Available online at www.sciencedirect.com
ScienceDirect
Computers, Environment and Urban Systems
31 (2007) 91–106
Computers, Environment and Urban Systems
www.elsevier.com/locate/compenvurbsys

Assessing different approaches to visualise spatial and attribute uncertainty in socioeconomic data using the hexagonal or rhombus (HoR) trustree

J. Kardos *, G.L. Benwell, A.B. Moore

Spatial Information Research Centre, Department of Information Science, University of Otago,
P.O. Box 56, Dunedin, New Zealand

Received 30 April 2004; received in revised form 21 July 2005; accepted 21 July 2005

Abstract

The visualisation of uncertainty for socioeconomic data is a necessary task so that map users may make well-informed and unambiguous decisions. This is particularly the case for choropleth mapping, which is well-known for implicitly highlighting the modifiable areal unit problem (MAUP), resulting in often spurious choropleth areal unit boundaries, not defined from geographical or socioeconomic phenomena. This paper addresses this form of spatial uncertainty by diffusing fixed and arbitrary choropleth boundaries and simultaneously, traditional *attribute* uncertainty through visualisation. The paper presents alternative ways of using the Hexagonal or Rhombus (HoR) quadtree tessellation (termed the *trustree*) for this purpose. Attribute uncertainty is expressed via tessellation size, and choropleth spatial boundary uncertainty via tessellation location. An Internet survey was conducted to assess the *usability* and *effectiveness* of six different trustree methods applied to New Zealand 2001 census data. Results are given and show that a transparent HoR trustree overlaying a choropleth map shown adjacent to the original choropleth is the most usable and the most effective way to express spatial and attribute uncertainty. Also, the HoR value-by-area display with its original areal unit boundaries overlain was almost as effective and usable. Future research should focus on assessing real-world map participants using a variety of uncertainty visualisation methods.

---

* Corresponding author. Tel.: +64 3 479 8142; fax: +64 3 479 8311.
E-mail addresses: jkardos@infoscience.otago.ac.nz (J. Kardos), gbenwell@infoscience.otago.ac.nz (G.L. Benwell), amoore@infoscience.otago.ac.nz (A.B. Moore).

0198-9715/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2005.07.007

Keywords: Uncertainty; Visualisation; Tessellation; Attribute; Choropleth; Spatial; Boundary

1. Introduction

Uncertainty in spatial (and attribute) data arises from a mismatch of that data with the real-world phenomenon it is trying to represent. This is inevitable and can be expressed in many forms, from the systematic error of a GPS measurement to the vague boundaries associated with certain natural land cover types (see Zhang & Goodchild, 2002; for a comprehensive overview of uncertainty). These two instances are examples of the state-of-the-art, employing current knowledge and technology, and are abstractions or *generalisations* of reality.

This paper looks into the hidden uncertainty associated with choropleth maps, and examines a visualisation technique—the quadtree hierarchical data structure (and variants)—in particular its ability to reveal uncertainty to decision makers. There is a different kind of generalisation at play with choropleth maps, a spatial and attribute data aggregation to polygons that is introduced despite the fact that the original point data operated on may be closer to reality.

In the case of census data this is a necessary process to observe confidentiality of individuals' information. However, the boundaries of the polygons that are chosen tend to be, for the most part, not semantically meaningful, yet they are displayed with an implied precision and accuracy. Although the principle of autocorrelation underlies the aggregation, with homogeneity assumed throughout a polygon (MacEachren, 1995), the implied continuity fails at or near to the polygon boundaries. Aggregated attribute values on either side of such a boundary can differ greatly, with no geographic discontinuity, such as a river or administrative district boundary, to explain it rationally (though sometimes a natural or human-made geographic barrier does form the polygon boundary). Of course, there are an infinite amount of polygons that can be formed in this way, each combination yielding a different map. This is an expression of the Modifiable Areal Unit Problem (MAUP—Openshaw & Taylor, 1981). This artificial scenario in the spatial plane is compounded by spurious choice of attribute class limits or breaks.

To explain the manifestation of uncertainty, and to demonstrate geovisualisation in a decision support context, take an example: a large supermarket chain-store cannot decide where to open a new store, with possible sites in many viable city districts. A random survey of 50 houses in each predetermined city district was conducted. The results were collated and inserted into a GIS, using all 50 participants to obtain a generalised value for each surveyed district. Unfortunately, when viewing the survey results two main assumptions will be prevalent: (1) all people living in each separate city district must have a similar response to the 50 survey participants, and (2) the city district boundaries define a natural survey break between areas. These two assumptions will not provide a truthful information display and will lead to *uncertainty* (both attribute and spatial). It would be natural for the survey result to be expressed as a choropleth map or visualisation, yet that map will not conventionally show the uncertainty possessed by the data—if low-uncertainty data was placed alongside high-uncertainty data on a map, there would be no visual property that could be used to distinguish between the two groups of data (that is not to say that uncertainty visualisation techniques do not exist—see Section 2). Therefore the uncertainty generally only becomes apparent after a decision has been made and an unexpected outcome arises based on that decision.