Computers, Environment and Urban Systems 33 (2009) 349–362

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

A functional perspective on map generalisation

Omair Z. Chaudhry [a,*], William A. Mackaness [a,1], Nicolas Regnauld [b]

[a] School of Geosciences, The University of Edinburgh, Drummond St., Edinburgh EH8 9XP, UK
[b] Ordnance Survey Research Labs, Ordnance Survey, Romsey Road, Southampton SO16 4GU, UK

ARTICLE INFO

Keywords:
Map generalisation
Multiple representations
Data modelling
Database enrichment

ABSTRACT

In the context of map generalisation, the ambition is to store once and then maintain a very detailed geographic database. Using a mix of modelling and cartographic generalisation techniques, the intention is to derive map products at varying levels of detail – from the fine scale to the highly synoptic. We argue that in modelling this process, it is highly advantageous to take a 'functional perspective' on map generalisation – rather than a geometric one. In other words to model the function as it manifests itself in the shapes and patterns of distribution of the phenomena being mapped – whether it be hospitals, airports, or cities. By modelling the functional composition of such features we can create relationships (partonomic, taxonomic and topological) that lend themselves directly to modelling, to analysis and most importantly to the process of generalisation. Borrowing from ideas in robotic vision this paper presents an approach for the automatic identification of functional sites (a collection of topographic features that perform a collective function) and demonstrates their utility in multi-scale representation and generalisation.

1. Introduction

From a database perspective a map can be viewed as a set of geometries rendered via a look up table of symbols, together with associated text. From the human view however, the map reflects a collection of concepts, often grouped or connected together in clearly defined ways – as a result of both physical and human processes. Thus the viewer does not see a twisty blue line, but sees a meandering river as it snakes through the delta on its way to the sea. The viewer does not see a dense collection of small angular polygons, but sees a collection of buildings, performing many different but related tasks that all contribute to the idea of urban space and the city. The cartographer takes advantage of the viewer's interpretative view when they come to generalise at smaller scales (Mackaness, 2007). An icon of an aeroplane substitutes the multi-storey car parks, hangars, terminals, aprons, and runways that make up an airport. The letter 'H' replaces the clinics, car parks, outpatient facilities, heating plant, and wards that typically constitute our understanding of what is meant by 'Hospital'. And a simple dot with the word 'Jakarta' next to it is used to locate and convey this vast megalopolis. So the key argument around 'functional site modelling' is that the generalisation process would be hugely facilitated by data that were 'functionally tagged' and structured in a similar manner. So that, at changing levels of detail, fine scale phenomena could be grouped together according to their collective function. In this manner, all the features that constitute an airport, a hospital, or a city could be replaced by a more generalised form (an aeroplane symbol, the letter 'H', or a dot with the word Jakarta next to it). Thus we can define a functional site as a collection of objects (natural or anthropogenic), usually in proximity to one another, which collectively perform a specific function. The term 'site' is considered to cover a range of geographies. Just a few examples might be: schools, retail parks, business districts, airports, docks, or cities – each of which we can associate a particular function or set of functions.

We argue that from a multi representational database perspective (Mustière & van Smaalen, 2007) we need to make explicit the nested connections that exist between these functional sites, as well as the components that constitute them. For example the car park, platforms, station building and shunting yards that constitutes a 'railway station' or the playing fields, classrooms, and sports hall, that constitute a 'school'. In this manner we can deliver appropriate visualisations of these functional sites or their 'components', at different levels of detail. Such an approach can also facilitate automated text placement (Barrault, 1995; Zhang & Harrie, 2006) for example where text is used to convey the extent, importance or function of something. Furthermore, by linking a functional site with its components we can automate the update process. For example as more suburban houses are built at the edge of the city, we can automatically update the extent of the city

* Corresponding author. Tel.: +44 (0) 23 8030 5044; fax: +44 (0) 23 8030 5072.
E-mail addresses: Omair.Chaudhry@ordnancesurvey.co.uk (O.Z. Chaudhry), William.mackaness@ed.ac.uk (W.A. Mackaness), Nicolas.Regnauld@ordnancesurvey.co.uk (N. Regnauld).
[1] Fax: +44 131 650 2524.

0198-9715/$ - see front matter
doi:10.1016/j.compenvurbsys.2009.07.002

boundary, because we have made the explicit link between the function of the boundary and the objects that constitute it.

A further benefit to this functional perspective is the idea that we can explicitly model the relationships between functional sites. For example, if we understand the function of a road, and that of the hospital, then we can define the notion of access points between the two. We can also model relationships between functional sites, such as the service provided by the hospitals to the city. This richer view provides a more intuitive framework by which we might combine third party data. Increasingly third party institutions are using the 'framework data' of National Mapping Agencies to answer themed questions that go far beyond conventional series mapping. These 'themed' questions might relate to evacuation strategies or access pathways to hospitals, to maintenance contracts of large civil installations, or to transportation planning. In these instances we need to precisely know what components constitute a particular site (an industrial estate, a shopping centre, a forestry reserve) – and in these contexts we need a database that supports this functional description of space, and its relationships between sites. In the next section we justify the need for an automated approach to the identification of functional sites. Section 3 characterises functional sites as a precursor to the development of a generic model. Section 4 details the methodology that falls from the framework described in Section 3. Section 5 presents results from the implementation, and the results and evaluation that point to future work.

2. Database requirements

Rich attribution of and the entities that comprise the functional sites is one way of explicitly modelling the functional site membership. For example we can use taxonomic and partonomic labelling that explicitly states that (say) (a) 'this is a playing field' and (b) 'it is part of Newbury Grammar School'. The problem with creating such multi representational databases is that in the past field surveyors have collected data in anticipation of their *cartographic* representation at a specific scale – where there has been no requirement to precisely and consistently categorise features, and where membership is self apparent (for example, that a terminal is part of Heathrow Airport) thus obviating the need for this type of attribution. This cartographic view has resulted in databases that (1) are poorly attributed; (2) are inconsistent; (3) use mixed approaches to geo-referencing and address labelling; and (4) do not make explicit the obvious association between the functional components and their parent site. In other words they do not support multiple representations. In the case of Ordnance Survey (the National Mapping Agency of Great Britain) data, we have cases where the address of the terminal does not contain the name of the airport, and the attribution associated with a car park is vague such that it makes it very hard to ascertain which functional site it serves, or even that it is a car park.

2.1. Solution

If each object was richly and consistently attributed, it would be a simple matter to determine the partonomic structures and functional descriptions of higher order concepts through the interrogation of those attributes. Using humans to attribute the database would be a huge undertaking; every object would have to include descriptions of multiple partonomic memberships. Given the enormity of this task, it is highly desirable to seek an automated solution to this problem – in other words to automate as much as possible, the process by which components are associated with a particular functional site. Because existing attribution may be inaccurate, and because the composition of functional sites can be complex (for example components may be dispersed in the way that city universities often are), it is unlikely that a completely automated solution is achievable, and therefore any solution should anticipate the involvement of a human in validating the partonomy and extent of a functional site.

3. A classification of functional sites

Functional sites are typically made up of components of different classes. For example a refinery might be made up of port facilities, storage vessels, bounded ponds, and tanker depots. Functional sites may be nested. For example fish processing, boat maintenance and sea rescue functions – can all exist within the functional site 'port'. Some sites have very crisp boundaries (the airport delimited by a security fence) whilst others are much more open to interpretation (for example the extent of suburbia or a mountainous area) (Smith & Varzi, 2000). There is considerable variation in the areal extent of functional sites (from a small railway station to a massive city), and this variation can exist even where they serve the same function (for example the educational function of a small primary school as compared with a large university campus). All these factors complicate the process of identification and validation. We discuss such factors later in this paper. Further, the task of classification is made difficult by the fact that our notion of a functional site varies with context and geography. One might ask how was the epithet 'Lake District' justified? Or at what point is a city deemed to have a 'financial district'? (indeed what is meant by 'District' in each of these cases). Collectively we see that the relationship between functional sites is multi-scaled and complex. Table 1 is a tiny subset of all functional sites that seeks to illustrate the variability of sites (by class and size). It illustrates that there are broad categories for which representation is appropriate at a range of scales (or levels of detail).

3.1. An object ontology framework

It is interesting to note that the human eye is able to examine the map, and often with no textual attribution is able to infer and identify a wide range of functional sites because of the shape of their components and the way they interact with surrounding features. For example dockyards can be said to interact with (or connect together) sea and land based networks. Peruse a map, and one can identify various ways by which you can discern: (1) the type of functional site, (2) its components, and (3) its geographical extent. From such observations, we can begin to identify qualities that can potentially be used to identify and discern the extent of functional sites in an automated context (Table 2).

Various research has explored ways of automatically enriching the description of entities. Thomson and Béra (2007, 2008) used descriptive logic reasoning to aid in the classification and enrichment of OS MasterMap® features into higher level concepts (terraced, semi-terraced and detached houses). Another project at Ordnance Survey used ontologies to aid in the identification of farming land data in OS MasterMap (Kovacs & Zhou, 2007). Lüscher, Weibel, and Burghardt (2008), Lüscher, Weibel, and Mackaness (2008) used an algorithmic approach for the automatic identification of instances of similar higher order concepts. The advantage of using an algorithmic approach being that it can handle fuzzy membership and uncertainty, as well as cope with large volumes of data.

When it came to modelling the relationships between function and form, we noted interesting parallels with work in robotic vision systems and understanding. For example, Wang, Kim, and Kim (2005) describe an 'object ontology framework' in which the vision system of the robot seeks to understand the function of an