TOWARD CADASTRAL LEVEL CARTOGRAPHIC ANALYSIS USING MULTI-SCALE SPATIAL DATA

David J. Cowen
W. Lynn Shirley
Patrick Noonan
Chris Wiesner
Humanities and Social Sciences Computing Lab
University of South Carolina
Columbia, SC 29208 USA
FAX: 803-777-7849 E-mail: cowend@hsscls.hssc.scarolina.edu

ABSTRACT

A major cartographic problem facing many jurisdictions is how to integrate nationwide demographic data with large scale cadastral databases. In the United States this problem is exemplified by the need to display and analyze decennial census information along with multipurpose cadastre data. This paper describes a systematic approach for developing a unified database that incorporates municipal and Federal geographic data sources. The first stage of the approach involves attribute matching from the two data sources. This provides a basis for combining Census based demographic data with data relating to property records such as land use. In the second stage, street center lines are automatically generated from property right-of-ways using a "vector to raster to vector" conversion methodology. This procedure provides a consistent basis for improving the positional accuracy of the 1:100,000 scale TIGER line files. It also creates a topologically consistent street network that can be used for navigation and routing.

1. INTRODUCTION

A common problem facing urban planners and other officials who deal with local government issues is the need to integrate multiple sources of jurisdictional information. In the United States this typically relates to city, county, state and federal administrative units that have overlapping jurisdictional authority. In the multi-purpose cadastral environment at scales such as 1:1,200 the individual property ownership units represent the most accurate large scale administrative units [1]. The parcel level files provide a detailed partitioning of property features based on legal ownership. A parcel level data base is legally maintained by tax records that include attributes of parcels such as value, ownership, zoning and land use. At the same time, the Federal Census is based on the TIGER line files at a scale of 1:100,000 [2]. The decennial census of population and housing enumerates the number of houses and population at the block level. In other words, the multipurpose cadastre is concerned with characterization of land within a block while the Census is interested in aggregated information about people and houses that characterize the entire block. Although both data sources provide useful information about people and land use for the same geographic feature they represent the feature in very different ways and scales [3]. The combination of these two sources of information provides a valuable basis for analyzing demographic and socio-economic conditions with property and land use information. For example, the assessor does not know how many people live in a block but is able to determine where the residential dwellings are within the block. Within a spatial data handling context this type of problem is generally associated with multiple scale data sources and the need for generalization of large scale boundaries [4] or adjustment of boundaries of one set of features to another set of the same features with more precise or spatial attributes [5].

The specific concern of this paper is to demonstrate a series of automated procedures that can be employed to build relationships between multipurpose land information system parcel level data and the US Bureau of Census TIGER data. In addition to gathering different types of information about blocks the Bureau of the Census and the local assessor differ in how they represent these same feature (fig. 1). From the multi-purpose cadastre viewpoint parcels are separated typically from the pavement at the right of way.

Therefore, a block level feature in the cadastre can be formed by dissolving all the parcel boundaries within the right of way. The pavement between the right of ways has no attributes, therefore, the blocks form detached polygons surrounded by pavement, water etc. The Census Bureau TIGER line files represent a city as a network of street center lines. This topologically complete network of one cell features bounds a set of two cells that constitute census blocks. Therefore, in this representation, there is no difference between the pavement and the property and the block boundary is the street center line. From a spatial data handling perspective the need to integrate these two representations requires two things: (1) form a relational linkage between the two representations of the blocks that combines the attributes of both data sources (2) create a set of street center lines from the larger scale parcel level data that creates a network with greater positional accuracy than the 1:100,000 scale TIGER data but maintains the Census attributes.

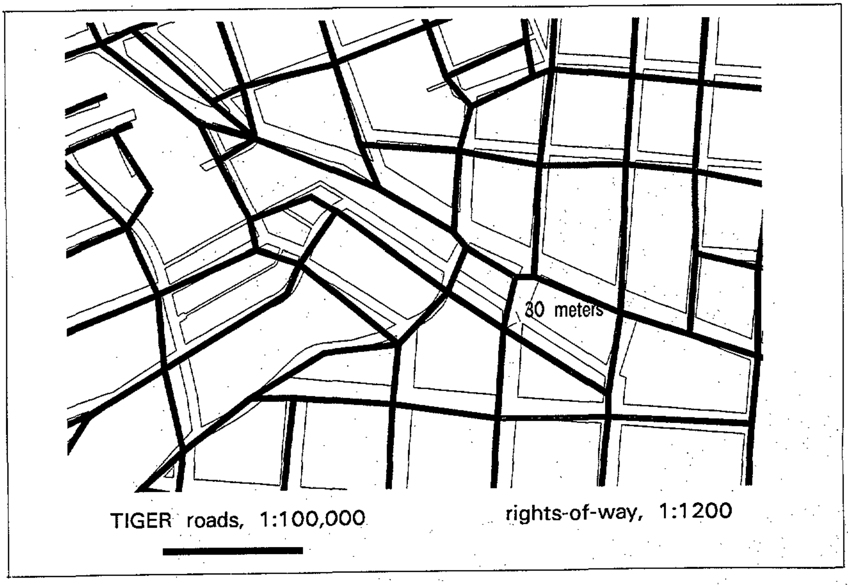

Fig. 1 Right of way blocks versus TIGER blocks

2. ATTRIBUTE LINKING

The joining of attribute data from two sources is a map overlay problem. The parcel level data are dissolved to create the set of blocks based on the right of ways. From the TIGER line file data it is possible to generate a centroid or label point. These points can then be overlaid on the right of way blocks. Even allowing for scale differences between the two sources all but a few of the census block centroids will fall within the correct right of way. Some visual editing can fix the remaining problems. Once this is accomplished all the Census block level data can be relationally linked to the right of way blocks. The parcel level property data can be aggregated within the block to create new attributes such as proportion of land area in different price, land use or zoning categories. Once this is accomplished it is possible to generate cartographic products that incorporate both sources of data. For example, it is possible to create graduated pies for the racial composition of the 1990 block level population that can be portrayed with the parcel level land use (fig. 2). This map is of great interest to the local planners and social service agencies.

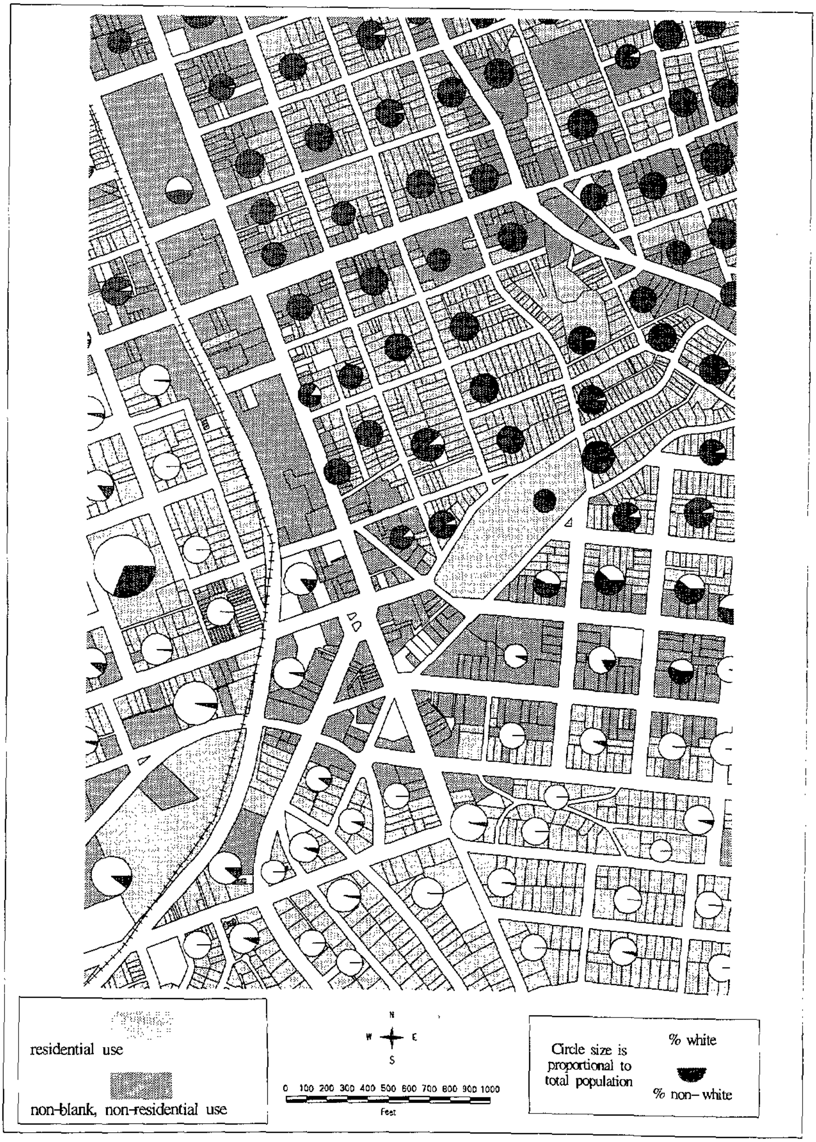

Fig. 2 1990 Census block level racial composition and parcel level land use.

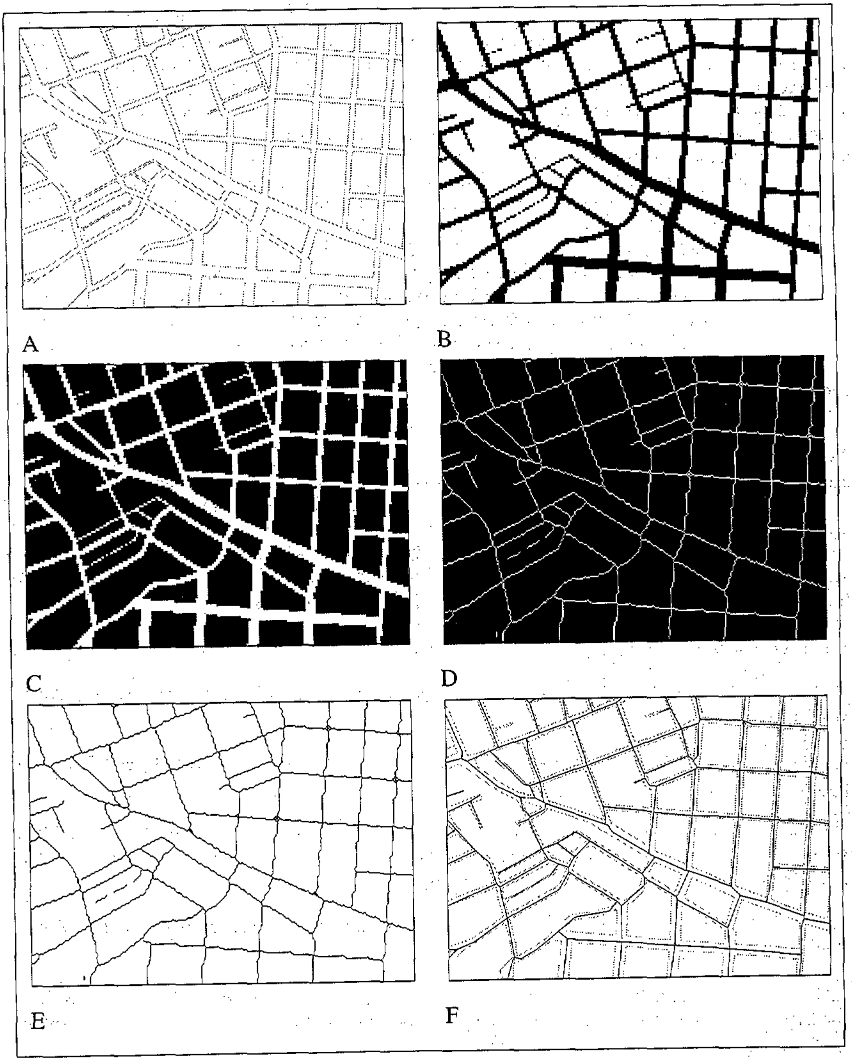

Fig. 3 A-F *Vector to raster to vector center line creation procedure.* A. Vector right of way blocks. B. Binary grid representation of land and pavement. C. Inversed binary representation. D. Thinned grid pavement. E. Initial vector center lines. F. Postprocessed vector center lines.

3. CENTER LINE GENERATION

The issue of creating more accurate street center lines is a more complex one than attribute matching. The logic followed in this procedure extends the work of Zhan [6] and is based on a vector to raster to vector conversion process (fig. 3). The first step in the process is to divide the parcel level data into binary categories of land and pavement. The binary polygon map is then converted into a grid cell representation. The goal of the grid process is to create a raster image of the map in which the width of the pavement is represented by several cells. In effect, this is analogous to raster image data that would be produced from scanning the vector map. In order to obtain reasonable results it is necessary to experiment with various cell resolutions. In practice, there has to be a sufficient number of cells spanning the street to have the line fit mid way between the two right of ways. The next step is to eliminate any large areas i.e. water bodies that should not be considered part of the street network. This process involves the use of standard cartographic modeling functions that categorize all the cells within a given radius of a cell. Large homogeneous areas are eliminated while cells with highly diverse values are found at intersections. Once the binary grid cell map is created the street center lines are defined as the linear skeletons of the raster pavement. This step is accomplished by thinning the grid structure and then converting to vectors. The resultant lines tend to have wave like shape and do not generally meet at sharp "T" junctions. These problems are automatically addressed by postprocessing the features with vector based GIS functions that generalize, spline and straighten linear features. These functions are now found in commercial GIS software aimed at raster to vector data conversion.

The resultant linear features form a topologically consistent network. The complete chains of this network can be assigned the unique TIGER one cell number and the attributes such as census geographical areas, street names and address ranges can be relationally linked to the corresponding features. This procedure can be automated but there is no assurance that there is always a one to one correspondence between the features. An alternative approach is to geometrically adjust the TIGER line segments to the corresponding links in the new network. This alternative is conceptually simpler and meets the long term needs of the Bureau of the Census and the creation of a National Spatial Data Infrastructure.

REFERENCES

[1] National Research Council. 1980. Need for a Multipurpose Cadastre. National Academy Press: Washington.

[2] Marx, R. 1990. The TIGER System: Yesterday Today and Tomorrow. Cartography and GIS 17(1) pp, 89-97.

[3] Egenhofer, M., Clementini, and P. Di Felice, 1994. Evaluating Inconsistencies Among Multiple Representations. Advances in GIS Research, Proceedings of the Sixth International Symposium on Spatial Data Handling, Vol. 2. pp, 901-920 Taylor and Francis: London.

[4] Buttenfield, B. and McMaster, R., 1991. Map Generalization: Making Rules for Knowledge Representation, Longman: UK .

[5] Cowen, D. and Shirley W.L. 1990. An evaluation of the use of Digital Line Graphs and TIGER Files for use in Multipurpose Geographic Information Systems. Proceedings Fourth International Symposium on Spatial Data Handling, vol. 2, pp. 621-631, Zurich.

[6] Zhan, C.,, 1993. A Hybrid Line Thinning Approach. Proceedings Auto Carto 11, pp, 396-403.