6-hL

REPORT NO. DOT-TSC-OST-74-7

DEPARTMENT OF TRANSPORTATION
COUNTY DIME FILE

TECHNICAL SUMMARY

Chris L. Davis

MARCH 1974

FINAL REPORT

Prepared for
DEPARTMENT OF TRANSPORTATION
OFFICE OF THE SECRETARY
Office of the Assistant Secretary
For Policy, Plans and International Affairs
Washington DC 20590

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-TSC-OST-74-7		

4. Title and Subtitle	5. Report Date
DEPARTMENT OF TRANSPORTATION COUNTY DIME FILE TECHNICAL SUMMARY	March 1974
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Chris L. Davis	DOT-TSC-OST-74-7
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
The Charles Stark Draper Laboratory, Inc. 75 Cambridge Parkway Cambridge MA 02142*	11. Contract or Grant No. OP4091/R4805
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address Department of Transportation Office of the Secretary Office of the Assistant Secretary For Policy, Plans and Internat. Af. Washington DC 20590	Final Report July 1973 to June 1974
	14. Sponsoring Agency Code

15. Supplementary Notes *Under Contract to: Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge MA 02142

16. Abstract

This document describes the DOT County DIME (Dual Independent Map Encoding) File available in sequenced and compressed versions from the Information Division of the Transportation Systems Center. The file provides the capability to produce maps of county-base data with display software systems currently in wide use. The Sequenced County DIME File has been arranged so that each county has a complete boundary whose segments are internally structured in a clockwise manner. The Compressed County DIME File is composed of variable length records that contain all the segment data for a county boundary common to two and only two counties.

The technical specifications necessary to run the Sequenced File on 32 bit machines and the Compressed File on IBM 360/370 equipment are included.

17. Key Words	18. Distribution Statement
Geocoding, County, County Boundary, County Segment, Geographic Display	DOCUMENT IS AVAILABLE TO THE PUBLIC

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	20	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PREFACE

This document describes the DOT County DIME (Dual Independent Map Encoding) File and provides the technical specifications necessary to use one of the two versions available on the appropriate hardware. The file was produced under the Information Program at the Transportation Systems Center sponsored by the Office for Policy, Plans and International Affairs. Work on the development of the file was performed under contract by the Charles Stark Draper Laboratory Inc.

The DOT County DIME File and the DOT National Geocoding File are avialable by contacting the Information Division, Transportation Systems Center, Kendall Square, Cambridge, Massachusetts, 02142.

TABLE OF CONTENTS

1. INTRODUCTION

Several software systems have been developed for the display of data. Certain of these systems such as SYMAP, CALFORM, DPS, etc. are capable of displaying data values associated with a geographic area. This is accomplished by density shading within the geographic area based on the relative magnitude of the data values to be displayed.

These software systems implicitly require certain capabilities and/or associated data files. These are:

a) a name or code file to associate each geographic area with alphanumeric or numeric identifiers;

b) a machine-readable file which contains the boundary coordinates for the named or coded geographic areas;

c) a machine-readable file of the data values to be displayed. This data must have the same name or code as used in the geographic area definitions.

2. GEOCODED DATA

Many federal agencies and industrial organizations collect or associate data by geographic area. Although there is an increasing tendency toward standardization, in practice, much useful data can be found coded in a wide variety of geographic reference systems. Therefore, the association of data coded to different systems requires a conversion capability.

3. GEOGRAPHIC CODE CONVERSION CAPABILITY

Because of the conversion problem, a U.S. Department of Transportation National Geographic Converter File (File 1) has been developed. This file provides conversion capability at the county or county equivalent level. It encompasses those codes listed in Table 1 for each county in the U.S.

File 1, therefore, provides the capabiltiy to associate data under several geocoding systems.

TABLE 1. ELEMENTS OF THE NATIONAL GEOCODING CONVERTER FILE 1

REGION	STATE	AREA	COUNTY CLUSTER	COUNTY
An areal unit smaller than the United States and larger than a single state. The number of regions within a File 1 geocoding system ranges from 8 units to 20 units.	An areal unit which is the first order political subdivision of the United States. The number of states within a File 1 geocoding system ranges from 48 units to 51 units.	An areal unit not necessarily smaller than a state but larger than the average county cluster. The number of areas within a File 1 geocoding system ranges from 120 units to 200 units.	An areal unit smaller than a state, generally consisting of two or more counties. The number of county clusters within a File 1 geocoding system ranges from 430 units to 550 units.	An areal unit which is the second order political subdivision of the United States. The number of counties and county equivalents within a File 1 geocoding system is approximately 3,142 units.
9 Region Codes: • Standard Federal Region • Bureau of the Census Division • Standard Point Location Code Region • Bureau of Public Roads Region II • National Location Code Region • Office of Emergency Preparedness Region • Water Resource Region • Land Resource Region • Census Freight Rate Territory	12 State Codes: • Federal Information Processing Standard State • Interstate Commerce Commission State • International Business Machine State • Bureau of Public Roads State I • Dun and Bradstreet State • General Services Administration State • Office of Business Economics State (Now BEA) • Bureau of the Census State • Office of Emergency Preparedness State • National Location Code State • Bureau of Public Roads State II • Standard Point Location Code State	5 Area Codes: • Office of Business Economics Region • Zip Code Market Area • Census Economic Sub-Region • Water Resource Sub-Area • Land Resource Area	9 County Cluster Codes: • State Economic Area • Interstate Commerce Commission Area • Office of Emergency Preparedness Area • National Location Code Area • Bureau of Public Roads Area II • Standard Point Location Code County Cluster • DOT Transportation Zones • Standard Metropolitan Statistical Areas • Census Market Areas	11 County Codes: • Federal Information Processing Standard County • Bureau of Public Roads County I • International Business Machine County • Dun and Bradstreet County • General Service Administration County • Office of Business Economics County (Now BEA) • Interstate Commerce Commission County • Office of Emergency Preparedness County • National Location Code County • Bureau of Public Roads County II • Standard Point Location Code County

4. BOUNDARY COORDINATE FILE

While File 1 allows data associates at the county level and the systems listed above provide for machine display, a county boundary coordinate file is necessary to complete an ability to display geographic referenced data.

The Bureau of the Census has developed a machine-readable file containing the coordinates of the county and state boundaries for the entire continental U.S. This file is in the Dual Independent Map Encoding (DIME) format. Each record in this file represents a county boundary segment, and contains information as shown in Figure 1.

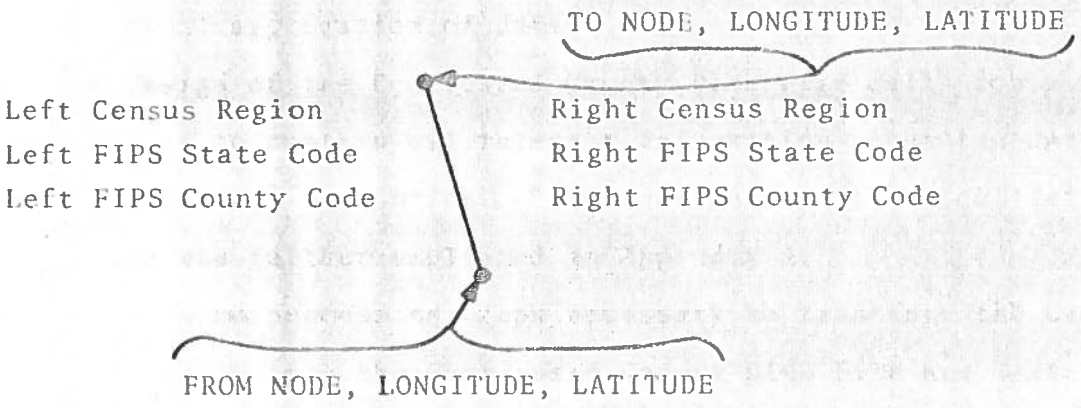


Figure 1. Census DIME File County Boundary Segment Information

5. BOUNDARY COORDINATE FILE FOR DISPLAY PURPOSE

The aforementioned software display systems generally require that the boundary coordinate file be in a sequenced format which is internally structured in a clockwise manner. Initial use of the census file demonstrated that an effective sequence for ordering the county boundary segments was not available in the County DIME File. Consequently, the County DIME File could not be used with display software.

A development effort was undertaken by the DOT in order to provide a county boundary coordinate file compatible with existing display software. The effort had as its objective, a compressed file which would facilitate:

a) information display

b) efficient information storage

c) efficient machine processing

d) areal aggregation of data.

The design of the Compressed County DIME File calls for a single record to contain all relevant information about the entire boundary between two counties. This idea and the technical details of the file are further explained in Appendix A.

The machine processing steps necessary to transform the Census County DIME File into the Compressed County DIME File are listed in Appendix B.

It is felt that the resultant file satisfies the objectives set forth. It is interesting to note that the organization and efficiency of the Compressed County DIME File become issues of

major significance as data display, information aggregation, and geocoding work proceeds below the county level.

APPENDIX A

DOT COMPRESSED COUNTY DIME FILE

A.1 DOT COMPRESSED COUNTY DIME FILE

Number of Records:	17,490
Tape:	9 Track, IBM 360
Density:	800 or 1600 BPI
Logical Record Length:	Variable, maximum 1272
Blocksize:	7294
Record Format:	Variable Blocked
Label:	None
Code:	Binary

A.2 FILE DESCRIPTION

The Compressed County DIME File is composed of records which contain the data for all boundary segments common to two adjacent counties. Figure A-1 and the following text describes the file and record structure.

Figure A-1 illustrates a hypothetical county (County A), its boundary segments labeled with small letters, the nodes for each segment numbered from 1 to 14 and adjacent counties. Since there are five counties which are adjacent to County A, there will be only five records in the Compressed County DIME File. The format and content of each of these records is shown in Table A-1.

As illustrated in Table A-1 each record can be thought to have two sections. The fixed section contains that information defining the right and left counties, the first node common to both counties and the last node common to both counties. A variable number of nodes defining intermediate boundary segments common to two counties is contained in the variable section of the record.

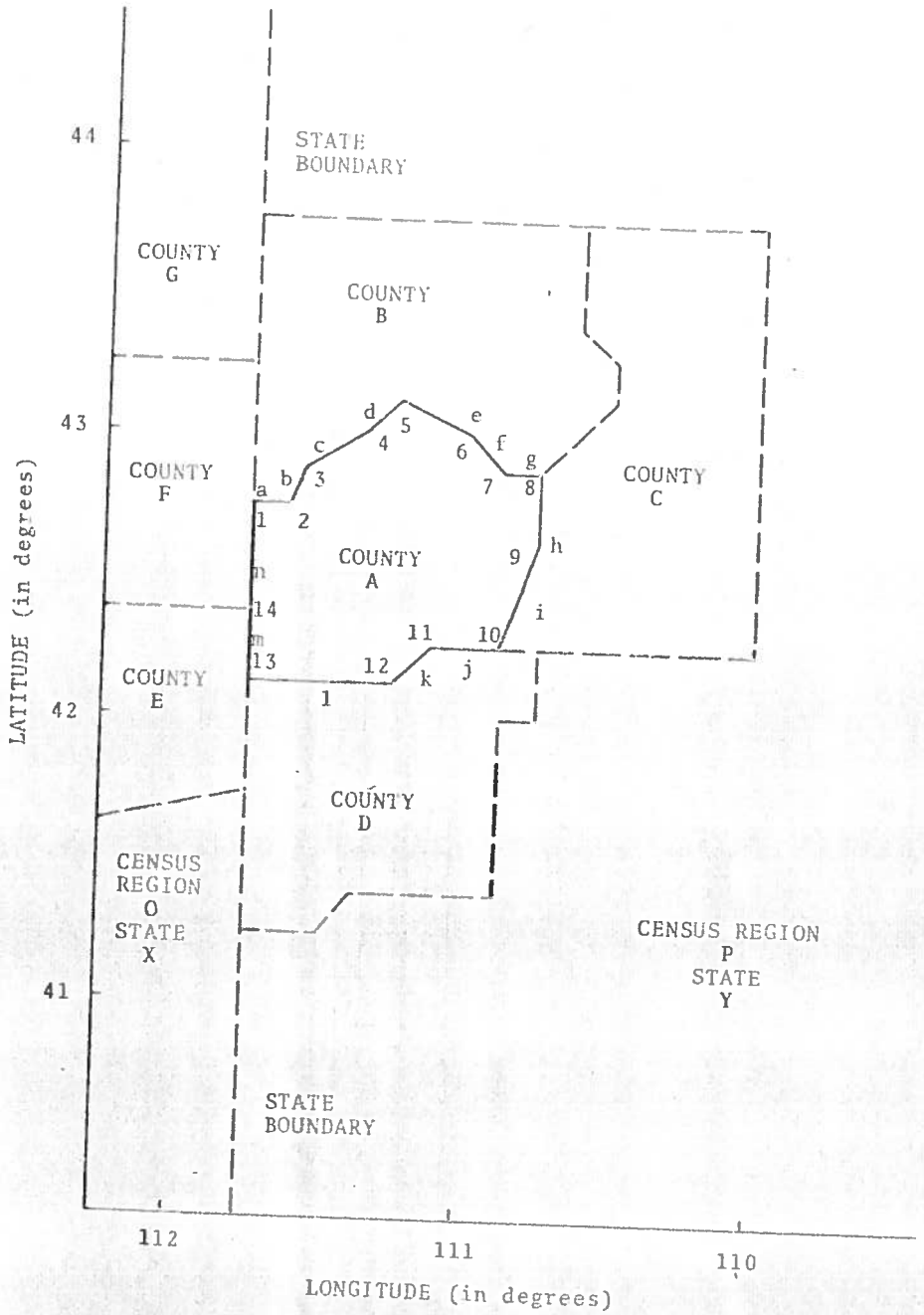
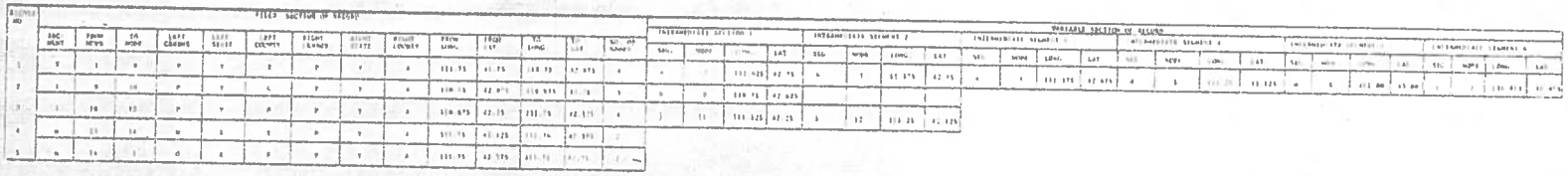


Figure A-1. Fictional County for DOT Compressed DIME File

TABLE A-1. RECORDS FOR FICTIONAL COUNTY IN DOT COMPRESSED DIME FILE



The Number of Nodes field is the sum of the number of nodes in the Fixed section of the record, which is always two (2), and the number of nodes in the variable section of the record.

Each field occupies four bytes (characters). The record of the Compressed DIME File is listed in Table A-2.

TABLE A-2. RECORD FORMAT FOR DOT COMPRESSED COUNTY DIME FILE

Field Description	Begin Position	End Position	Length	Format
Segment No.	0	3	4	Fixed Binary
From Node	4	7	4	"
To Node	8	11	4	"
Left Census Region	12	15	4	"
Left State	16	19	4	"
Left County	20	23	4	"
Right Census Region	24	27	4	"
Right State	28	31	4	"
Right County	32	35	4	"
From Longitude	36	39	4	Float Binary
From Latitude	40	43	4	"
To Longitude	44	47	4	"
To Latitude	48	51	4	"
Number of Nodes[1]	52	55	4	Fixed Binary
Segment Number[2]				
Node			4	Fixed Binary
Longitude			4	"
Latitude			4	Float Binary
			4	"

[1] This field indicates the number of nodes contained in the record

[2] A series of this field sequence representing intermediate nodes will be appended to the record. The number of these field sequences is equal to Number of Nodes field less two.

APPENDIX B

MACHINE PROCESSING STEPS

FOR OBTAINING

DOT COMPRESSED COUNTY DIME FILE

The following machine processing steps were required to produce the Compressed County DIME File.

1. Record (segment) numbers were added to each Census DIME File record in order to uniquely identify each segment.
2. Each Census record was duplicated with the TO and FROM and the RIGHT and LEFT information inverted. This procedure permits sorting on RIGHT state and county and LEFT state and county FIPS code positions to form a file which has all boundary segments for a given county sorted into adjacent records.
3. The adjacent records (segments) representing a complete county boundary were sequenced in clockwise order.
4. It was observed that the actual boundaries for Park and Teton counties in Wyoming and Fremont in Idaho were not in the file. In lieu of these boundaries, the Yellowstone National Park (YNP) boundary had been included. These segments of the YNP boundary which were not actual county boundary segments were deleted from the file. The coordinates for the above county boundaries were manually obtained (± .01 degree accuracy) and inserted into the file.

The file resulting from the steps 1 through 4 is referred to as the DOT Sequenced County DIME File. This file contains all of the boundary segments for any given county in adjacent, clockwise sequenced, segment-oriented records. A technical description of the tape file is listed in Appendix C.

5. Those records in the DOT Sequenced County DIME File common to two adjacent counties are compressed into one record in the DOT Compressed County DIME File. A more detailed description of the file is contained in Appendix A.

APPENDIX C

DOT SEQUENCED COUNTY DIME FILE