properties of spatial information, including scale-invariance and others:
Spatial graph model. *For a given set of points and a parameter $\rho > 1$, we construct outgoing edges for each point in the following way: For a point p, the point with the smallest distance q is identified, and edges from p to every point with a distance smaller than $\rho \cdot |q - p|$ is introduced. The resulting direct graph is called spatial graph model.*
As can be seen in figure 1, the spatial graph model connects only neighboured points in contrast to random graphs and is thus visually similar to the original data.



(a) Gr. representation of the timetable (b) Spatial graph model (c) Random graph

Figure 1. The spatial graph model for *Länstrafiken Sörmland* (640 stops) in Sweden, compared with a graph representation of the timetable and random graph with the same stops.

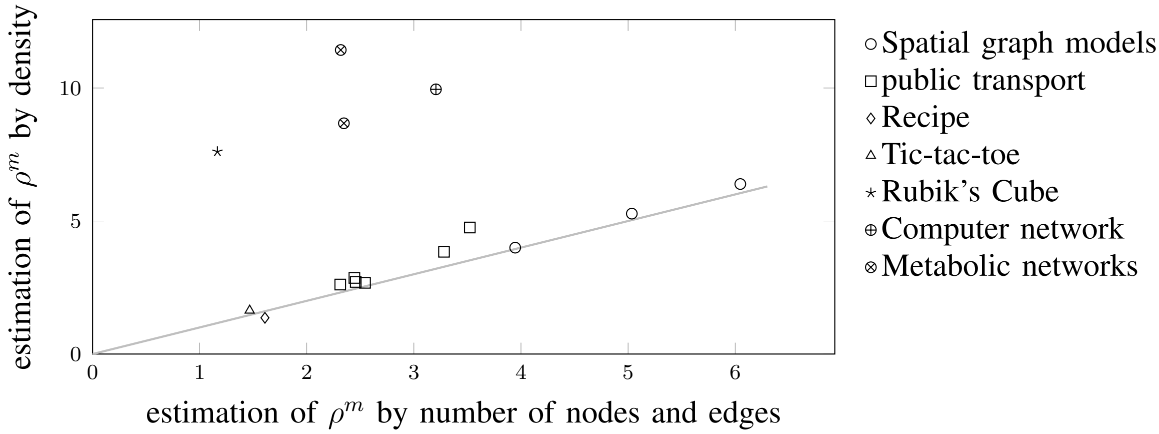


Figure 2. Comparison of two different estimations for ρ^m: the spatial graph model and public transport have a spatial structure; the recipe and Tic-tac-toe have a temporal structure

Spatial or not?

Properties. *The spatial graph model (with n points, uniformly distributed in a d-dimensional region) has the following properties for $n \to \infty$:*

1) *We expect ρ^d to equal the number of edges divided by the number of nodes.*
2) *We expect the density of a subgraph with m nodes to have density $\rho^d/(m-1)$.*

The first property allows us to directly, the second one to indirectly determine an estimation for ρ^d. For spatial graphs, we expect both estimations to be equal. As can be seen in figure 2, comparing both estimations allows us to distinguish between non-spatial and spatial (or temporal) data which is located on the diagonal. We can even distinguish temporal from spatial data as the value of ρ^d is smaller for temporal data (time is 1-dimensional) and larger for spatial data (space is usually 2- or more-dimensional).

Conclusion

Space causes some patterns and structures to be found in spatial information. I have presented some important properties of spatial structures, and demonstrated how they can be detected. This offers numerous possibilities: data can be automatically categorised according to its relation to space, some semantics can be added, complex search engines like Wolfram Alpha can be enhanced in spatial aspects, algorithms can be designed and modified to take care of the spatial structure to be much more efficient, etc.

Acknowledgment

The author would like to thank Andreas Frank for the discussions and his support.

References

[1] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic Geography*, vol. 46, p. 234–240, 1970.