A GENERAL APPROACH TO MAP CONFLATION

Anthony E. Lupien
William H. Moreland

Environmental Systems Research Institute
Redlands, California 92373

Abstract

Map conflation is approached as two generic geoprocessing problems, feature alignment and feature matching. Techniques for their solution are presented. By taking a general approach, match relationships between maps can be exploited by relational database functions of a geographic information system to create various products. Examples are presented from the Arc/Info implementation.

Introduction

Lynch and Saalfeld define conflation as a "combining of two digital maps to produce a third map file which is 'better' than each of the component source maps (1985)." In their pioneering conflation experiments at the Census Bureau they developed a method of conflating maps from two particular series, GBF/DIME and USGS scanned maps. Their reports of this development provide a vocabulary for conflation research and numerous suggestions for implementing conflation in geographic information systems (GIS) software.

Many GIS applications can be served by a conflation capability. These include transfering attributes from old to new street network maps, insetting maps into existing spatial databases, unifying maps of different geographic features and evaluating maps' coordinate distortion. The edge-matching problem can even be treated as a subset of conflation.

This paper presents the method we have developed for conflation, which is treated as the solution of two generic geoprocessing problems within a relational database environment. Solution of the feature alignment and feature matching problems, it is shown, creates database relationships between two maps that may be used to generate any desired "third map."

In conflation, the "best" elements of two maps are combined to form a third. Definitions of what is "best" in a map vary with application, but whatever conflation product is sought conflation is generally the same problem: Identifying features in two maps that represent the same earth feature or are in the same location, then selectively merging features and attributes of both into a third.

Conflation as Two Generic Problems

Conflation requires solution of two generic geoprocessing problems: Feature alignment and feature matching. By "features" we mean the lines, points and polygons that comprise topological digital maps, or "coverages".

Features are aligned by transforming the coordinates of those in one coverage to fit another's. Rubber-sheeting techniques are used to solve this problem. Our implementation of rubber-sheeting involves simple extensions to an existing graphic editing program and a coordinate transformation program developed for modeling triangulated irregular networks (TIN). A data structure called a "link", which can be treated as both an arc and a point, is used to bridge the requirements of the two programs.

Once features of two coverages have been aligned they may be matched. Rosen and Saalfeld report success with iterative applications of topological criteria that compare features based on their relationships to others in the same coverage (Rosen and Saalfeld, 1985). Our research, however, suggests that simple feature-to-feature distance measurements between coverages suffice. In the Arc/Info implementation most map features that should ideally match generally do, without significant false-matching.

Solving the Feature Alignment Problem

We would like to believe our success in conflation with simple feature-matching rules is due to the strength of our feature alignment solution. The first component of this is an interactive graphic system for displaying two coverages and adding "links" to one of them. Each link can be thought of as an arc from a location in the coverage to be transformed to a corresponding location in the other (fig 1). As arcs, these links can be added, deleted, modified and symbolized using common GIS software. Link management capabilities were, therefore, easy to add to the graphic editing subsystem of Arc/Info.

These links are passed to a coverage coordinate transformer, without disturbing the graphics environment. This program treats links as points with delta-x and delta-y attributes and builds from them a triangulated irregular network. From this network two fifth-order distortion surfaces are constructed, one for x and one for y. The distortion surfaces are used to transform coverage features, which are redisplayed (fig 2). A non-linear (bivariate quintic) interpolation algorithm is used to produce smooth surfaces (Chen and Guevara, 1986).

Links are added and transformations run iteratively. Alignment of 500 to 1000 features, a useful working set size (Lynch and Saalfeld, 1985), is controlled with only a few links. Experiments (Lupien, 1986) suggest that one well-placed link per 50 features yields close alignment (fig 3).

After initial alignment further global transformations tend to cause aligned features to drift apart. To avoid this drifting phenomenon subsequent transformations are limited by use of zero-length, "identity", links. An identity link effectively "nails down" a coverage location, preventing it from moving in transformation.

Surrounding an area with identity links localizes the impact of links within the area (fig 4). Features inside the boundary will move; those outside will not. Identity links may also be used to limit transformation to a subset of coverage features selected by some attribute. After initial alignment and matching, transformations are often limited to unmatched features by placing identity links along matched features (fig 5).

The software for building TINs and using them to generate surfaces was, like the graphics editor, developed for Arc/Info without anticipating its usefulness in conflation. But by structuring the software as generic modules, and by recognizing links to be a conceptual bridge between graphical arcs and points of distortion, the actual development of strong feature alignment tools was made quite economical.

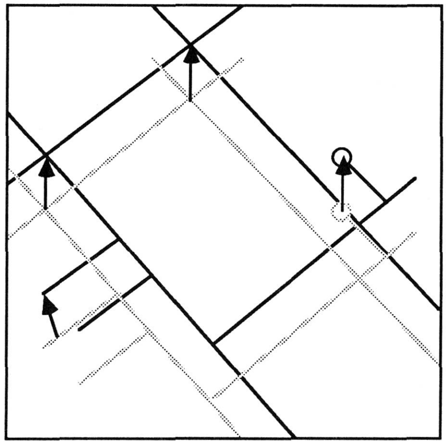

Figure 1

Each link can be thought of as an arc from a location in the coverage to be transformed to a corresponding location in the other.

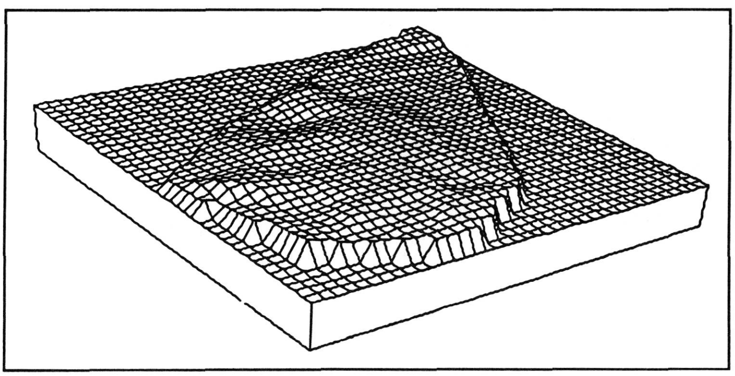

Figure 2a

50 link distortion surface for X

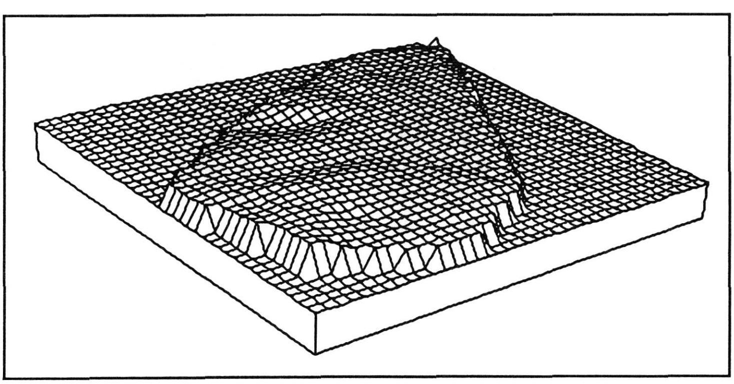

Figure 2b

50 link distortion surface for Y

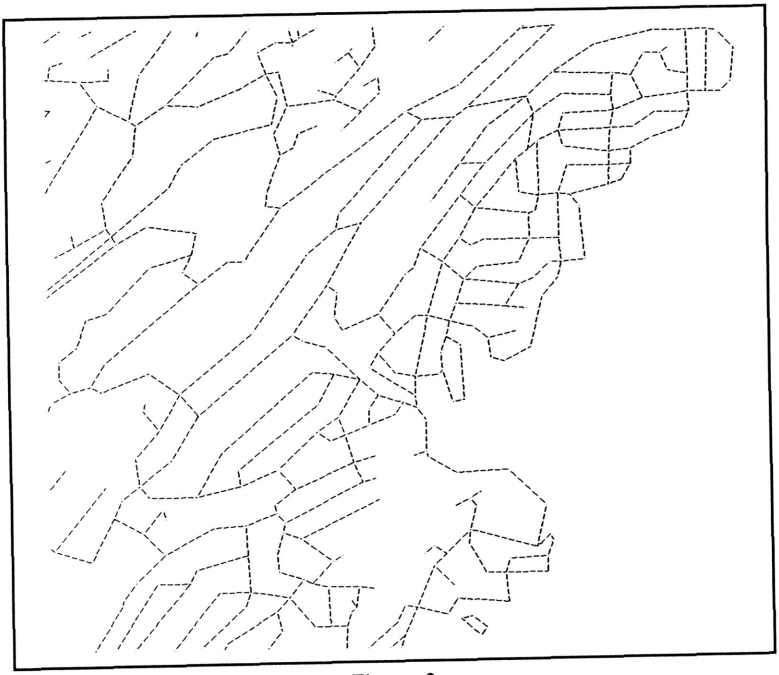

Figure 3a
Partial street network coverage, Atlanta GBF/DIME file.

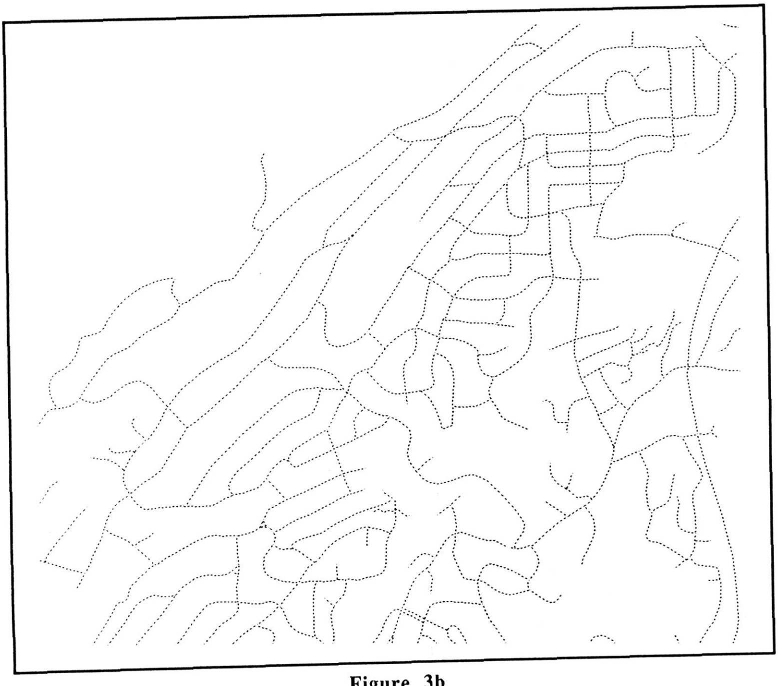

Figure 3b
Partial street network coverage, Courtesy BellSouth Services Co.

Figure 3c
Original coverage alignment (untransformed).

Figure 3d
Coverages aligned with 5 links.

Figure 3e
Coverages aligned with 50 links.

Figure 3f
Coverages aligned with 50 links, then locally adjusted (fully processed).

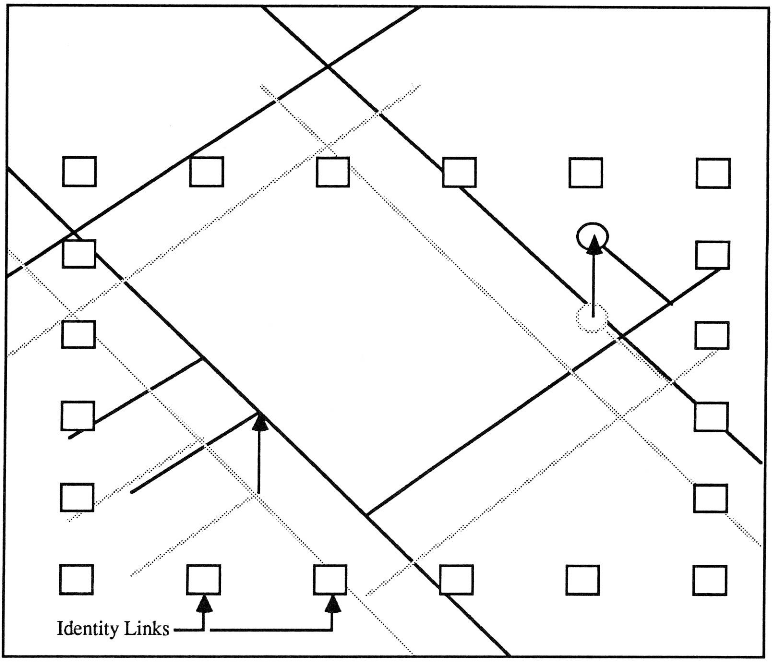

Figure 4

Surrounding an area with identity links localizes the impact of links within the area.

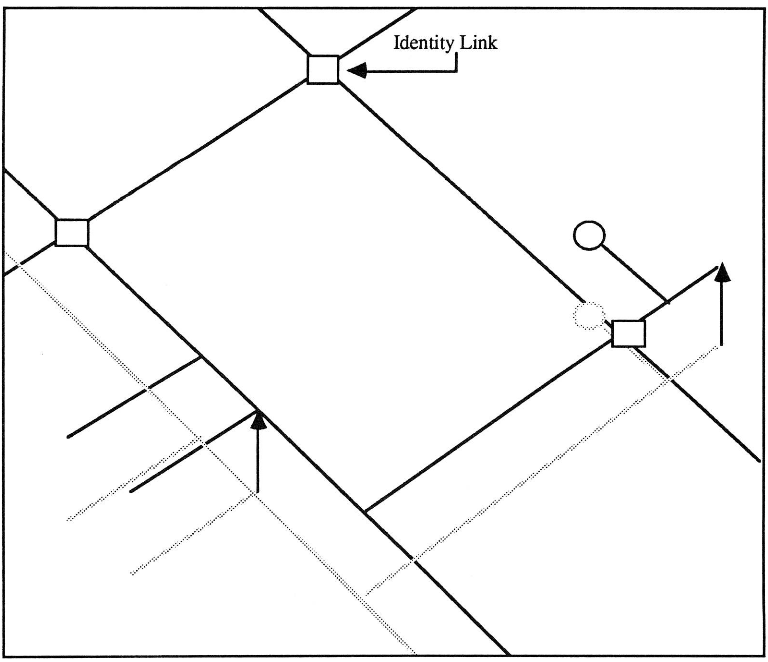

Figure 5

Transformations are limited by placing identity links along matched features.

Solving the Feature Matching Problem

Simple geometric distance calculations may be used to identify matching features in two closely-aligned coverages (fig 6). These calculations are computationally effecient, which is particularly desirable because of conflation's iterative nature. The criteria presented below are used to match point and arc features. In Arc/Info, polygon feature locations are represented by points and are best matched using point-in-polygon testing.

A point feature in coverage A is matched if only one point in coverage B is within a specified tolerance distance. This tolerance is relaxed in successive runs that follow alignment iterations, as suggested by Rosen and Saalfeld (1985).

Arc features are matched in the same way, except that distance calculations are made from each arc vertex, including its from- and to-nodes, to the nearest point along candidate arcs. If each vertex of an arc in coverage A is within tolerance of an arc in coverage B, and if this is true for only one arc in B, then a match is recorded.

These match criteria produce many-to-one relationships, which are sometimes desirable. To produce one-to-one matches only, the procedure is run in both directions (coverage A to coverage B, then coverage B to coverage A), and duplicate pairs extracted from the two lists of matches.

Products of Conflation

Once match relationships have been determined they may be used to transfer attributes between coverages, or to merge the unique features of conflated coverages into a third, or to create maps in which features of two coverages are selectively symbolized.

Many GIS applications, including the Census Bureau/USGS joint project (Lynch and Saalfeld, 1985), seek to transfer attributes from one coverage to another. Often this is done to improve the cartographic quality of an otherwise useful database.

Attribute transfer in the Arc/Info implementation is performed by relating the feature attribute tables of matched coverages to the normalized list ennumerating feature matches. The attribute tables of matching features are then joined.

Other applications seek to merge unique features of two coverages into a third. Many such efforts are aimed at updating a database with new information. Conflation is used to identify new features that have no equivalent in the existing database.

Conflated coverages are merged in Arc/Info by relating their feature attribute tables to the list of matches, then selecting those features in each coverage that have no matching feature in the other. The two selected sets are saved to a third coverage.

Conflation may be used in still other GIS applications to improve the quality of maps produced from several coverages. A map might be made, for example, using features from hydrographic and street-network coverages. Such a map might be produced at a scale that would place a river and a road in the same location. This can be avoided by conflating the coverages and using match relationships to control map symbolization.

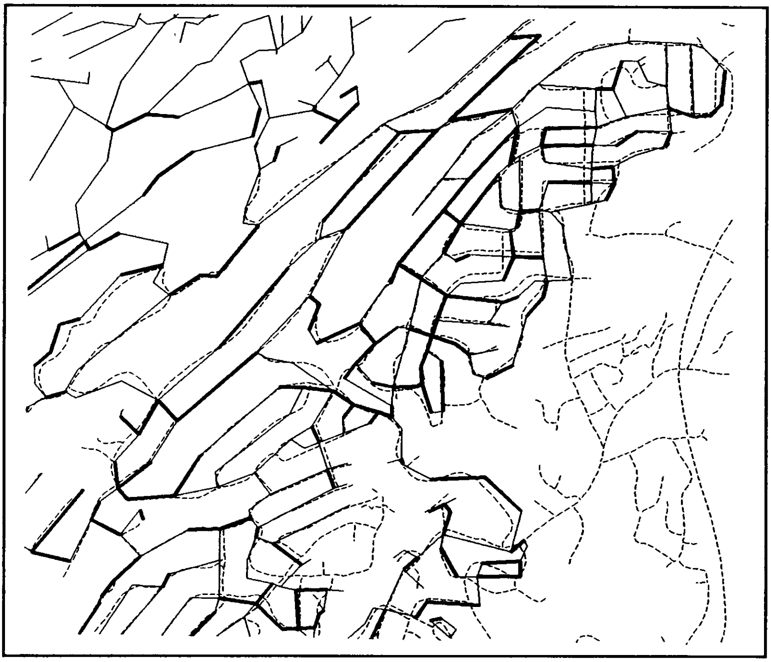

Figure 6a
Matching features after initial alignment.

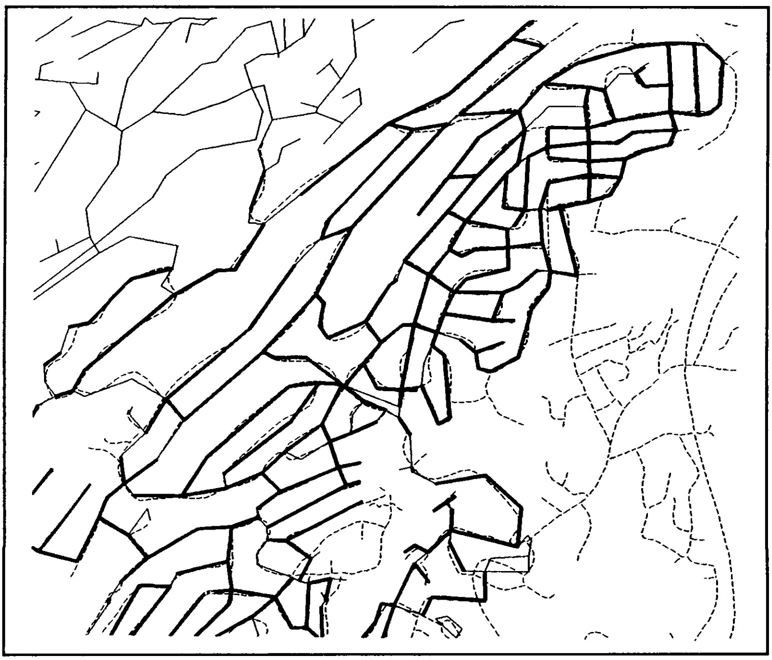

Figure 6b
Matching features after full processing.

References

Chen, Z. and A. Guevara, 1986, "System Selection of Very Important Points from Digital Terrain Model for Constructing Triangular Irregular Networks," *Proceedings, Auto-Carto VIII.*

Lupien, A., 1986, "Conflation with Arc/Info", internal Environmental Systems Research Institute document.

Lynch, M. and A. Saalfeld, 1985, "Conflation: Automated Map Compilation, a Video Game Approach", *Proceedings, Auto-Carto VII.*

Rosen, B. and A. Saalfeld, 1985, "Match Criteria for Automatic Alignment," *Proceedings, Auto-Carto VII.*