THE CONGRUENCE BETWEEN U.S. CENSUS DIME TECHNOLOGY AND OBSERVED LOCAL OPERATING AGENCY GEOGRAPHIC INTERESTS AND NEEDS

Charles E. Barb, Jr.
University of Oklahoma

INTRODUCTION

The purpose of this paper is to briefly review the basis of U.S. Census DIME technology and to compare its applicability to geographic information system needs of local public agencies. The paper is drafted in semi-technical terms that should be understood by geographers and cartographers without extensive background in DIME technology.

EVOLUTION OF DIME TECHNOLOGY

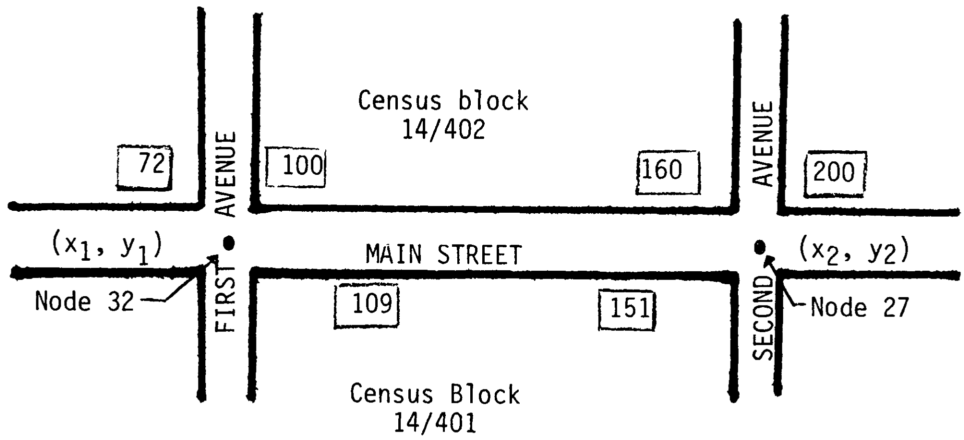
The current basis of DIME technology is best understood in a historic context and through comparison with the earlier SACS system. Both DIME and SACS are based upon automated geographic reference directories, or "geographic base files," at the street segment scale of detail. Figure 1 summarizes the record content of the two files. From a functional perspective, both types of segment records include an address element and a geocode element. The process of street address geocoding involves searching the address element of the records in a file to find the segment along which the address falls, and extracting selected geocodes from the segment record found. This search-extraction process is computer-resident.

SACS

SACS stands for Street Address Conversion System, which originated at the University of Washington about 1963. 1/ The system was developed in recognition of a need to computer geocode street addresses, particularly the large number of addresses being collected in regional transportation origin-destination surveys.

SACS was based upon a file of street segments which were described by segment end "node" coordinates. (The file of street segments was limited to those along which street addresses would logically occur.) The focus of SACS was to translate street addresses to point coordinates which would then enable point plotting and,

Editor's Note: Dr. Barb's background in the technology was derived through eight years of systems and applications research at the Urban Data Center, University of Washington.



SACS Street Segment Record	Illustration Content	DIME Segment Record
Address element		
Street name	MAIN	Segment name
Street type	STREET	Segment type
Low left address	100	Low left address
High left address	198	High left address
Low right address	101	Low right address
High right address	199	High right address
Geocode element		
- (not included) -	14/402	Left census tract/block
- (not included) -	14/401	Right census tract/block
Low node number	32	Low node number
High node number	27	High node number
Low node x,y	x_1, y_1	Low node x,y
High node x,y	x_2, y_2	High node x,y

Figure 1. Conceptual Content of SACS and DIME record.

through point-in-polygon processing, 2/ subsequent translation to area codes. The elegance of the system lay in its efficient segment file structure, the interpolation of coordinates for individual addresses from the segment record, 3/ and the emphasis upon coordinate rather than area code manipulation of data. At the time the system also maximized the potential of limited computer power available.

Subsequent University of Washington research in street segment file construction methods 4/ emphasized cartographic accuracy both in the selection of underlying system map resources and in policy governing the abstracting of street networks to line segments. The coordinate precision sought for the system was ± 10 feet which matched the accuracy of commonly available municipal map resources of the era and which also allowed for a unique coordinate to be interpolated for individual addresses.

By 1968 SACS had been conceptually promoted with the Dominion Bureau of Statistics and National Capitol Commission in Canada; Portland, Oregon and the U.S. Bureau of the Census. In that year a prototype city-wide Seattle system became operational and was broadly advertised for local use. Significantly, public agency response was negligable which was a first indication that there was a problem with broad local adaptation of the technology.

It is also useful to parenthetically note that not until 1969 was network minimum path application of the SACS street segment file recognized.5/ This application necessitated an additional set of criteria for street segment record definition and directory completeness, one of several redefinitions of the file's purpose and content.

U.S. Census Activities

The principal geographic problem recognized in the 1960 decennial census was the lack of a comprehensive and uniform metropolitan map base with which to conduct the field enumeration and later to geographically describe small area statistics.6/ The solution to the Bureau's map problem was the Metropolitan Map Series. While derived from a U.S. Geological Survey map base, the Metropolitan Maps were only coarsely drawn to meet the Bureau's limited representational purposes defined at that time. The Bureau's program in automated, segment-scale geocoding evolved later.

Early Bureau experimentation in automated geocoding, at the census tract-scale, occurred during the 1963 Economic Census.7/ The experiment employed a street side approach which represented a direct computer translation of the manual census tract street index. In 1965 the Bureau embarked upon the Address Coding Guide (ACG) program which employed the same segment-side approach but at the individual street segment level.8/ The limited focus of the ACG was to translate street addresses to nominal census tract and block area codes. Point coordinate geocoding and street network applications were not part of the ACG concept.

In 1966 the New Haven (Connecticut) Census Use Study was initiated, in part, to explore the potential of computer mapping local data geocoded through an ACG. As a result of this research the street segment-based DIME file was proposed as an alternative to the ACG.9/ The DIME approach had two advantages: (1) it eliminated the redundancy of the ACG (DIME essentially merged the two segment-side records into one segment record), and (2) it enabled, through graph editing, a method of examining the completeness and coding accuracy of a file.10/ This latter attribute was particularly important to the Bureau because it was becoming apparent that local efforts at coding ACG's were turning out badly. The DIME approach was accepted in 1968 and at the same time the Metropolitan Map Series was designated as the map base for the files. Local agency coding of the DIME files occurred in 1969 and 1970 and, again, due to inadequate funding and supervision, a poor quality job was done.

It is significant to note that the DIME file, within the Bureau's frame of interest, constitutes a census area boundary segment file rather than a street segment file. This difference becomes significant in street network-related minimum path applications. (The extraction of a usable network file from a Census DIME file is not a trivial undertaking.) It is also important to understand that the basic DIME file -- without node coordinates -- fulfills the Bureau's basic need for file editing and interest in translating street address to nominal census area codes. The Bureau subsequently inserted node coordinates into their DIME files as a low priority project by digitizing nodes on Metropolitan Maps. The Bureau's interest in the coordinate precision of a DIME file remains limited to coarse representational purposes.

In summary, SACS was conceived to translate street addresses to point coordinates to enable point mapping and, through point-in-polygon processing, area coding. The system was initially conceived as a tool for micro urban socio-economic analysis of street address-identified data resources. DIME technology, also segment-based but representing a description of census area boundaries, was developed by the Census Bureau to support internal needs and also to facilitate small area chorologic analysis. Coordinate precision and network application of DIME files is secondary. Minimal Federal funding and project supervision of DIME file coding by local agencies has resulted in a marginal file which requires considerable additional local investment before it can be used. The Bureau's current thrust is to clean up the files for application in the 1980 decennial census.

OBSERVED LOCAL AGENCY INTERESTS AND NEEDS

Experience in Seattle mentioned above, and research into long-term geocoding system developments elsewhere, 11/ suggests local adaptation of DIME-type technology is proceeding slowly. A significant question is -- why? It is proposed that, in part, the reason is that we have been distracted with the notion that DIME technology is the "answer," and the Federal Government has been preoccupied with selling a packaged system for self-serving reasons. This course has been taken in place of analyzing existing and evolving local agency interests and needs in geographic description and analysis, and developing necessary technology to meet them.

Local agency operational interests and needs and their congruence with DIME technology is suggested by a brief examination of a few familiar municipal functions. Time and space limit discussion to merely a suggestion of the type of examination that should be undertaken.

ASSESSMENT

The assessment function involves land and its development. The unit of interest is a legally and precisely defined land parcel. Of course in many cases a parcel does not have an assigned address and in most cases today is geographically referenced in legal rather than street address terms. Obviously inherent to the interests and needs of this function is a parcel data bank including extensive attribute descriptors of property and particularly its legally described location and extent.

BUILDING INSPECTION

The building inspection function involves structures under development and, possibly, ongoing monitoring of their condition and use. The unit of interest is the structure and sometimes dwelling and commercial units within a structure. The inherent interest and need of this function is a mixed structure/unit data bank also including extensive attribute descriptors.

TRAFFIC ENGINEERING

The traffic engineering function includes street and "street furniture" inventory and maintenance. Street furniture broadly includes on-street parking spaces and meters, traffic signs, and traffic channelization. These are generally described in an engineering precision. Inherent to the interests and needs of this function is a mixed resource data bank including engineering quality geographic descriptors.

DISTRIBUTION UTILITIES

Distribution utilities include telephone and electrical and gas power utilities. The utility function requires description of distribution networks which are frequently not congruent with streets and census statistical area boundaries. For sub-surface networks in particular, in-place facilities must be described with engineering precision. 12/ Inherent to the interests and needs of this function is a network-based, engineering-quality facility data bank including extensive attribute descriptors and probably including customer information.

In summary, local operating agencies have concern and responsibility for a range of geographically located land and facilities which they need to have described to various degrees of geographic precision. In most cases their needs for geographic precision are far greater than the ± 40 foot accuracy claimed for the Metropolitan Map Series or obtainable from a DIME file. Local agencies are also operationally and administratively concerned with a broad range of attributes of the land and facilities beyond their geographic location. These interests and needs are more closely met by special purpose, facility-oriented data banks than geocoding systems conceived principally for socio-economic analysis.

CONCLUSION

The brief description of the origin of census DIME technology reveals that it is a tool for relatively coarse, urban-scale socio-economic analysis, particularly as related to the decennial census. A brief analysis of local agency functions suggests that local agency operational and administrative interests and needs are more geographically precise than census DIME files can support and more substantively based than geographic location. Two principal conclusions are drawn from the discussion.

First, given the current basis of U.S. Census DIME technology, local public agencies should not look to it to meet their broad and complex operational geographic coding and cartographic interests and needs, directly.

Secondly, to meet local agency needs and interests in most cases, a relatively broad local technology infrastructure will have to be developed. The Census Bureau's DIME system will probably be recognized as a small element in this infrastructure. At this time of initial infrastructural development, greater emphasis should be placed upon the transfer of concepts such as DIME graph-theoretic coding rather than the transfer of packaged but often inapplicable systems.

REFERENCES

1. SACS research was initiated and directed by Dr. Edgar M. Harwood, Professor of Civil Engineering and Urban Planning. See: Robert B. Dial, Street Address Conversion Systems, Urban Data Center Research Report No. 1 (Seattle: University of Washington, 1964). The University of Washington's decade-long research program in automated geocoding technology has been principally funded by the National Science Foundation.

2. By point-in-polygon analysis is meant a computer-resident process of determining in which polygon, described in terms of perimeter vertex coordinates, a point geocode falls. The approach enables flexible tabulation of data by statistical area. See: Helaman R.P. Ferguson, <u>Point-in Polygon Algorithms</u>, Urban Data Center Research Report No. 7 (Seattle: University of Washington, 1971).

3. Point geocodes obtained from SACS commonly included "block face" coordinates representing the midpoint between segment-end node points and including on appropriate setback distances from the segment, and "unique address" coordinates which represent a proportional interval along (but set back from) the segment based upon the house number. Block face coordinates, being assigned consistently to all addresses falling along a segment side, represent a sub-block aggregating key.

4. Hugh W. Calkins, <u>Operations Manual for Street Address Conversion System</u>, Urban Data Center Research Report No. 2 (Seattle: University of Washington, 1975); and Richard Arthur O'Neal, "The Geographic Base File: Its Specification and Development," an unpublished master of urban planning thesis, Department of Urban Planning, University of Washington.

5. Matthias H. Rapp, <u>The Minimum Path Assisgment System and Its Application to the 1969 Seattle Shelter Allocation Project</u>, Urban Data Center Research Report No. 4 (Seattle: University of Washington, 1969).

6. William T. Fay and Robert L. Hagen, "Computer Based Geographic Coding for the 1970 Census," a paper presented at the American Institute of Planners Conference, Portland, Oregon, August 14, 1966.

7. Herman H. Fasteau and George Minton, <u>Automated Geographic Coding System</u>, 1963 Economic Censuses: Research Report No. 1 (Washington, D.C.: U.S. Bureau of the Census, 1965).

8. Fay and Hagen, <u>op</u>. <u>cit</u>.

9. Donald F. Cooke and William H. Maxfield, "The Development of a Geographic Base File and Its Uses for Mapping," <u>Urban and Regional Information Systems for Social Programs</u>, Papers from the Fifth Annual (1967)Conference of the Urban and Regional Information Systems Association, ed. by John E. Rickert (Kent State University, 1969).

10. The acronym DIME stands for Dual Independent Map Encoding which underscores the concept of dual or redundant coding of census area boundaries as a graph. Each segment is encoded as a link in a network and again as an edge between adjacent census areas. This coding enables the editing of a Dime file for closure around a census area or around a network intersection.

11. Charles E. Barb, Jr., <u>Automated Street Address Geocoding Systems: Their Local Adaptation and Institutionalization</u>, Urban Data Center Research Report No. 9 (Seattle: University of Washington, 1974).

12. Interestingly, some information system designers are turning to describing subsurface distribution networks in textual rather than cartographic form due to the limitations of recording information on and extracting information from maps.