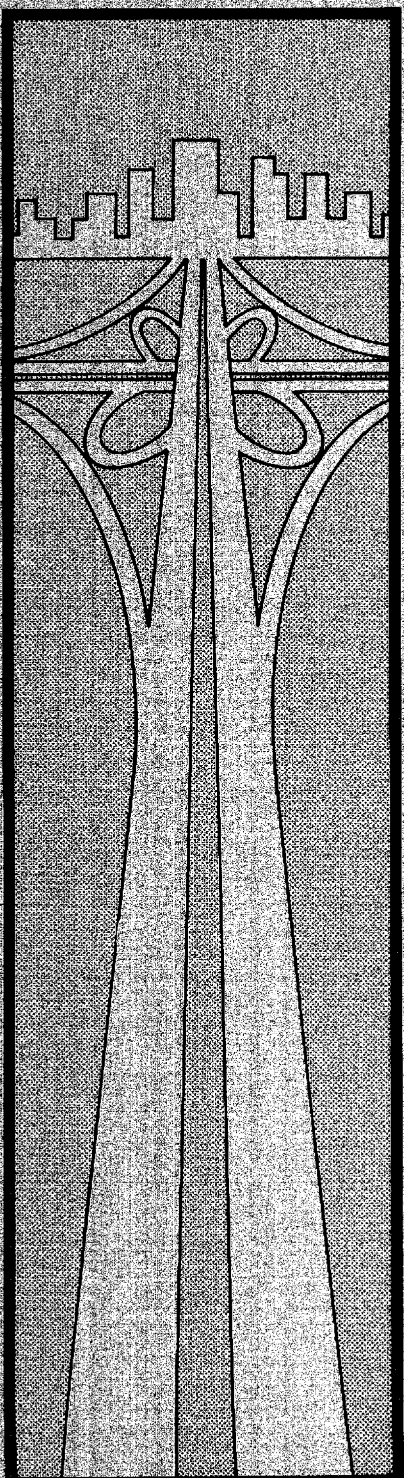
HIGHWAY PLANNING TECHNICAL REPORT



Number 7
May 1968

Coordinated Data System for Highway Planning

U.S. DEPARTMENT OF TRANSPORTATION
Federal Highway Administration
Bureau of Public Roads

CONTENTS

U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL HIGHWAY ADMINISTRATION
BUREAU OF PUBLIC ROADS
WASHINGTON, D.C. 20591

Coordinated Data System for Highway Planning

By
William E. Blessing, Highway Research Engineer
Methods Branch
Current Planning Division
Office of Planning

INTRODUCTION

The Current Planning Division conducted a review of highway planning electronic data processing in 43 State highway departments during the past year and a half. This review showed that, in many of the planning activities, improvements could be made in the manner in which computers are being used.

Two reasons that possibly explain why such improvements have not been made are:

1. Planning people have not been sufficiently familiar with computer systems analysis or programing techniques. Consequently, few changes have been made in the processing steps when a manual operation has been converted to a computer oriented operation. Insufficient consideration has been given to the interdependencies of data requirements among the different planning activities.

The data processing people doing the conversion work have known little about the interdependencies, usually because they had not been advised by the planning people and because they did not have time to study the planning requirements. The result has been that a program developed to convert an operation merely performs by computer the same steps that had been done manually.

2. Until about two years ago, many planning divisions had access only to computer systems that essentially restricted users to card input and output.

Since the processing of planning data generally involves a need for a large input and output of records, this restriction caused considerable difficulty to planning people using those computers. Many States now have computers, or are getting computers, with the potential to handle practically all planning processing needs, but planning people are still trying to use the equipment in the same way they did with older, less versatile computer systems. This report describes a system that should result in more efficient use of computers to process planning data.

TYPES OF DATA IN THE SYSTEM

Planning data may be grouped into two categories - that which can be related to the roadway and that which cannot.

1. The first group includes road inventory, traffic characteristics, condition rating, road life, historical, maintenance and accident information. It is suggested that a separate file be made for each of these kinds of information. However, road inventory data should probably be broken into four files: one file made up of cultural data, one with roadway characteristics, one with geometric data, and one containing administrative system information and locations of jurisdictional boundaries. The system should also include a file of map coordinates, which are related to the roadways, to provide for graphically displaying any of the planning data.

2. The second group includes information on finances, cost indexes, and highway statistics. Although these data should also be in separate files, the retrieval of data from them will be done on a different basis from that for the first group. These are essentially individual files that are not directly related to other files as in the first group. The basis for retrieving data items from files in the first group is always the roadway, regardless of the items wanted. For the second group, the basis is whatever item of data is needed and that item will vary from user to user.

INDIVIDUAL DATA FILES

A "Coordinated Data System" is based on establishing coordinate data files. A definition of coordinate is "equal in rank or order." The several classes of highway planning data, such as road characteristics and traffic volumes, are considered to be equal in importance and usefulness. Each class of data should, therefore, be stored in a file that is equal to, but separate from files containing other classes of data. Each file should not contain more than one class of data and all data elements in the file should pertain to that one class.

Usually, there are several sections in the planning divisions, each responsible for collecting and processing some class of data and for supplying these data to various users. In many cases, these users are coding, storing, and processing the same data elements, thus duplicating each other's efforts. In the Coordinated Data System, the section which collects the data is responsible for maintaining the basic computer file of those data.

Many of the several classes of planning data are related to roadways. In recording observations of these data along the roads, there are points where characteristics of the data change, causing new values to be recorded. These points of change, or breaks, must be recorded relative to one another.

When the data are related to the roadways, the breaks should be represented in a computer file as the information about the roadway actually exists. This is illustrated in Figure 1. The breaks, or lengths between breaks, are not necessarily the same for one class of data as they are for another, nor are they the same for one user as they are for another. It is difficult for the data recorders and users to agree on breaks compatible to each of their needs. Thus, each one ordinarily changes the breaks to fit his own requirements. However, too often breaks are based on user considerations rather than on what actually exists. This causes data observations to be broken into more lengths than necessary, resulting in duplications, extra coding, and a loss of flexibility.

An advantage of separate files is that they allow the breaks to be recorded and stored for each class of data in the way they are determined when data are collected.

By having separate files for each class of planning data, duplication in data handling can be reduced, the data file can be more current, and uniformity can be established in the basic data that are used by different groups. This is possible because newly collected information does not have to be distributed to all users, some of whom may not update their files with the new information immediately.

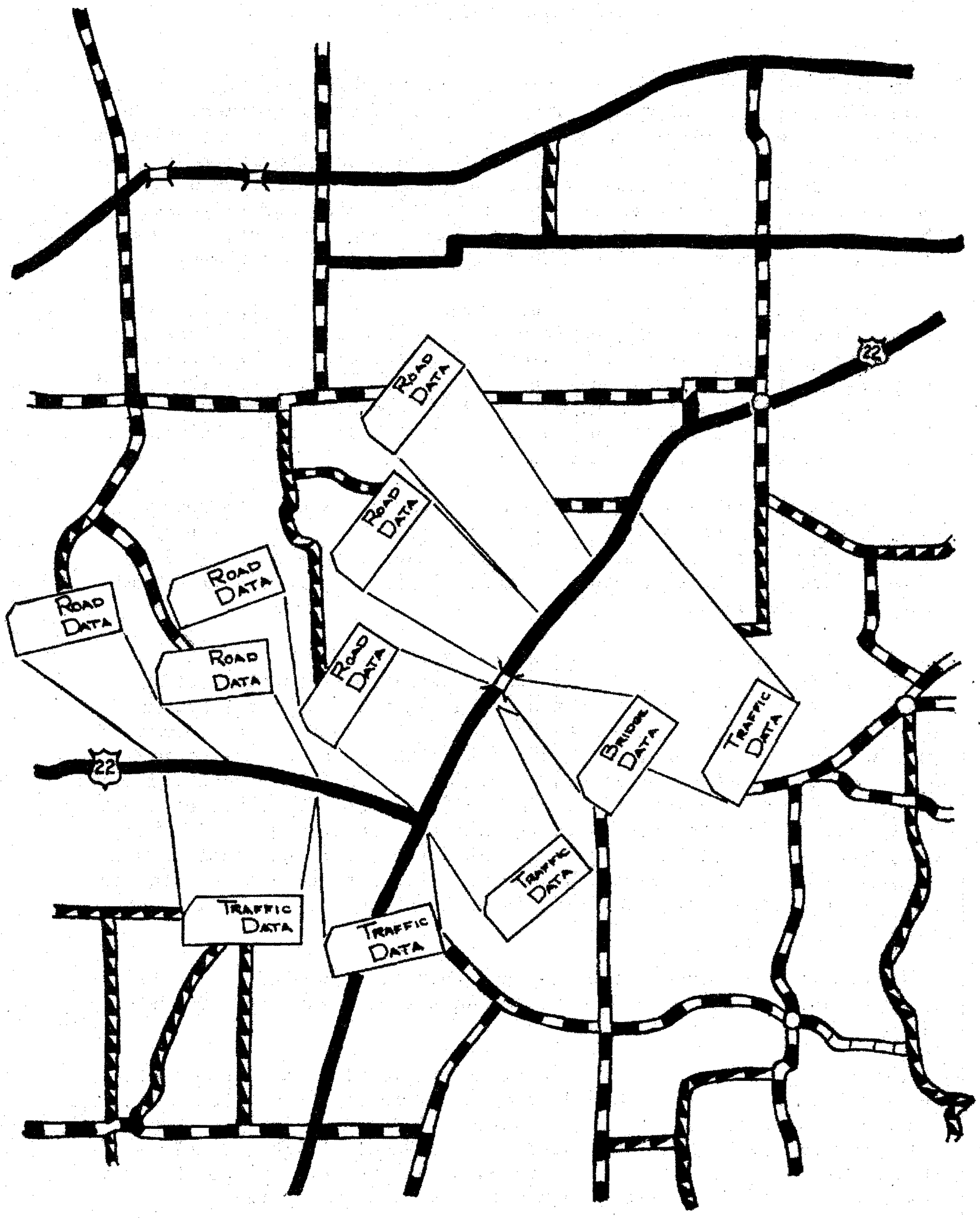


Figure 1. Storing Roadway Information as It Exists

Several States use to advantage a single planning data file containing all classes of data. However, it should be noted that the more kinds of data in a file, the more times the file is out of circulation for updating and processing and the more times the entire file is not available for either of these uses. This point is illustrated in Figure 2.

CORRELATION OF THE DATA

1. Common reference base.

There must be a way to tie the separate files together so that data can be combined by computer to produce analyses and reports. This is done in the Coordinated Data System by relating all data to a common base of reference. Since the data are related to the highways, it follows that the reference base should be the highways.

The method used to relate data to the highways should be simple to use for persons collecting and recording the data and flexible enough to allow correlation of several classes of data through computer processing. The following sections are concerned with approaches to this problem.

2. Location reference system.

Persons recording data observations along routes must have a method for locating themselves along the routes. This method may be termed a location reference system. It provides a means of uniformly reporting locations of data observations, for finding the locations of previous data observations, and for giving motorists information about their highway location.

A location reference system provides information about the location of specific points along a route. The location of non-specific or in-between points is determined by measuring from the specific points which is discussed under "Location of Data Using a Common Base of Reference."

Different kinds of location reference systems presently being used or under consideration are reviewed in the following sections.

a. Uniformly spaced mileposts. The most familiar kind of reference system is the one in which mileposts are placed along the roadway at even milepoints, the distances between the mileposts usually being the same. This has been the most widely used location system. Mileposts are normally installed on toll roads and have been placed on many Interstate highways as well as on Primary State highway routes, in some States.

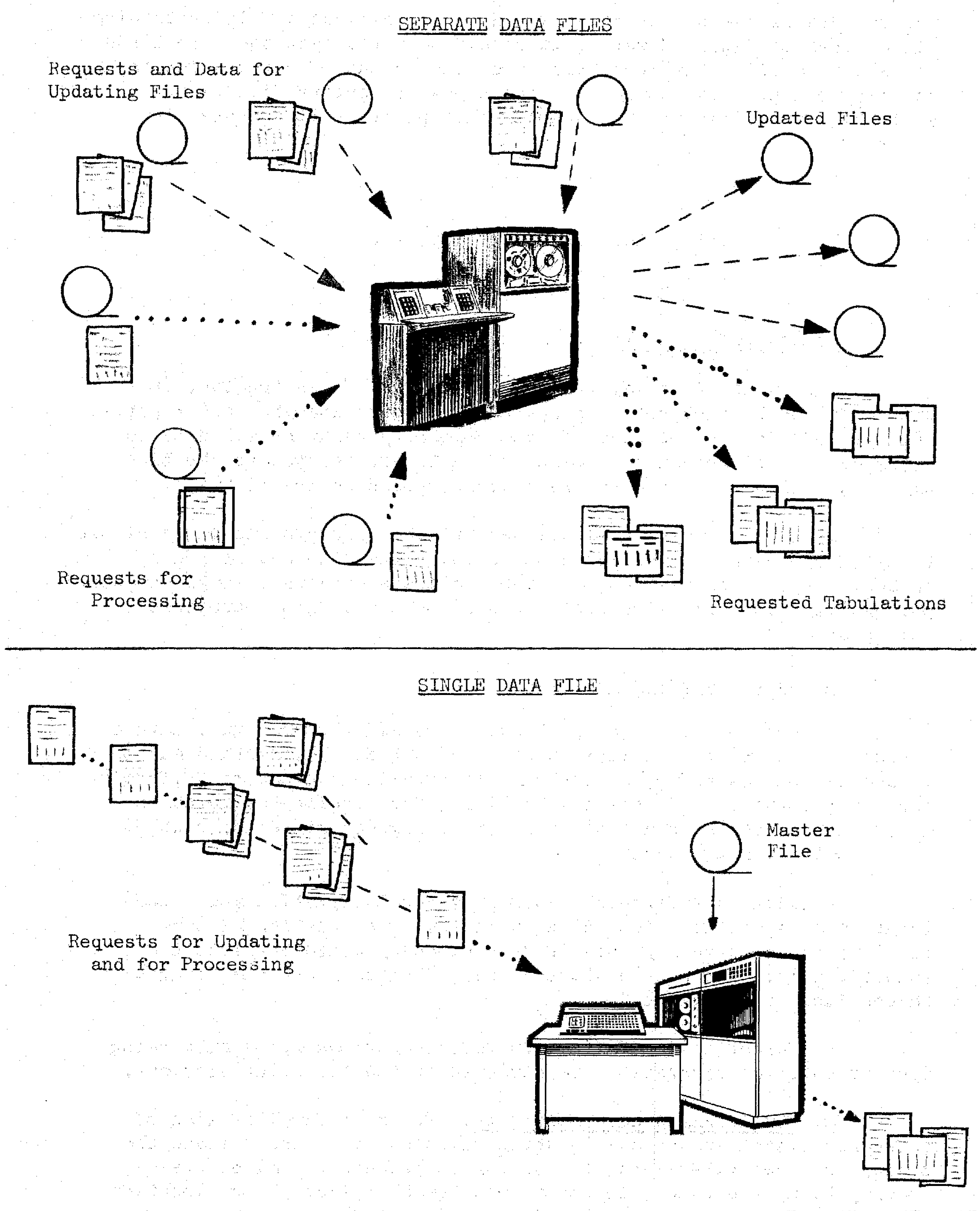


Figure 2. Separate Planning Data Files versus A Single File

The mileposts are usually signs that vary in size and shape from one State to another. The information provided on the signs also varies from State to State. The route number and the milepoint are always included. Usually a county name is given, although this may be a county number. A description of the type of signs used in 19 States may be found in a publication of the Insurance Institute for Highway Safety entitled, "A Summarized Review of Mileposting on State Maintained Highways in the United States."

Advantages of the uniformly spaced milepost method are: the method is familiar to many people; the location information is provided for the motoring public; and the uniform spacing means that one knows he has to go no farther than a fixed distance, usually one mile, to find a milepost.

Two disadvantages associated with uniformly spaced mileposts are: any construction change that shortens or lengthens the route will require a relocation of all mileposts that follow the change, and the placement of the milepoints creates a problem for maintenance forces that must work around them.

b. Nonuniformly spaced mileposts. 1/ A system similar to the uniformly spaced mileposts is the nonuniformly spaced mileposts system. The only difference is that signs are not necessarily at even mileposts, but are spaced irregularly. The type of installation is the same as noted for the uniform mileposts.

One disadvantage of this method is that construction changes necessitate a change in the sign, although in this case, the sign information alone can be changed and not the sign location. Again, maintenance forces have the problem of working around the signs unless existing signs and features are used. If the latter is used, it is probably difficult for the motoring public to find the information unless they are educated in the system.

Two States have made use of existing signs along the roadway to serve as mileposts. Mileage information is printed on some material which is then attached to the back of the existing signs. (One State uses a reflective material and the other aluminum strips).

c. Reference posts. The reference post method is similar to both the uniformly spaced and the nonuniformly spaced milepost methods. The big difference is that reference posts do not carry route or milepoint numbers but carry instead an identification number. The actual

---

1/ The report, "Highway Design and Operational Practices Related to Highway Safety," by the Special AASHO Traffic Safety Committee, dated February 1967, refers to this kind of milepost as reference markers.

locations of the signs are kept in central office records which show the county, route, and milepoint associated with each identification number.

The major advantage of reference posts is that changes in lengths of a route or in a route designation do not cause changes either in the location of the reference posts or the information shown on them. The central records are revised to show new route or milepoint numbers for the identification numbers (reference posts).

A disadvantage is that location information is not provided to the motoring public for their direct use. The problem of the maintenance forces having to work around the signs exists for this method too.

d. Route special feature log. There are no signs or posts connected with the route special feature log method. Instead, a straight line diagram or a log is kept in the central office of the milepoints associated with special features on a route. These are usually intersections and bridges, but could include monuments and other well established points along a route.

The log or straight line diagram, which is used by field personnel of highway departments and law enforcement officials, gives basically the county, route, name and milepoint for each special feature. These special features are analogous to the mileposts in the irregularly spaced milepost system and the log is analogous to the information on these mileposts.

The advantages of this system lie in the fact that no signs are needed. This means no installation costs, no changes because of construction or renumbered routes, and no maintenance problems.

This method does not give location information to the motoring public. The log or straight line diagram must be maintained on a current basis. This requires that revised sheets of the log or straight line diagram be sent to people using the system and that they replace old sheets in their possession with the new ones.

e. Coordinates. Another procedure that has been tried is that of using coordinates, usually State plane coordinates, to locate points of data observations. This method has been considered in connection with accident location. There have also been attempts to record locations of traffic and inventory sections by coordinates in order to tie these data to accident records.

Location of data in the field by use of coordinates requires a map along with a template or some other means of scaling the coordinate location on the map. A variation, which has been considered in one or two States, is to record a distance and direction from a point having known coordinates and then compute the actual coordinates of the desired point in the central office.

The major advantage of using coordinates is that a coordinate is a fixed point in space and is unaffected by changes in route length or name. Thus, a point on a route located by a coordinate identification is always located at the same place regardless of route changes.

There are four major problems associated with the use of coordinates as location identification:

(1) Special equipment and, in many cases, training is required to get the coordinate of a position.

(2) The process of getting a position coordinate is a slow one when compared to other methods. This makes it impractical for a field crew or even an office force to use this procedure for recording locations of planning data along highways.

(3) From a data processing standpoint, coordinate identification makes it difficult to sequence records in their proper relation to one another. Further, if data are wanted between any two points, a manual effort is needed to determine the coordinates of those two points.

(4) There is no location information available to the motoring public.

f. Most favorable system. Whatever system is used, an education process is needed to teach people to use the system properly and an updating process is needed to keep the location information current.

Thus, making a decision as to which system is best requires consideration of the costs of installation, the educational effort required, the flexibility of the system (since it must be a practical procedure for locating all types of data), and the costs of keeping information current and accurate. A consideration of the extent to which the motoring public is served may also be pertinent.

Probably the most economical and practical method is to use reference posts. An educational effort can teach people, including the motoring public, to measure and report locations using reference post numbers as well as they can using mileposts. Additionally, the reference post numbers can be a combination of route numbers and milepoints, or just milepoints such that they closely reflect the actual mileages, unless a route was lengthened or shortened by a large distance. Generally, the numbers will be sufficiently accurate that the motoring public will still be provided with useful information. Also, the numbers could be the State plane coordinates of the reference post location, if desired.

It may be easier, or possibly necessary, to locate certain data by route and milepoint. It should be recognized that data can be related to the roadway by a combination of any of the methods described and still be compatible since the reference base is the same.

3. Location of data using a common base of reference.

Part 2 under CORRELATION OF THE DATA described ways to locate roadway characteristics in the field. This part of the report will cover ways of identifying in data records the location of data relative to the roadway for data processing purposes. The following procedures provide the means of coordinating several different planning data files effectively even though the files are physically separate.

a. Route and milepoint. Probably the simplest and most familiar method of relating data to routes is to use route numbers and milepoint identifications in all data records.

With this method, every record in the data files is related to a route by recording a route number and to a section of that route by recording a beginning and ending milepoint, or a milepoint and a section length. Figure 3 illustrates this procedure by adding a beginning and ending milepoint to the records. It should be noted that there is no record in either of the files having both beginning and ending milepoints that are the same as the beginning and ending milepoints in any record of the other file. The milepoints in Figure 3 were chosen to illustrate how a group responsible for one of the files may record necessary milepoints and to show how it is unnecessary to match the breaks that were chosen by another group that is responsible for maintaining another file. The milepoints are the distances from an origin point on a route.

There is some opinion that the origin point should be at the beginning of each route in a State. However, many people think the origin points should be at every county line. If the latter procedure is used and if data are to be represented in storage as they exist on a route, a code must be provided in every data record to sequence the counties along that route. Such a sequencing code is necessary also for the control sections that are presently being used in several States.

(1) Problem of changes in route length. Two problems are associated with the use of milepoints and route identification.

a. Construction changes which lengthen or shorten a route make obsolete the milepoints previously used in the data records affected by the changes.

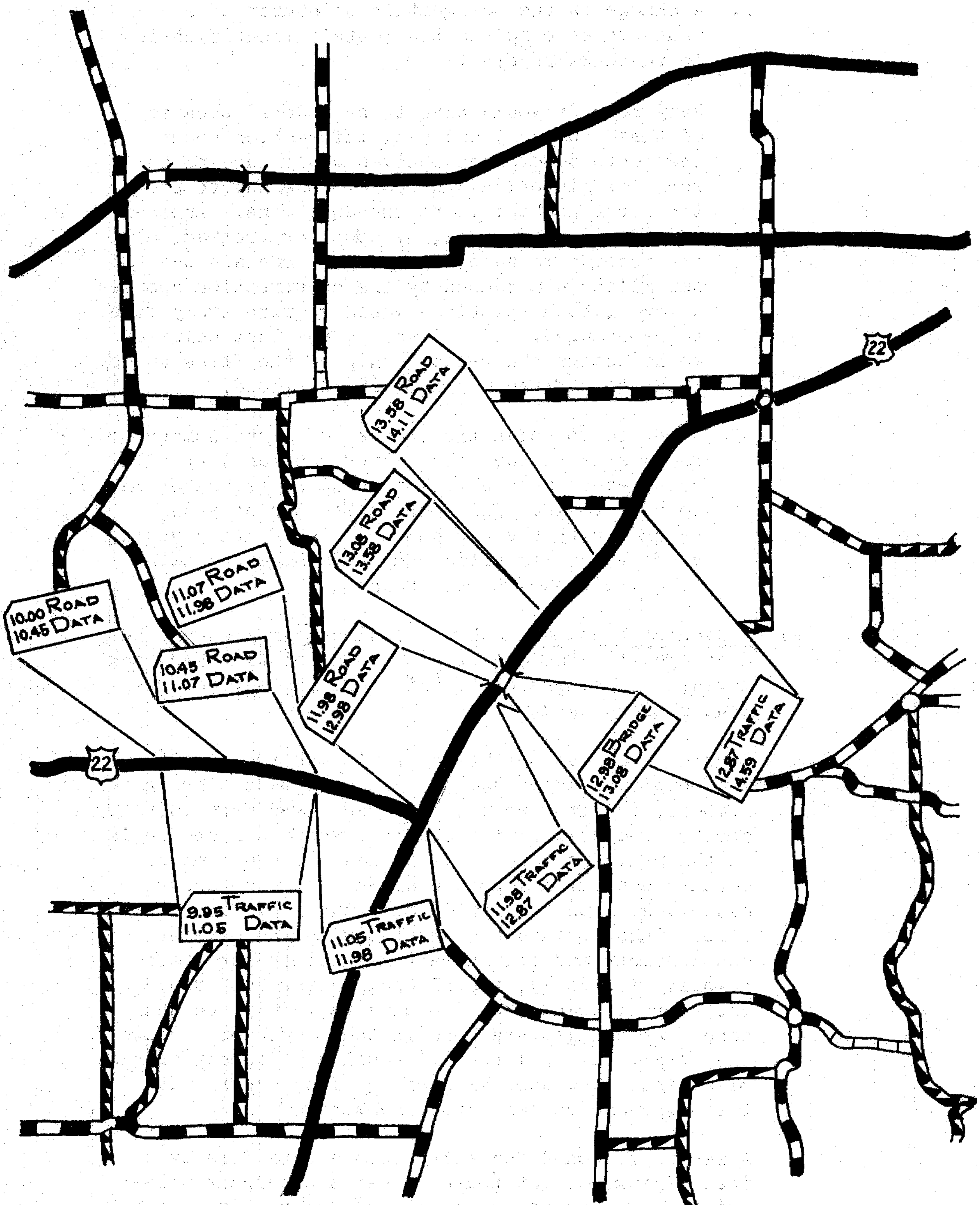


Figure 3. Relating Information to the Roadway With Beginning and Ending Milepoints.

b. A change in the designation or number of a route makes obsolete the route's identification in the data records.

Many records would have to be updated because of these changes; and yet, information about the route before the changes would need to be kept for historical purposes. One way to solve the first problem is to use equations. Another solution is to change, by computer program, all the records in each file so they contain the new milepoints caused by the construction changes. However, both solutions would require every file to be changed. Additionally, the last solution would change the relationship of the files to any location reference system in the field.

A method to solve the problem of route number changes is to have a computer program that will write every record with the old route number into a historical file and update the existing records with the new number. Again, this would cause a loss in efficiency because every data file would need to be changed.

(2) Missouri's method. Other solutions can probably be developed for these problems. One solution has been devised by members of the Missouri highway department's planning division.

Corridors across the State are formed from latitude and longitude coordinates. These corridors are given numbers, odd from west to east and even from south to north. The number of a corridor where a route begins in the State is used to get a basic route number. Certain control milepoints along the route are associated with the basic route number in a control file. These points are equated with the current route number and the current milepoints for those points. Initially, the milepoints for both basic and current numbers on the same route will be the same. As changes are made in the length of a route, the basic milepoints will equal a different current milepoint. The same kind of equality occurs with the basic and current route numbers.

A record is coded for entry into a data file by identifying current route number and current milepoints. The updating program converts this identification into the basic route number and basic milepoints from the control file and enters the record into the data file with basic identification. Data are retrieved from the file in the same manner.

This method has two good points.

a. Only one file must be updated with changes in route lengths and numbers.

b. No matter what the current milepoints and route designations change to over the years, data are stored and retrieved for the same sections of roadway. This is illustrated in Figure 4.

With a method such as this, it is possible to determine the period of useful historical data for each file in the system, two, three, five, or more years, and carry data for these periods on the same file as the current year's data.

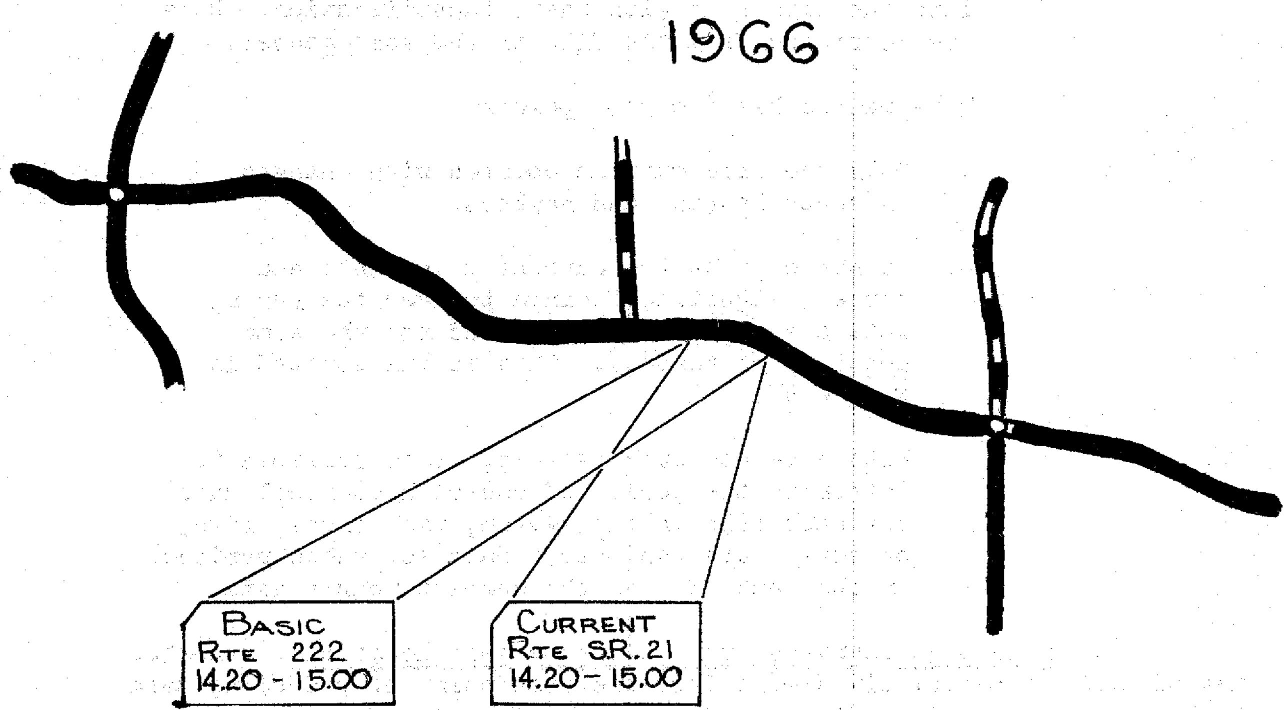
b. Milepoint computed from reference post locations. Another method that is relatively simple from the standpoint of recording data in the field involves coding a distance and direction from a reference post and the reference post number. This information is then used by a computer program in conjunction with the reference post location file to compute and store route and milepoint identification in the data records.

The statements made in section 2. (c) on page 7 also apply to the records developed under this method. The principal advantage of this method is the ease of location and recording of data in the field. The field crews do not have to know what the actual location or route number is, only the reference post number.

A disadvantage is the fact that every different route must have at least one reference post, including very short side routes.

c. Length identification using reference posts. A variation of the method just discussed gives the same simplicity and greater flexibility. The procedure is to store in the data records, as the location identification, the distance and direction from a reference post and the reference post number. This may be done for both ends of a road segment or for just one end. In the latter case, the length of the segment is stored.

The milepoint locations of the segment ends are computed by using the reference post location file. Thus, this method does allow for retrieving data with a milepoint specification.



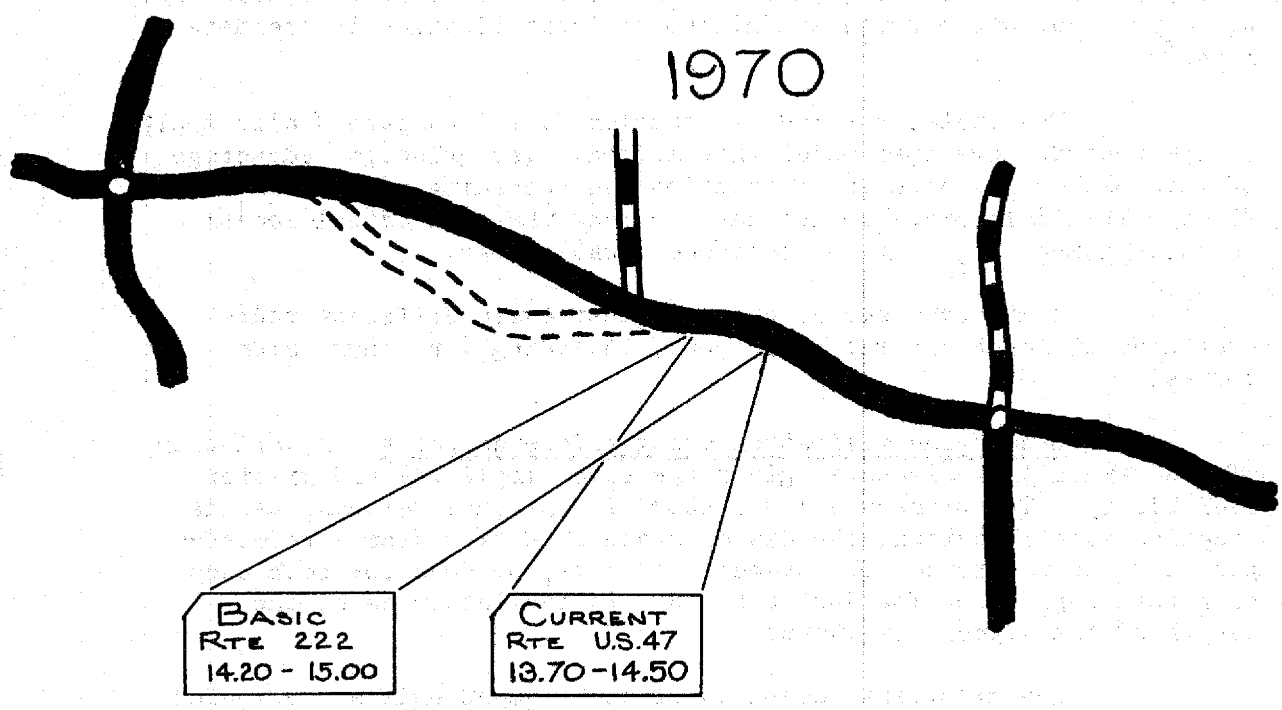


Figure 4. Showing relationship of Basic identification to Current identification before and after route and construction changes

An advantage of this method, besides the ease of recording in the field, is that the records in the data files do not have to be changed when there is a change in the route length or in the route number. Such changes would only affect the reference post location file. This would then require only making changes in one file as compared to having to change many other data files that would be affected under other methods.

d. Point identification using reference posts. Another method is to store in a data record the distance and direction from a reference post and the reference post number, of the point where some data characteristic changes, that is, at a break in the data. The data in the record is representative of the roadway until the point is reached where the data changes. At this point, another record is stored with the point located as described above.

As in the previous method, the actual milepoint would be computed from a reference post location file.

This method allows simple field recording, and data records in the files do not have to be changed because of changes in route location and number. There is an additional advantage with this method, new data records can be inserted anywhere in any file without changing any other record. With any other method discussed, insertion of a new record would cause at least one existing record and possibly more to be changed with regard to length.

e. Most favorable location identification method. Of the methods discussed, either of the reference post methods is believed to have the most flexibility. The reference post method using point identification is probably better for updating or file maintenance, since records in a file can be added with no affect on other records. The reference post method using length identification may have an advantage from a processing standpoint, since a segment length is inherent in each record and it is not necessary to read the next record to find the segment length.

Either of these methods can be used in conjunction with records identified by route and milepoint. The common reference base makes the identification methods compatible for data processing.

EXAMPLES OF HOW THE COORDINATED DATA SYSTEM MIGHT BE USED

1. Route and milepoint identification

Assume that a series of separate data files have been established as discussed in this report. Also assume that two of the files contain road characteristics data and traffic characteristics data that are related to the roadway through route and milepoint identification in the data records. This is shown in Figure 3.

Although not shown in Figure 3, assume that there is a file of accident data for accidents that have occurred at several points along the route, and a file of condition ratings that includes records for the route between milepoints 10.00 and 12.00, between 12.00 and 12.98, between 12.98 and 13.08, and between 13.08 and 14.59.

Finally, assume that some department head wants to know something about Route 22 between milepoints 11.00 and 13.00. Say that he wants to have information about the accidents that occurred during the past year in this 2-mile section. He also wants to know about the surface widths and the adequacy ratings between the given points.

A computer program takes the given milepoints and the information wanted as control parameters. It then retrieves the appropriate data, and prints the desired information. The printed output for this case could look something like that shown in Figure 5, keeping in mind that the numbers shown are arbitrary.

The example used was to answer, via a special purpose computer program, a request for certain information that is derived from a particular section of road. The termini of the section was given as request parameters. It should be remembered that many routine and periodic programs can be developed with such parameters built in. It is likely that such programs would not list information as shown in Figure 5 but would make some computation or analysis with the retrieved data. The output in these cases may be anything from a single line to a complete tabulation.

2. Reference post and distance identification

This example shows how distances to reference posts can be used as identification in data records to relate data to a roadway.

For comparison with the use of route and milepoint identification, assume, for this example, the same data files and records used in the route and milepoint example above.

ROUTE U. S. 22

BETWEEN MILEPOINTS 11.00 AND 13.00

ACCIDENT NUMBER	LOCATION	DATA
XXXXX	11.01 TO	
XXXXX	11.90 TO	
XXXXX	11.95 TO	
XXXXX	11.97 TO	
XXXXX	11.97 TO	
XXXXX	12.00 TO	

SURFACE WIDTH	LOCATION
22.0	11.00 TO 11.07
23.0	11.07 TO 11.98
24.0	11.98 TO 12.98
24.0	12.98 TO 13.00

TRAFFIC VOLUMES	LOCATION
1555	11.00 TO 11.05
1500	11.05 TO 11.98
1950	11.98 TO 12.87
1975	12.87 TO 13.00

ADEQUACY RATINGS	LOCATION
75	11.00 TO 12.00
90	12.00 TO 12.98
85	12.98 TO 13.00

Figure 5. Sample output for example of Coordinated Data System usage.

Assume a reference post location file exists that stores the actual locations of reference posts by route, county, milepost (from beginning of route in the State) and suppose that the reference posts placed as shown in Figure 6 have the following information in a reference post location file:

Reference post	Milepost	Route	County
380	9.00	US 22	Jason
384	11.25	US 22	Jason
388	14.00	US 22	Jason

Two ways in which reference post identification can be established were discussed in the section, CORRELATION OF THE DATA, 3.c. and 3.d. Assume that the data records of the route and milepoint example above are identified as described in the section, CORRELATION OF THE DATA, 3.c., in which both ends of a road segment are stored in a data record. Figure 7 shows the identification information used in the route and milepoint example and the corresponding identification used in this example. The beginning points on each line are the same and the ending points are the same.

As in the previous example, assume that a department head wants to know how many accidents have occurred between milepoints 11.00 and 13.00 and what the surface widths and adequacy ratings are between these points. As stated before, a computer program takes the given milepoints and the information wanted as control parameters. The printout would be the same as that shown in Figure 5 for the previous example.

A variation of this procedure, which is described under the section, CORRELATION OF THE DATA, 3.d., is to identify in the data records the beginning points only of the road segments they represent. If this is done, the printout would still be as shown in Figure 5 and the control parameters would still be as shown in the two examples.

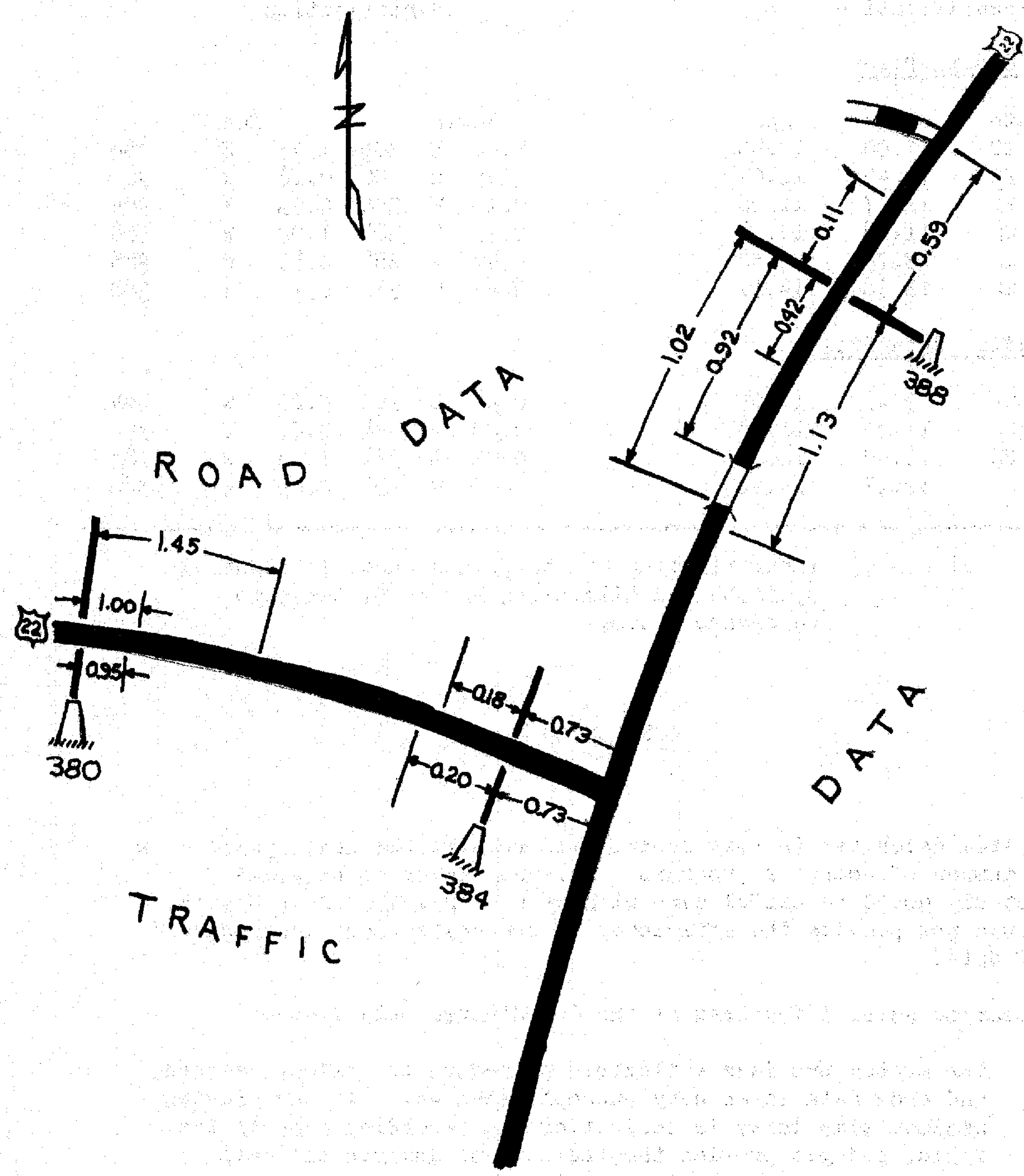


Figure 6. Locating Points by Relating Them to Reference Posts

Route and milepoint identification			Reference post and distance identification					
Road data file								
Route	Begin	End	Begin			End		
US 22	10.00	10.45	1.00	E	380	1.45	E	380
US 22	10.45	11.07	1.45	E	380	0.18	W	384
US 22	11.07	11.98	0.18	W	384	0.73	E	384
US 22	11.98	12.98	0.73	E	384	1.02	W	388
US 22	13.08	13.58	0.92	W	388	0.42	W	388
US 22	13.58	14.11	0.42	W	388	0.11	E	388
Traffic data file								
US 22	9.95	11.05	0.95	E	380	0.20	W	384
US 22	11.05	11.98	0.20	W	384	0.73	E	384
US 22	11.98	12.87	0.73	E	384	1.13	W	388
US 22	12.87	14.59	1.13	W	388	0.59	E	388

Figure 7. Identification of the points shown in Figure 3 by Route and Milepoints and by Distance to Reference Posts.

CONCLUSION

The system described in this report will require the development of a large number of computer programs. A large number of programs undoubtedly would be useful even without a Coordinated Data System, but would not provide the efficiency in the collection, storage, and use of data.

To summarize several features of the Coordinated Data System:

1. The system provides a flexible procedure to collect, record, and code data in an easy uncomplicated way. It can provide maximum simplicity in collecting and recording data in the field, and yet provide the facility to analyze all data sources with any relationship desired.

2. Data can be collected as it exists without having to consider how a user needs the data.

3. Data values are stored in a computer file in the way they exist in relation to one another along a route.

4. Data can be retrieved according to the user's specifications. The user can establish control sections or segments of any length for the purpose of summarizing data or retrieving information.

5. The system is modular. It can be developed one file at a time. Additional files (that is, new data) can be added to the system at any time without disturbing any other file. Such additions would also require development of new computer programs.