Geographic Information System Inventory Data Preparation: Assigning Spatial Properties to Highway Feature Files Using Independent Data Sources

SCOTT A. KUTZ

Many transportation organizations are considering the use of Geographic Information System (GIS) technology for part of their overall approach to managing infrastructure data. They will often find that existing data sets developed and maintained over long periods of time may lack spatial properties needed for inclusion into the GIS environment. Even though these data sets represent the core inventory of features (highways, bridges, signals, and so on), the pre-GIS uses for the data typically never required any type of spatial information. However, most of the data sets did incorporate some type of location reference such as street and address range in urban areas or route and milepost in nonurban areas. This paper discusses the procedures used to assign spatial properties to features in one data set of highway features using an independent data set as its source of spatial information. The highway features were in an existing, nongraphic, address-delimited data set for the city of Chicago. The data set included street centerlines, bridges, viaducts, intersections, traffic signals, and vertical clearance/underpasses. The data set that contained the spatial properties was an independent base map data base of right-of-way, "midlines." The overall approach to reconciling these data sources and assigning the spatial properties of Illinois State Plane X,Y coordinates, network connectivity, and graphic representation to the highway Street Centerline features is discussed. The successful conversion achieved to date has produced a large number of valuable and useful highway features that have been loaded into the geographic data base. The difficulties encountered in matching between these two independent data sources developed by two different organizations are also presented.

The applicability of Geographic Information Systems for Transportation (GIS-T) is well documented in the literature, with three of the more prominent publications cited here (*1–3*). Establishing a geographic base map that serves as a common geographic reference is an essential step for implementing a GIS-T. In the context of this paper, the geographic base map is intended to be some representation of the base highway network. The location referencing system(s) used by an organization relate any stated location (route and milepost, control section and milepoint, street address, and so on) back to the adopted geographic base map. It is through consistent references to this common geographic base that a GIS-T is able to integrate information that may be physically stored in several different data bases (*4*).

The meaning of the "geographic base map" will be different for each implementing organization. The source and content of the data selected to form the "base" can be expected to vary with the organization. These variations are typically in response to differing needs for such factors as accuracy, required highway classes, the need to represent divided highways or multi-tiered highways as a single or multiple centerlines, and length of time needed to acquire, create, or correct inaccuracies in the data. Some possible sources from which base map data is derived include Topographically Integrated Geographic Encoding and Referencing (TIGER) files, aerial photography, or locally digitized data from original hardcopy source documents. Content variation often takes the form of determining which classes of highway are included versus those that are excluded from the data base. One approach may be to include any class of highway for which the implementing organization has some type of responsibility. This approach works well from the Federal or State perspective. However, it can result in some conspicuous gaps in network coverage when this approach is used at the local level.

It is often the case that any combination of problems can arise when actually creating a geographic base map. Some of the problems encountered include: difficulties in associating geographic coordinates to highway segments delimited with control section and milepoint references (*5*); inconsistencies between the locations of network links and road segments in a region (*6*); and difficulties when matching two data sets that are both encoded at the street block level in an urban area (*7*). Another problem can be that the organization does not have too few data but may actually have too many data. This is particularly a problem when redundant data sets exist. For example, the city of Los Angeles was identified as having two different versions of its 300,000-link street network during discussions at a recent GIS-T workshop (*8*).

This paper discusses a project that, although successfully converting a large portion of its data, has encountered its own set of problems when trying to construct the geographic base map for the street network in the city of Chicago. The problems are conceptually similar to the types of problems described above (*5–8*).

Constructing the geographic base map for the Chicago Department of Transportation (CDOT) entailed use of the existing public rights-of-way base map (the official city base map controlled by the Department of Planning and Development, DPD) as the basis for assigning spatial properties to the CDOT street inventory. The CDOT street inventory has been developed over many years as a tabular data set containing street and address range information but without the spatial properties of Illinois State Plane X,Y coordinate data, network connectivity data, or graphic representation data. Without these spatial properties, the CDOT data could not be integrated into a GIS data base.

Camp Dresser and McKee Inc., The Sears Tower, Suite 450, 233 South Wacker Drive, Chicago, Ill. 60606-6306.

The DPD Base Map features represented public right-of-way midlines ("centerlines") that did have spatial characteristics and also contained street name and address range data. However, data sets created and maintained by different organizations with different missions typically are not the same because the organizations each have their own approach to data structures and information management objectives. For example, the perspectives of the two organizations relative to the "streets" are somewhat different. These different perspectives are contrasted in Table 1. The resulting differences between the two data sets being discussed here complicated the process of matching the highway features to the Base Map right-of-way features.

This work was done as part of creating a Physical Inventory data base for the Chicago Citywide Infrastructure Management System (CWIMS), a GIS-based decision support system being developed for planning and coordinating capital improvement projects. Constructing the Physical Inventory portion of the CWIMS data base requires merging data from independent sources, mostly from various city departments that have developed these data sources independently over the years. Specifically for the CDOT data, the process involved the use of address and address range matching. Various techniques were attempted before adopting an address range-matching approach that yielded the highest match rate.

PROJECT OVERVIEW

The CWIMS project was started in May 1992, and its first release is scheduled for June 1996 (followed by a 1-year warranty period).

Objective

The objective of the project is to implement a computerized data base, mapping, and decision support system for planning and coordinating capital improvement projects. It is a GIS that will improve infrastructure planning capabilities within the city's public rights-of-way. CWIMS is intended to improve the city's ability to perform the following tasks:

- Identify and coordinate capital improvements to the water, sewer, street, and bridge infrastructure;
- Allocate capital resources for more effective management and maintenance of the $30 billion infrastructure network; and
- Identify potential infrastructure problems/conflicts before they become critical.

The CWIMS project is being performed by Camp Dresser & McKee, Inc., (CDM) for the Mayor's Office of Budget and Management. Participating departments include: Department of Water (DOW), Department of Sewers (DOS), Department of Transportation (CDOT), Department of Streets and Sanitation (DSS), the Department of Planning and Development (DPD), and the Department of Management Information Services (MIS). Several subconsultants are also participating in the project.

Geographic Coverage

The geographic area covered by the CWIMS project (the limits of the city of Chicago) is completely contained within the Illinois-East zone (number 1201) in the U.S. State Plane Coordinate System. The data base coordinates are based on the North American Datum of 1927 (horizontal datum). The data base units are "0.0305 meters" ("tenths of feet"); that is, a coordinate change of 100 data base units in either the X or Y coordinate represents a distance of 3.05 m (10 ft).

System Architecture

The CWIMS project is divided into the following contract deliverables: Project Design a series of data bases (Base Map, Physical

TABLE 1 Contrasting Perspectives of "Streets" by DPD and CDOT

Department of Planning and Development (DPD) Perspective	Chicago Department of Transportation (CDOT) Perspective
Base Map represents public rights-of-way (ROW) in the city	Inventory represents streets within the public rights-of-way
Includes ROW information for Expressways	Does not include data on Expressways
Does not include information for Ramps	Includes data on Ramps
View is "flat" (planimetric), i.e., ROW which cross also intersect	View is "non-planar", i.e., overpassing streets can cross over other streets without intersecting
Base Map is non-directional	Streets are directional (i.e., recognize a direction of travel)
A ROW "corridor" is viewed as a single entity	A ROW "corridor" can contain multiple instances of street features, such as the east-bound lanes and west-bound lanes of a median-separated boulevard
Instances where there is a short "jog" in the ROW where two adjacent sections of ROW do not exactly intersect are often explicitly represented in the DPD data (but not always)	Instances where there is a "jog" in the street centerline on two sides of a general area of intersection tend to not be explicitly shown in the CDOT data.

Inventory, Condition Assessment, Capital Projects, and Current Replacement Cost); a series of implemented systems (Capital Planning Decision Support, Systems Data Transference, Application Programs, Security, and Mainframe Diagnostics); and training.

The focus of this paper is in the area of the Physical Inventory Data Base deliverable. More specifically, this paper discusses the experience gained while using the official city base map provided by the DPD as the basis for associating spatial properties to the Physical Inventory data provided by the CDOT. The CDOT data was nongraphic with locations specified by street name and address range. The spatial properties that were assigned to the CDOT features are listed below:

- Illinois State Plane X,Y coordinates for the location of CDOT features,
- Network connectivity (which also defined the logical relationship between features), and
- Graphic representation.

The hardware and software environment for the CWIMS project reflects the city's interest in maintaining a central data base as the repository for all infrastructure data while taking advantage of the benefits offered by workstations and personal computers for system users. Data access is provided by use of a city-wide Transmission Control Protocol/Internet Protocol (TCP/IP) network with departmental servers. The central data base resides on the city's IBM mainframe computer in DataBase 2 (DB2) relational data bases. The GIS portion of the infrastructure data base is managed by the IBM product geoManager, part of the IBM Geographic Facilities Information System (GFIS) product set. Many other DB2 relational tables of infrastructure data exist external to geoManager. These tables can be accessed either directly or through geoManager data retrievals using relational "JOIN" operations.

The project plan calls for the delivery of analysis, planning, and decision support capabilities to CWIMS users via the ArcView2 desktop GIS product from Environmental Systems Research Institute, Inc. The data will be extracted from the city-wide corporate data base. A data translator will convert the GFIS data model stored in the geoManager data base into ArcView2 format during data extraction.

CDOT DATA MODEL

The CWIMS project is building its data model consistent with the GFIS architecture. The GFIS data model is a two-dimensional hierarchical (network) model that includes individual feature instances grouped into user-defined layers. Within each layer, each feature's network connectivity is explicitly modelled by permitting one or more nodes to exist at any X,Y location. This approach implements the nonplanar aspect of the data model, permitting multiple features to exist at the same X,Y location without the need for them to be network connected. Graphic representations are also explicitly defined for each feature, with the flexibility of defining any number of scale-dependent pictures that can be displayed at any location required by the application, even if that location is different from the feature's data base coordinate location. The GFIS data model also provides for up to two levels of "child" (dependent) features that inherit their location and their existence from their "parent" features (*9*).

The CWIMS CDOT Physical Inventory data model includes the features outlined in Table 2. A view of the CDOT features converted to date in the Chicago central business district are shown in Figure 1.

METHOD FOR ASSIGNING SPATIAL PROPERTIES TO CDOT DATA

In addition to the tabular address-delimited data provided by CDOT, the spatial properties of Illinois State Plane X,Y coordinates, network connectivity, and graphic representation must be added before the CDOT Physical Inventory data can successfully be incorporated into the CWIMS data base. This section outlines the process used to assign this spatial data.

DPD Midlines (Right-Of-Way Limits)

For the purpose of this conversion, the portion of the DPD Base Map data that provided the basis for assigning spatial properties to the CDOT data is the "midline" of the public rights-of-way (ROW) limits. These ROW midlines are modelled in the CWIMS data base as span features. The midline attributes of interest for matching the incoming CDOT data are listed below.

- Common street name,
- Unique midline identifier, and
- Address ranges: low/high even addresses, low/high odd addresses.

By virtue of its being in the geoManager data base, each instance of a midline feature also had Illinois State Plane X,Y coordinates for its endpoints, network connectivity at each end (expressed as nodes), and its graphic representation consisting of a polyline between the two endpoints. The spatial data associated with the Midline features are depicted in Figure 2*a*.

Corresponding CDOT Street Centerline Data

Although the CDOT Physical Inventory data consist of more feature types than street centerlines, all other CDOT data are located relative to the converted street centerlines. Therefore, the primary focus for converting the CDOT data into CWIMS was to assign spatial properties to the CDOT street centerlines by matching to the DPD midline features representing the official City Base Map already in CWIMS. Similar to the above discussion about the relevant midline attributes, the street centerline attributes that were used in matching to the ROW midlines follow:

- Street name,
- Street ID number (unique for CDOT, different from midline ID), and
- Address range (low/high addresses, *not* separated by even and odd).

Conversion Process for the CDOT Features

At first glance, it appeared that the CDOT street centerline data contained sufficient information to successfully identify the corre-

TABLE 2 Features in the CWIMS CDOT Data Model

Identifier	Type	Description
CSTREET	Span feature	CDOT Street Centerline: typically represents a portion of a street between two intersections (with an accompanying address range). In some cases, a Street Centerline feature may span a distance less than a complete block or cover a distance of multiple blocks depending on the locations of intersecting Streets. Network connectivity: All Streets at an intersection are connected to the same network node. This approach will be modified somewhat to assign different nodes for different "levels" when the conversion effort expands to include multi-level (multi-tiered) streets.
BRIDGE	Point feature	Bridge: elevated structure which "carries" a CDOT Street segment across a body of water, other street segments, railroads, etc. Network connectivity: Attached to the closest network node of the Street Centerline feature which it "carries".
VIADUCT	Point feature	Viaduct: elevated structure which "carries" a railroad, pedestrian walkway, or anything other than a CDOT Street over a CDOT Street segment. Network connectivity: Attached to the closest network node of the Street Centerline feature on which it is located.
CINTER	Point feature	Intersection: point at which two or more streets attach or converge. Network connectivity: Attached to the common network node shared by all Street Centerline features at the Intersection location.
SIGNAL	Dependent feature	Traffic Signal: Child feature of an Intersection which stores data about the number, status, and operation of the signals at the Intersection. Network connectivity: None, not applicable for dependent features.
UNDRPASS	Dependent feature	Vertical Clearance/Underpass: Child feature of a Street Centerline which stores data about the type and clearance between the pavement surface and the restricting structure above the street (such as a bridge or viaduct). Network connectivity: None, not applicable for dependent features.

sponding DPD midline features. Both data sources contained street name and address range on specific streets, so matching that information would provide access to the spatial data needed for the CDOT street centerlines.

It quickly became apparent that the mapping of Street Centerlines to ROW midlines was not one-to-one. Almost 1,600 more DPD ROW midlines existed than CDOT street centerlines. Not surprisingly, there were also several cases in which the spelling of the street names differed between the DPD data and the CDOT data. The spelling inconsistencies were fixed before continuing with the remainder of the conversion process.

The discussion in this paper focuses on converting the street centerline features. Once the centerlines were converted, assigning spatial properties to the remainder of the CDOT features was primarily an exercise in geocoding the point feature's street address.

Converting Street Centerline Features

Three different approaches were evaluated for the street centerline conversion.

Approach 1: Identify Matching Midlines Based on From/To Intersecting Streets

This approach attempted to use the from/to intersecting street information in the CDOT data to identify the midline feature that represented "one end" of the CDOT street centerline and likewise for the "other end." This approach was not selected because it only matched about 65 percent of the streets, took several hours of mainframe CPU time to complete, and could not handle the many-to-one case in which multiple midline features exist along the extent of a CDOT street centerline. This latter limitation could result in assignment of incorrect X,Y endpoint coordinates and a "gap" in the graphic representation.

Approach 2: Identify Matching Midlines Based on Low/High Addresses

The DPD midline data included low/high address ranges for both the even and odd addresses. In contrast, CDOT data included only low/high address ranges without distinguishing between even and

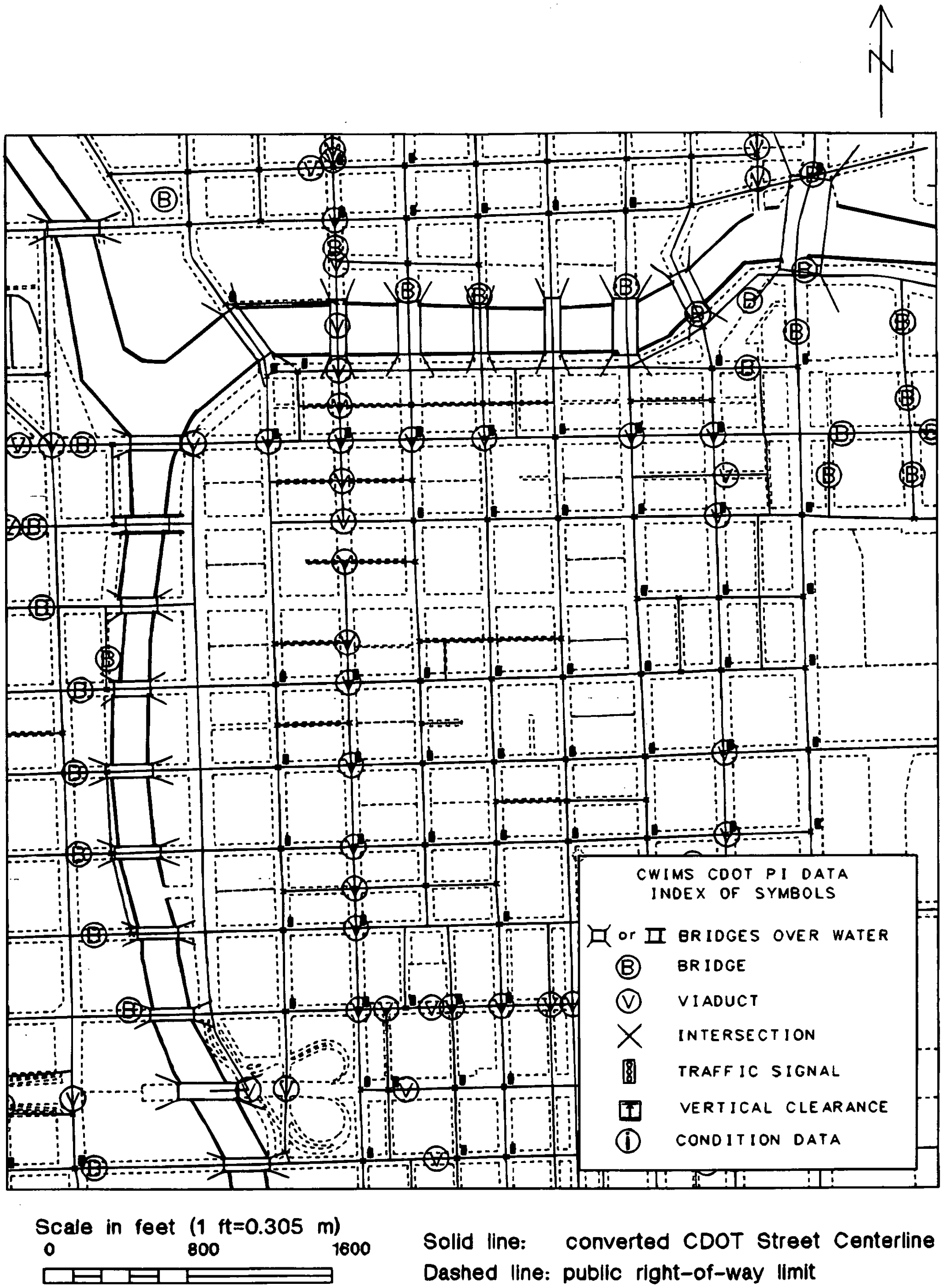

FIGURE 1 Converted CDOT highway features in the Chicago central business district.

odd addresses. Approach 2 relied on pairing streets with midlines based on similar low and high address ranges. The concept of "similar" was implemented in multiple passes through the data, with the second pass relaxing the criteria somewhat in an attempt to improve the match percentage. The match criteria for the two passes follows:

- Pass 1: CDOT street matches a DPD midline with the same name and the street has a low address that matches either the low odd or even midline address and a high address that matches either the high odd or even midline address.
- Pass 2 (for any street centerlines not matched on Pass 1): CDOT street matches a DPD midline with the same name, and all four midline addresses (low/high even and low/high odd) fall into an address range calculated as 5 below the CDOT street low address and 5 above the CDOT street high address.

This approach was more successful, but was not adopted. Pass 1 resulted in a 35 percent match rate, and Pass 2 increased the match percentage to 74 percent. This approach proved to be susceptible to anomalies in the DPD midline address range data. These anomalies

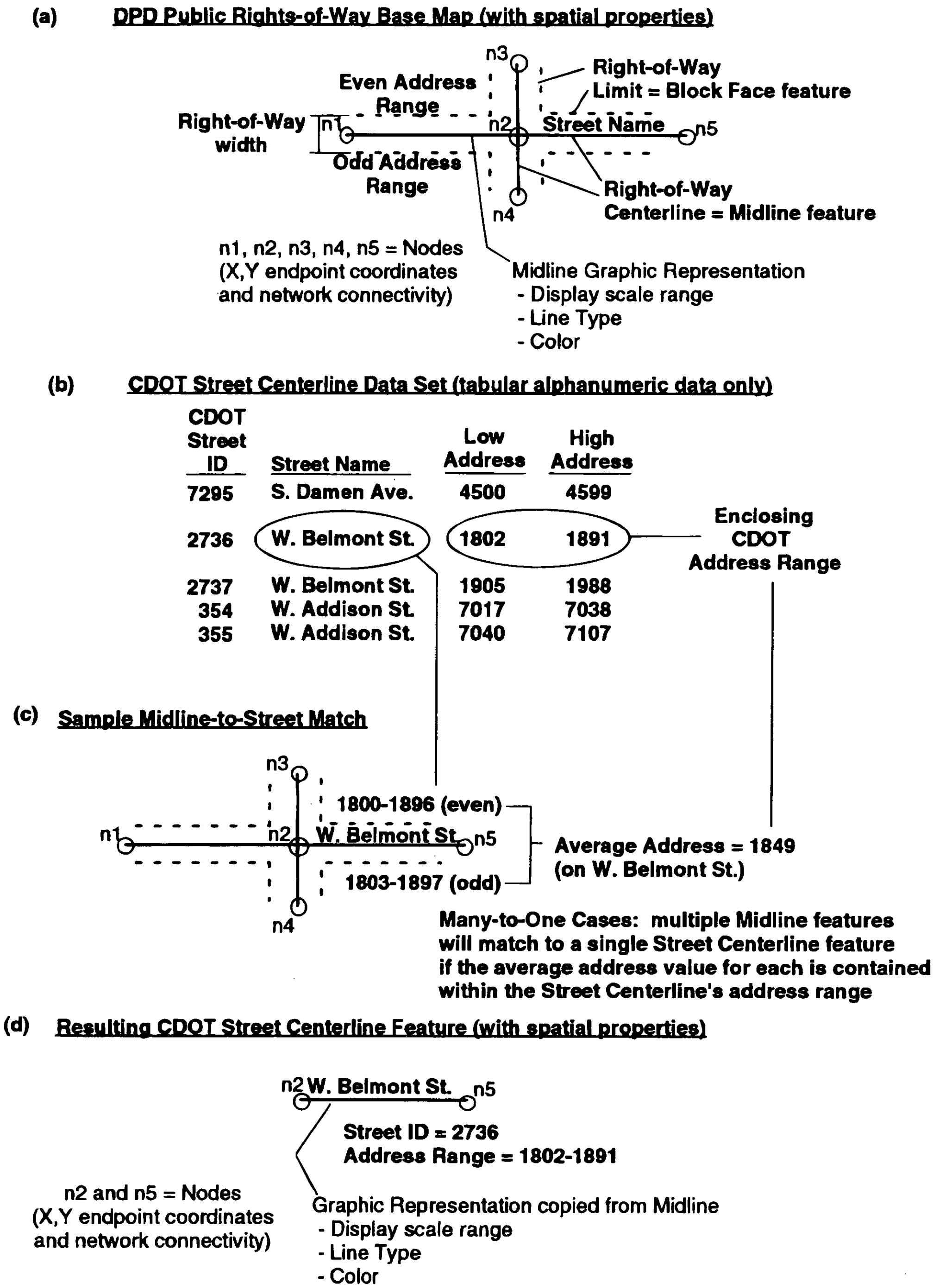

FIGURE 2 Use of average midline address range to match street centerline.

included cases in which the low even and odd addresses were more than 10 addresses apart. For example, the low even address may be 400, but the low odd address might be 425. Similarly, cases existed in which the high even address was different from the high odd address by more than 10 addresses. A few cases were also encountered in which the low even/odd address was larger than the high even/odd address.

Approach 3: Identify Matching Midlines Based on Average Address

This approach first calculated the "average address" for each DPD midline feature across the four values of low/high even and low/high odd addresses. This midline feature was then matched to the first CDOT street centerline that had the same street name and whose address range included this average midline address. Approach 3 proved to be the most flexible and resulted in a match percentage of 92 percent. One major advantage is that it permitted more than one midline feature to be matched to a single CDOT street centerline, which was the desired action in those cases in which the extent of a street centerline included multiple midline features. Approach 3 was selected for the actual CDOT data loading into the CWIMS Physical Inventory data base. An overview of Approach 3 is shown in Figure 2. Figure 2*a* depicts the existing spatial characteristics of the DPD Base Map data, including the midline features. Figure 2*b* shows that the incoming CDOT street centerline data were tabular alphanumeric data. Figure 2*c* shows the use of the average midline address (along with a common street name) as the method for matching to the CDOT street centerline data. Note that in many-to-one cases in which multiple midline features correspond to a single street centerline, multiple midlines would be matched to a single street if their average address value fell into the range of a single street centerline feature. Figure 2*d* depicts the resulting street centerline feature with its newly assigned spatial characteristics of X,Y coordinates for its endpoints, nodes for its network connectivity, and its graphic representation (copied from its matched midline feature or features).

Large areas of the Chicago street network are characterized by a well defined regular grid. In those areas, almost 100 percent of the CDOT street centerline features were successfully matched to the DPD ROW midline features. The adopted technique (Approach 3) did prove to give some "false-positive matches" in densely developed areas of the city where traffic control medians or triangular-shaped dividers exist.

Street Centerline Spatial Data

The spatial data assigned to a matched street centerline is summarized in Figure 2*d* and were determined as follows:

- The street centerline feature was added to the "C" layer (CDOT).
- The X,Y coordinates for the endpoints were assigned to be the Illinois State Plane X,Y coordinates for the corresponding endpoints of the matched midline feature or features.
- All street centerline features were assigned the same "node value" at their respective X,Y endpoints (on the C layer). This guaranteed a connected street network because all street features with an endpoint at any common X,Y coordinate will occupy the same network node at that location. The node assignment process will be modified to some extent in the future when it becomes possible to process information for multi-tiered (multiple level) streets. In this case, multiple nodes will be required at the same X,Y coordinate to ensure proper network connectivity across the different tiers.
- The graphic representation of the street centerline was obtained by copying the existing graphic representation of the midline feature or features that had been combined to represent the street.

Converting the remainder of the CDOT features was dependent on the successful conversion of the street centerline feature or features with which the bridge, viaduct, intersection, or vertical clearance underpass was associated. Once the corresponding street centerline(s) was converted, assigning spatial data to the other types of CDOT features was accomplished by converting the point feature's street address into an Illinois State Plane X,Y coordinate, assigning it to a node at the near end of the street centerline feature, and assigning a designated type of symbol for its graphic representation. The focus of this paper is on the process and issues associated with converting the street centerline data, so no additional discussion will be provided for the other CDOT features.

ISSUES AND OBSERVATIONS IN STREET CENTERLINE CONVERSION

It would be naive to expect that the algorithmic approach outlined in the previous section would result in a perfect match between the two independent data sets. In fact, several issues were identified as the result of problems encountered during the matching process.

Encountering problems during the conversion process typically meant that one or more CDOT street centerline features could not be converted into the CWIMS Physical Inventory data base. To focus attention on these cases and recommend corrective action, listings of all the "no match" features were compiled into a *Source Data Errors Report*, and that report was provided to CDOT. Recommendations have also been made to the DPD for enhancements to the DPD Base Map representation of midline features to better handle the cases of divided streets at the same level and streets that exist at multiple levels (i.e., tiered streets). The work to resolve and correct these identified issues is an ongoing process within the CWIMS project.

Note that there is also a domino effect whenever matches are not successful for any features that have dependents or relationships to other features. For example, an underpass/vertical clearance is modelled as a child feature to a street centerline. Any time a street centerline cannot be matched, then it is not possible to convert any underpass/vertical clearance features associated with that centerline. Similarly, intersection point features are located by identifying the point where two or more street centerline features come together. Any time street centerline features cannot be converted, then it also will not be possible to convert one or more intersection features. Traffic signals are modelled as child features of an intersection. As a result, any intersection features that are not converted will typically prevent the conversion of one or more traffic signal features.

The sources of inconsistencies when assigning spatial properties to the street centerline features in the CDOT data files follow. To generalize the information, the two different data sets are referred to as the Reference Source (for the DPD data set containing spatial properties) and the Input Source (for the CDOT data set to which spatial properties must be assigned):

1. Street centerline features existed in the Input Source whose street name could not be located in the Reference Source.
2. Street centerline features existed in the Input Source whose street name was located in the Reference Source but for which there were no associated instances of midline features. This is the situation in which the Reference Source has provision for a named ROW (midline feature) in its common street name table. However, there were no instances of features for that ROW midline. As a result, it was not possible to obtain the spatial information needed to convert the corresponding street centerline.
3. A single ROW midline feature existed in the Reference Source and contained within its address range the entire address range of two or more street centerline features in the Input Source, as depicted in Figure 3*a*. Per the project design, it was intended to handle any given midline as a complete feature. There was no intent

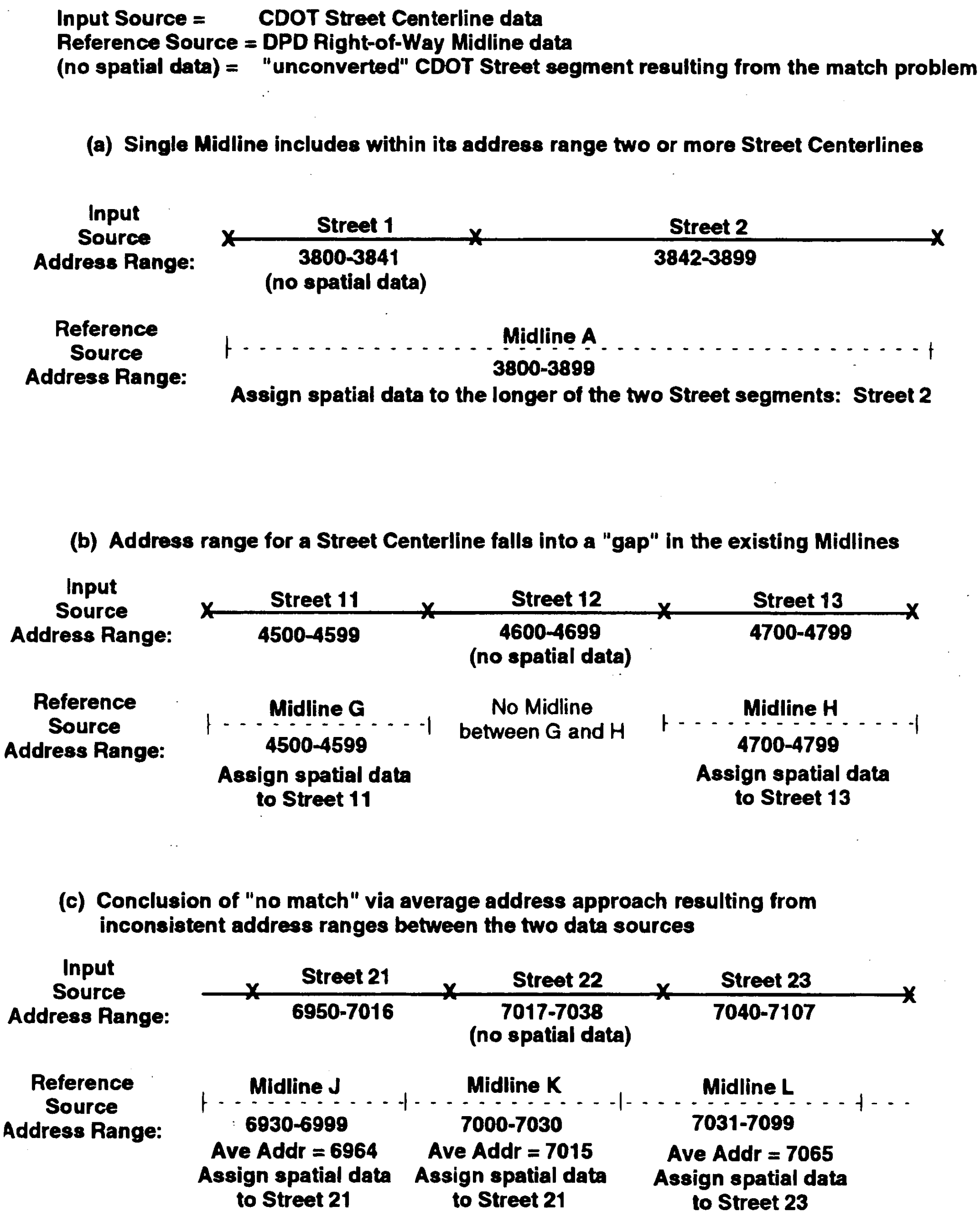

FIGURE 3 Overview of problems encountered with the "average midline address" approach.

to subdivide any midline features when matching to the street centerlines because it is the city's intention to make revisions in the raw data to resolve many of these differences. Accordingly, for this phase of the data conversion, the spatial properties for the entire span of a midline were assigned generally to the longer of the multiple street centerline features. This approach left as unconverted the other (shorter) street centerlines that were also completely contained within the midline address range.

4. Even though the street centerline from the Input Source found a match on a common street name with a ROW midline in the Reference Source, the address range of the street centerline was inconsistent with the address range of the midline in one of the following ways:

a. Address range too high: street centerline had an address range that was greater than the address range of any midlines with the same common street name.

b. Address range too low: street centerline had an address range that was less than the address range of any midlines with the name common street name.

c. Address range fell into a gap: although the street centerline features may have spanned a continuous set of address ranges, the corresponding ROW midline features had gaps such that there was no midline feature (and consequently no spatial data) corresponding to one or more sets of address ranges along the street centerline. An example of this case is provided in Figure 3*b*.

5. Additional anomalies in address ranges between the street centerlines in the Input Source and ROW midlines in the Reference Source follow:

a. Street centerline segments with very short address ranges seldom resulted in a successful match to a midline. The short address ranges typically resulted when a diagonal street intersected another street and created a very short block. Usually, the conversion process was unable to find a midline whose *average address* was within the address range of the short street centerline segment. This is similar to the case shown in Figure 3*a*, in which *Street 1* represents the short street that was not converted.

b. Address ranges on "even" and "odd" sides of ROW midline features were sometimes very different. When considering the entire address range (across both the even and odd sides) of a midline, it was sometimes the case that the address range for a street centerline segment may be completely contained within the address range for two different adjoining midlines.

c. The approach of using the "average midline address" sometimes resulted in situations in which no match was concluded, even though the high or low address of the midline was somewhat contained in a street centerline feature. This often resulted from inconsistent address ranges in the two different data sets, as shown in Figure 3*c*.

CONCLUSIONS

The approach developed to this point in the project for assigning spatial properties to the CDOT street centerline features from the independent DPD midline features has resulted in the successful conversion of a large percentage of the CDOT data into the CWIMS Physical Inventory data base. Although there is still room for improvement, recall that the main objective for CWIMS is to support the capital planning process. The CDOT street centerline spatial data already converted represent major progress because they support a level of comparative analysis for capital projects planned by different departments in the city in a manner never before possible.

Work continues on the CWIMS project. With the benefit of the initial analysis for using the DPD midline and CDOT street centerline data sets together, the city is evaluating the types of changes that are appropriate in both sets of data to improve data consistency. The following conclusions can be cited as the result of experiences to date:

1. The expected "consistency" between the existing data set that requires assignment of spatial properties and the candidate data sets that can be used as source data should be assessed at the start of a conversion project. Choose the candidate source data set that is expected to have the smallest amount of inconsistency. The more years for which the two data sets were maintained independently, the more likely it is that inconsistencies will exist.

2. It is likely that this process of using one data set to assign spatial properties to another data set will be the first time that any type of "independent analytical work" has been performed on either of the data sets. The process may uncover problems in the raw underlying data. A focused effort is required to ensure that this process is used as a learning experience with the goal of making improvements in the overall quality of the data. Expect an iterative process, with revisions to the conversion process between each iteration.

3. Once inconsistencies are identified, it is important that the reason for the inconsistency be determined and that processing procedures are adopted to avoid the same problems in the future.

4. Jurisdictional considerations may affect the content of one or both of the data sources. This is more noticeable with municipal or local governments in which the agency often has no control over, or responsibility for, federal or state transportation infrastructure features. The federal and state-owned features tend to be the larger features on the transportation network, so their omission can create some conspicuous gaps in the geographic data base at the municipal or local level. It is appropriate to evaluate whether transportation features for which an organization has no jurisdiction or responsibility should at least be accounted for in the municipal or local geographic data, perhaps in the context of read-only data obtained on a periodic basis from a federal or state transportation agency.

5. It can be expected that some amount of data typically will not be successfully converted when using automated techniques. However, the features that are converted still represent significant amounts of usable, valuable data. Further, this successfully converted data provide a framework for future work to resolve discrepancies and to improve the conversion success rate.

6. When using address ranges as a basis for matching between sets of span features, identify if there are any local conventions that can simplify the process of determining which end of a street (or ROW) segment is the "low end." For example, the city of Chicago uses a well defined quadrant system for assigning street prefixes (North, South, East, and West). These quadrants are based on an origin at a designated intersection of two major streets in the downtown area. By convention, all address ranges are assigned such that the "low" address is always the end of a street (or ROW) segment closest to the origin street intersection.

ACKNOWLEDGMENTS

The project described in this paper is being performed by Camp Dresser & McKee, Inc., (CDM) for the Office of Budget and Management, city of Chicago. Carl Johnson and Daniel Hudson of CDM provided many valuable review comments during the writing of this paper, and their insights are appreciated.

REFERENCES

1. Nyerges, T. L., and K. J. Dueker. *Geographic Information Systems in Transportation.* The Center for Urban Studies. Portland State University, Portland, Ore., 1988.
2. *NCHRP Research Results Digest 180: Implementation of Geographic Information Systems in Transportation Agencies.* TRB, National Research Council, Washington, D.C., 1991.
3. *NCHRP Report 359: Adaptation of Geographic Information Systems for Transportation.* TRB, National Research Council, Washington, D.C, 1993.
4. Nyerges, T. L. Locational Referencing and Highway Segmentation in a Geographic Information System. *ITE Journal,* Vol. 60, No. 3, March 1990, pp. 27–31.
5. Guthrie, M. F. The Digital Control Section Atlas at the Michigan Department of Transportation: Implementing Segmentation and Refining Geometry. In *Proc., Geographic Information Systems for Transportation Symposium,* Orlando, Fla., 1991, pp. 163–176.

6. O' Neill, W. A., and B. Akundi. Automated Conversion of Milepoint Data to Intersection/Link Network Structure: An Application of GIS in Transportation. In *Transportation Research Record 1261,* TRB, Washington, D.C., 1990, pp. 27–34.
7. Azad, B., J. Brown, and J. N. Brown. Matching Census and Local Geographic Boundaries: A Practical, Repeatable Approach. *Geographic Information Systems,* Oct. 1991, pp. 40–45.
8. Petzold, R., S. Lewis, and D. Fletcher. *GIS as a Tool for ISTEA Management.* Fourth Annual TRB Workshop on Application of Geographic Information Systems to Transportation, Washington, D.C., Jan. 1994, p. 74.
9. IBM Corporation. *Graphics Program Generator Application Developer's Guide.* Publication Number SH20-6891-02. Mechanicsburg, Pa. Jan. 1990, pp. 25–31.

The views expressed in this paper are solely those of the author and not of the city of Chicago.

Publication of this paper sponsored by Committee on Transportation Data and Information Systems.