Roadway Modeling and Data Conversion for a Transportation Facilities Information System

W. K. BOTTIGER AND W. P. KILARESKI

Many transportation agencies are in the process of developing roadway management systems to assist with their rehabilitation and maintenance programs. All require a centralized data base to manage the large amount of data collected for each highway section. Described in this paper is a research project that used the utility industry (gas, electric, and so on) interactive graphic data base for highway applications. As a roadway section is similar to an electric line, many of the modeling concepts are therefore also similar. A computer model of the highway system was developed that used five types of facilities. Type 1 is a point facility (sign, signal), whereas Type 2 represents a highway span facility. The Type 4 and 5 facilities represent the data elements and the pictures associated with the interactive graphics. All data are stored at appropriate *X* and *Y* coordinates. The roadway model developed with the five facility types allows the user to trace the connectivity of the highway route as well as to obtain the requested highway information. Also described in this paper is a case study that was conducted to test the model. U.S. Geological Survey maps were digitized and merged with the Pennsylvania Department of Transportation's Systematic Technique to Analyze and Manage Pennsylvania Pavements condition information for roadways in a county in Pennsylvania. It was found that the facility models adequately describe the highway system. The user, however, must be careful in the digitizing process. Map document shrinkage and expansion (due to humidity) is enough to change the document location reference. It was also found that manual digitizing is extremely time consuming, and it is necessary to develop computerized interface data conversion routines.

The highway network in the United States represents a capital investment of more than one trillion dollars. The operation of the network has a direct effect on the social, economic, and political well-being of the nation. For many years the nation's highways have served the public efficiently; today, however, the roadway infrastructure is deteriorating rapidly. Unfortunately, the rate of deterioration is much greater than that at which repairs and rehabilitation can be accomplished. Curtailed revenues, increased construction costs, and increased truck traffic have created a dilemma for highway administrators. Often, the highway engineer has so many projects to undertake, he does not know which one to program first. Consequently, many transportation agencies are developing pavement management systems (PMS) to help manage their highway network.

W. K. Bottiger, Arthur Andersen & Co., 5 Penn Center Plaza, Philadelphia, Pa. 19103. W. P. Kilareski, The Pennsylvania Transportation Institute, The Pennsylvania State University, University Park, Pa. 16802.

The term "pavement management," or "roadway management," describes strategies used at various levels of highway administration. Pavement management systems usually encompass "all activities involved in providing and maintaining pavements at a satisfactory level of service. These activities range from initial data gathering to planning, design, construction, maintenance, rehabilitation, and periodic monitoring of existing pavement condition" *(1, 2)*. A pavement-management system can also provide the information required to determine alternative strategies as well as the optimum treatment required for a particular highway segment. During the past several years many PMS have been developed and implemented at the state and local level. Some PMS are designed for the network level while others concentrate on project-level needs. Although the structure of a PMS can be complex, all PMS use a computerized data base. In fact, all pavement management systems are useless without the data base. Data processing needs range from large mainframes to microcomputer systems.

Highway agencies, especially ones with extensive highway networks, require large amounts of diversified information to manage their systems. This information may contain data about a pavement's history, such as construction records, material specifications, and as-built standards. It may also include the current status of the highway with respect to longitudinal roughness, skid resistance, and distress conditions. Such ancillary information as accident location, culvert location, guiderail status, and bridge data also serve useful purposes.

Some highway agencies have been collecting this inventory and condition information for decades while others have only recently begun to gather information about their system. Because each agency may be responsible for thousands of miles of highway, the problem usually is not whether there are data available, but rather that there is too much information to evaluate. Frequently, data are found at many different locations within an organization. For example, accident information is maintained in the safety unit, truck traffic data in the traffic engineering section, material records in the construction section, and maintenance activities in the maintenance group. Yet the highway administration needs all of this information to make sound decisions.

For a PMS to be effective, the data collected throughout the agency should be integrated both to support analysis applications and to provide for ease of information access for everyday uses. One means of providing this ease of access is through use

of computer graphics as a key to information in the data base. In this scenario, the information available to the highway administration can include not only the descriptive data (structural properties, maintenance history, and so on) but also the location and connectivity (logical relationship of one location to another) of the various facilities that make up the highway and bridge network.

Three major problems result from the volume and diversity of the data required:

1. As data are collected from a number of sources over a period of years, it is difficult to maintain an accurate, up-to-date, and centralized data base.
2. Because of the diversity of input sources, it is difficult to maintain an accurate and complete data base when revisions and updates are required.
3. The large amount of data often makes retrieval of a specific subset of data a time-consuming and tedious task.

These problems may seem insurmountable to a highway agency that has always worked with isolated and independent data sets. However, highway agencies can benefit from the experience of the utility industry, which has a successful history of managing information on geographically distributed facilities, as is discussed in the next section.

UTILITY DATA-BASE MANAGEMENT SYSTEMS

For many years, utility companies (e.g., power, gas, water, telephone) faced the problems of maintaining information about their assets, which are distributed over a large geographic area. Circuit information, pole status, transformer location, and other data were maintained on paper records or maps. Consequently hundreds, and even thousands, of maps and records had to be manually maintained. The problems faced by the utilities were the same as the data-base problems now facing highway agencies. As most utility companies allocate considerable resources to keeping their facility records current and accurate, considerable work went into the solution of these problems. The result of a joint research effort between the utility industry and IBM was the development of the distribution facilities information system (DFIS).

The primary goals of a DFIS are to reduce the cost of maintaining facilities records, to store the records in standard form on a computer data base, and to make the facilities data more readily available in the form best suited to user requirements *(3)*. The DFIS data processing includes two concepts: interactive graphics, which is used to maintain a defined data structure, and a geo-facilities data base.

The interactive-graphics capability is provided by the Graphics Program Generator (GPG) software. GPG is a "set of programs designed to create, maintain, and display information about facilities, their locations and relationships to one another" *(4)*. Through an interactive conversion with the graphics work station, GPG stores data in a structure customized by the using agency. This data structure contains the attribute data for each facility, as well as that facility's location and connectivity to adjacent facilities. The work space developed by one user on a single work station represents a subset of all the agency's data. Once the user has completed the adding or modifying data in his geographic area, the data are transferred to the geo-facilities data base. GPG provides the means for sending data to, and receiving data from, the data base. The maintenance of the geo-facilities data base is provided by the Geo-facilities Data Base Support (GDBS) software. It maintains the hierarchical data model (structure) developed by GPG.

Experience has shown that the majority of data-retrieval requests are by geographic area, and the data model used by GDBS is designed to facilitate such requests. Facilities located near each other in the field are also stored near each other in the data base. An additional component of the data structure permits the connectivity of facilities to be explicitly represented. With this concept incorporated in the data structure, data retrieval for a network can be performed quickly.

The data base maintained by GDBS is continuous. One data base represents the entire area for which an agency is responsible, so there are no breaks introduced by map edges. This centralized data approach ensures that, as data are entered into the system, the information is immediately available to all users. Thus, all data are current and the question of "What data are up-to-date?" can be avoided. The area retrieval capabilities mean that data can be obtained for specific geographic areas, such as political boundaries, service regions, engineering districts, and others. The network-retrieval capabilities, in contrast, permit a user to obtain data for any network of facilities available in the data base.

At the present time the DFIS is being used primarily by utility companies. Consequently, all of the development and implementation work has involved the writing of menus, pointing rules, and so on, for utility applications. As the utility land base and other facilities are similar to transportation facilities, it is logical to manage pavement management system data with GPG. Therefore, the objective of this research was to build the foundation for a transportation facilities information system (TFIS) that will address the issue of collecting and maintaining data on an agency's highway and bridge network.

Whenever interactive graphics is used with a data base the fundamental question of data modeling must be addressed. Too often, little thought is given to the structure of the data base. Consequently, the highway agency is forced to accept a data base that was designed for some other application. The users quickly learn that they must continually modify their operation to make the data base system work, rather than have the data base support their needs. A basic question for PMS is how the roadway system should be modeled in the data base. The research discussed in the paper was designed to create a roadway data base that realistically models the highway system.

ROADWAY MODELING CONVENTIONS IN TFIS

The GPG and GDBS software products allow considerable flexibility in how data can be modeled within a TFIS. This is desirable, as transportation organizations need to manage information that varies widely in form and purpose. Some types of information will rarely need to be updated, including information on the geometry of the highway section, which typically remains the same until major reconstruction is performed. Other types may need to be updated yearly or even more often.

There are five facility- and data-modeling options available within the framework of the GPG/GDBS system. These conventions are named Type 1 through Type 5, and each has a

specific data-modeling purpose. This discussion will outline each of the conventions and describe what facilities and data records in a highway and bridge network have been modeled using each convention, as well as provide a description of the attribute data associated with that facility.

It should be noted that this discussion involves only those facilities that were developed specifically for the TFIS by the project team. It does not cover such items as work entities or the entities used for the plotting of data. The facilities discussed are specifically those that would be of interest to a transportation organization implementing TFIS, and are items that are expected to be found as parts of the physical highway and bridge network.

Type 1 Facilities

Type 1 facilities are those that exist at a single set of *X*, *Y* coordinates and are attached to a single point connector at that location. Point connectors are defined as a method of modeling explicit expressions of connectivity between facilities at the same location *(5)*. The logical data model associated with Type 1 facilities is shown in Figure 1. The Type 1 convention is typically used to model facilities that have the ability to stand alone at a single physical location. Type 1 facilities do not have the ability to control connectivity across a single *X*, *Y* location. In a utilities application, items that are typically modeled as Type 1 facilities include poles, manholes, landmarks, and lot and block locations of utility service customers.

Type 1 facilities may also have subfacilities or repeating data groups associated with them. A subfacility (Type 4) is a facility that is subordinate to a Type 1, 2, or 3 facility in the data hierarchy and that has a picture associated with it. A repeating data group (Type 5) does not have a picture associated with it, but is capable of storing a number of records of the same format, typically of a periodic nature. The example commonly used in utilities applications is the situation in which a power pole is modeled as a Type 1 facility, has a guy wire (a subfacility of the pole) modeled as Type 4, and has a series of yearly inspection records for the pole (a repeating data group) modeled as Type 5. The facilities that were modeled as Type 1 for highway application include: intersection, culvert, sign, signal, and railroad crossing.

Type 2 Facilities

Type 2 facilities are those that exist between two sets of *X*, *Y* coordinates and are attached to a point connector at each end. The logical data model associated with Type 2 facilities is shown in Figure 2. The Type 2 convention is typically used to model facilities that make up the network and for which connectivity information must be maintained. A Type 2 facility is typically any facility that can be represented as a line, straight or otherwise, between two points. As most facilities of interest to transportation organizations are of the span type, this is a heavily used convention in TFIS. In the utilities application area, items that are typically modeled as Type 2 facilities include pipes, primary and secondary circuits, and property lines. Type 2 facilities may also have subfacilities or repeating data groups associated with them, as shown in Figure 2. It should be noted that most of the Type 4 and Type 5 facilities used in TFIS are subordinate to Type 2 facilities. Subfacilities and repeating data groups are described in detail in the following sections on Type 4 and Type 5 facilities. The facilities that were modeled as Type 2 for highway application include: highway, tunnel, bridge, railroad, river, and boundary.

Type 3 Facilities

Type 3 facilities are those that exist at a single set of *X* coordinates and are attached to two point connectors at that location. The logical data model associated with Type 3 facilities is shown in Figure 3. In utilities applications of GPG/GDBS, Type 3 facilities are typically used to model those facilities that control flow through the network, whether it be

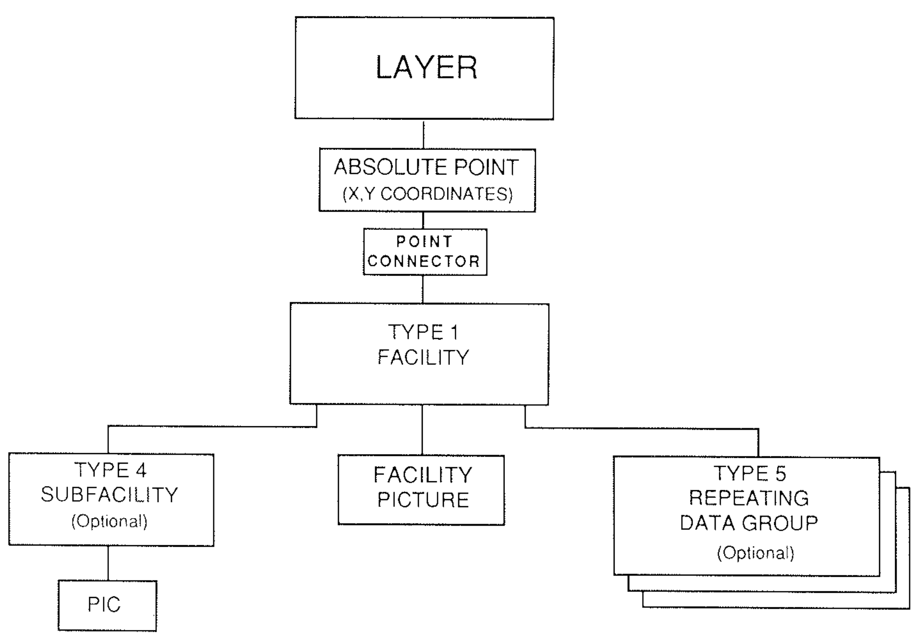

FIGURE 1 Data model for Type 1 facility, point facility.

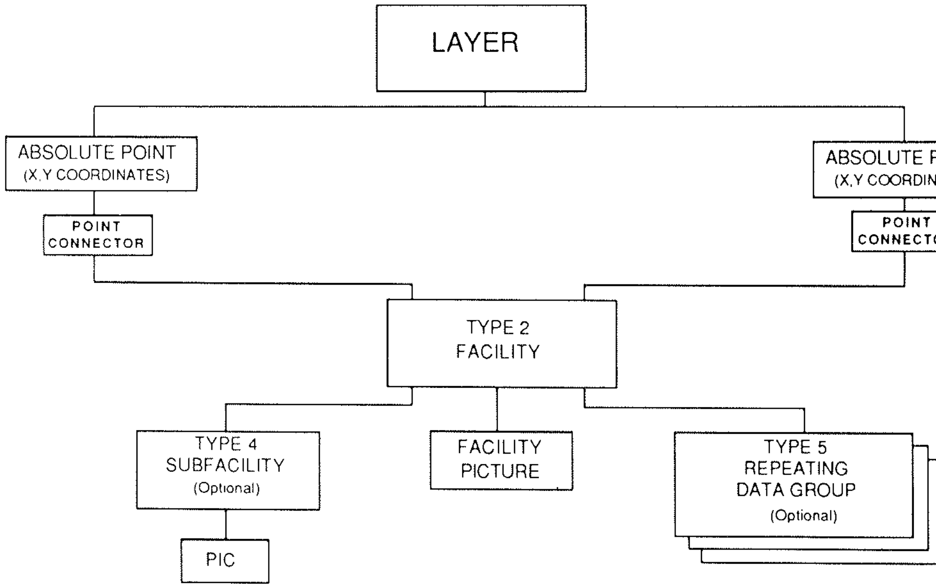

FIGURE 2 Data model for Type 2 facility, span facility.

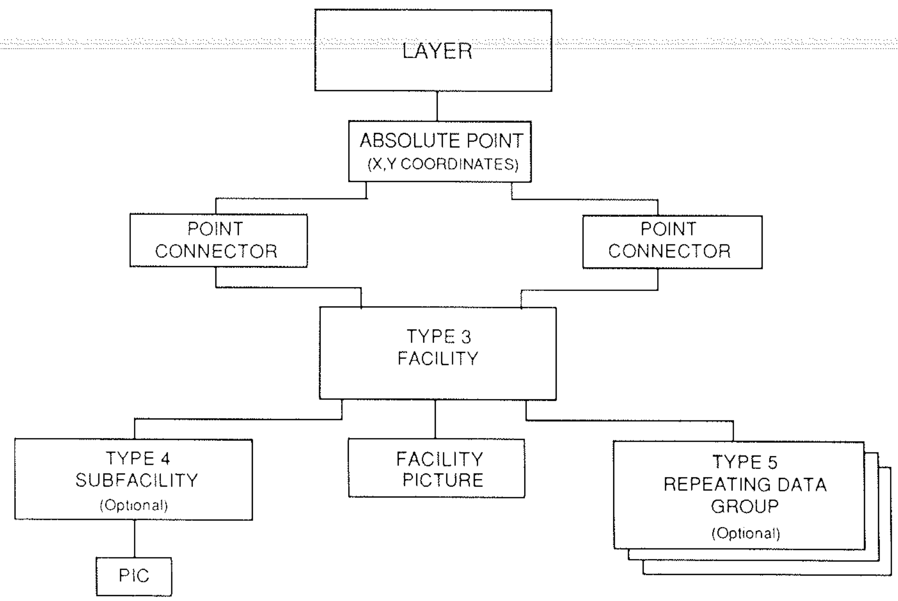

FIGURE 3 Data model for Type 3 facility, connector.

electricity, water, or natural gas. Such items as values, transformers, and switches are normally modeled as Type 3. In general, Type 3 facilities are useful for modeling items that control some type of flow across a point.

No use for the Type 3 facility has been found in the TFIS project to date. Although some consideration was given to the possibility of modeling intersections as Type 3 facilities, it was discovered that the subroutine that expands the centerline on highway facilities (draws curb lines parallel to the centerline at one-half the width) would not work properly if intersections were modeled as Type 3. It was subsequently decided simply to attach all of the approaches to an intersection to a single point connector, and model the intersection as a Type 1 facility attached to the same point connector. An organization implementing TFIS should, of course, evaluate its particular data-management needs to see if any facilities should be modeled as Type 3.

Type 4 Facilities

Type 4 facilities are typically subfacilities of facilities modeled as Type 1, 2, or 3. Thus they are subordinate to their parent facilities, and if the parent facility is deleted from the work space, the subfacility is also deleted, as it is lower on the data hierarchy than the parent facility. Type 4 facilities can be accessed by scanning the data hierarchy; that is, it is possible to look for all subfacilities associated with a particular facility, and process them en masse. Since Type 4 facilities also have pictures associated with them, it is also possible to access a particular Type 4 facility by pointing to its picture on the graphics screen or map.

Type 4 facilities (subfacilities) may also have Type 4 or Type 5 facilities associated with them lower on the data hierarchy. This ability to nest subfacilities and repeating data groups allows a considerable volume of information to be associated

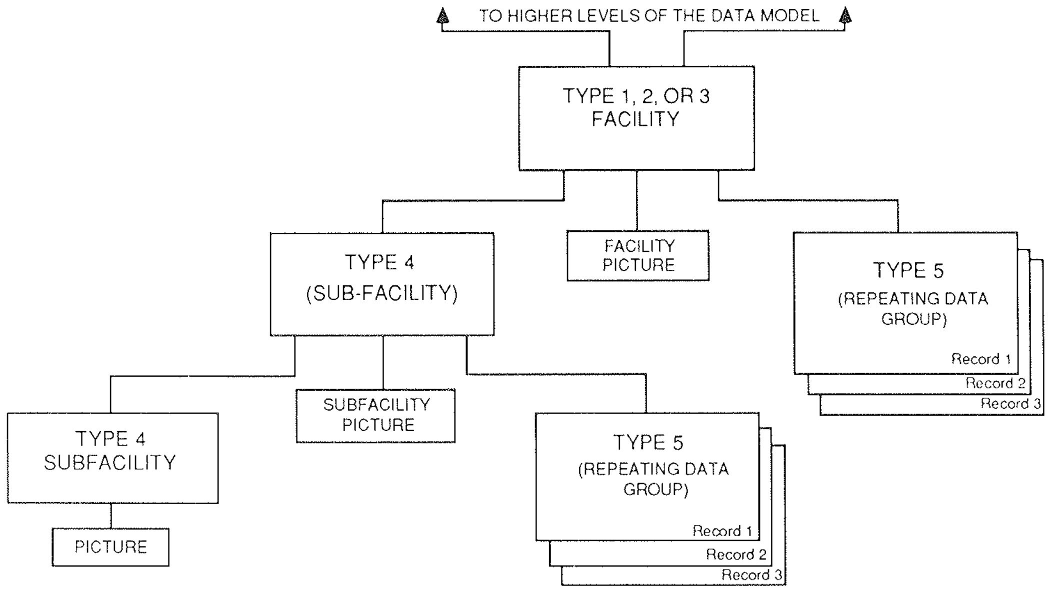

FIGURE 4 Nesting of information in the data model.

with a given facility in the work space. The concept of nesting in the data structure is illustrated in Figure 4.

In utilities applications (DFIS), several examples of the use of Type 4 facilities are found. Perhaps the best is the modeling of a transmission pole guy wire as a subfacility of the transmission pole. If the pole (Type 1) is deleted from the work space, it is not desirable to have a guy wire facility in the work space with "nothing to guy." This problem is alleviated by modeling the guy wire as a Type 4 so that it is deleted from the work space if the pole is deleted.

In the TFIS, the following entities have been modeled as Type 4 facilities: manual patching, shoulder cutting, pipe replacement, and surface treatment. All routine maintenance activities are modeled as a Type 4 facility as the maintenance treatment is a subfacility of the highway section (Type 2) or the point facility (Type 1).

Type 5 Facilities

Type 5 facilities are referred to as repeating data groups. They are subordinate in the data structure to Types 1, 2, and 3 and can also be subordinate to Type 4 subfacilities. Repeating data groups are especially useful in situations where periodic records must be kept on a particular facility or subfacility. Such items as the periodic inspection of transmission poles for signs of deterioration are likely candidates for modeling as Type 5 in utilities applications. A repeating data group record may be repeated any number of times for a particular facility. The Type 5 facilities have no pictures associated with them, as do the Type 4 facilities.

With reference to a roadway management system, the Type 5 facilites are where the majority of data elements are stored. For example, geometry is a repeating data group. The group contains information on the geometry and alignment of a highway section. The following data fields are part of geometry: length, grade, shoulder type, shoulder width, median type, median width, control areas, degree of curves, and superelevation. Each Type 5 facility can have almost unlimited data fields. The other Type 5 facilities developed in this project include: traffic, pavement design, safety, pavement monitoring, and programmed maintenance.

DEALING WITH ROADWAY-MANAGEMENT SECTIONS IN TFIS

After the framework for the TFIS was developed and the categories defined for the facilities, a case study was performed to determine the feasibility of the system. In this scenario, the code developed was tested on a highway network involving part of Centre County, Pennsylvania. Specifically, the sample application included the area covered by the State College and Julian, Pennsylvania, 7.5-minute U.S. Geological Survey (USGS) topographic quadrangle maps (see Figure 5). It was expected that, by using an existing highway network structure, deficiencies in the developed software control files would be more quickly uncovered than if the code were tested on hypothetical networks.

These expectations have indeed been borne out, and many problems or potential problems have been discovered in the data-modeling and programming conventions. Although not all of these problems have been solved, their discovery has at least given the project team a chance to document them, so that a user implementing TFIS will not make the same mistakes or need to duplicate the efforts of the project team.

One critical data-modeling aspect that has come to light in the Centre County sample application is the need to be able to logically unify a series of digitized highway facilities into a single, homogeneous entity known as a roadway management section (RMS). An RMS is defined by most highway agencies as a section of highway having a defined beginning and ending point and homogeneous properties throughout its length of 500 to 2,500 ft. These properties may include pavement type, pavement roughness and degree of deterioration, geometry, amount and distribution of traffic, number of accidents, and numerous other items. The end points of the RMS are defined in terms of *X*, *Y* state plane coordinates.

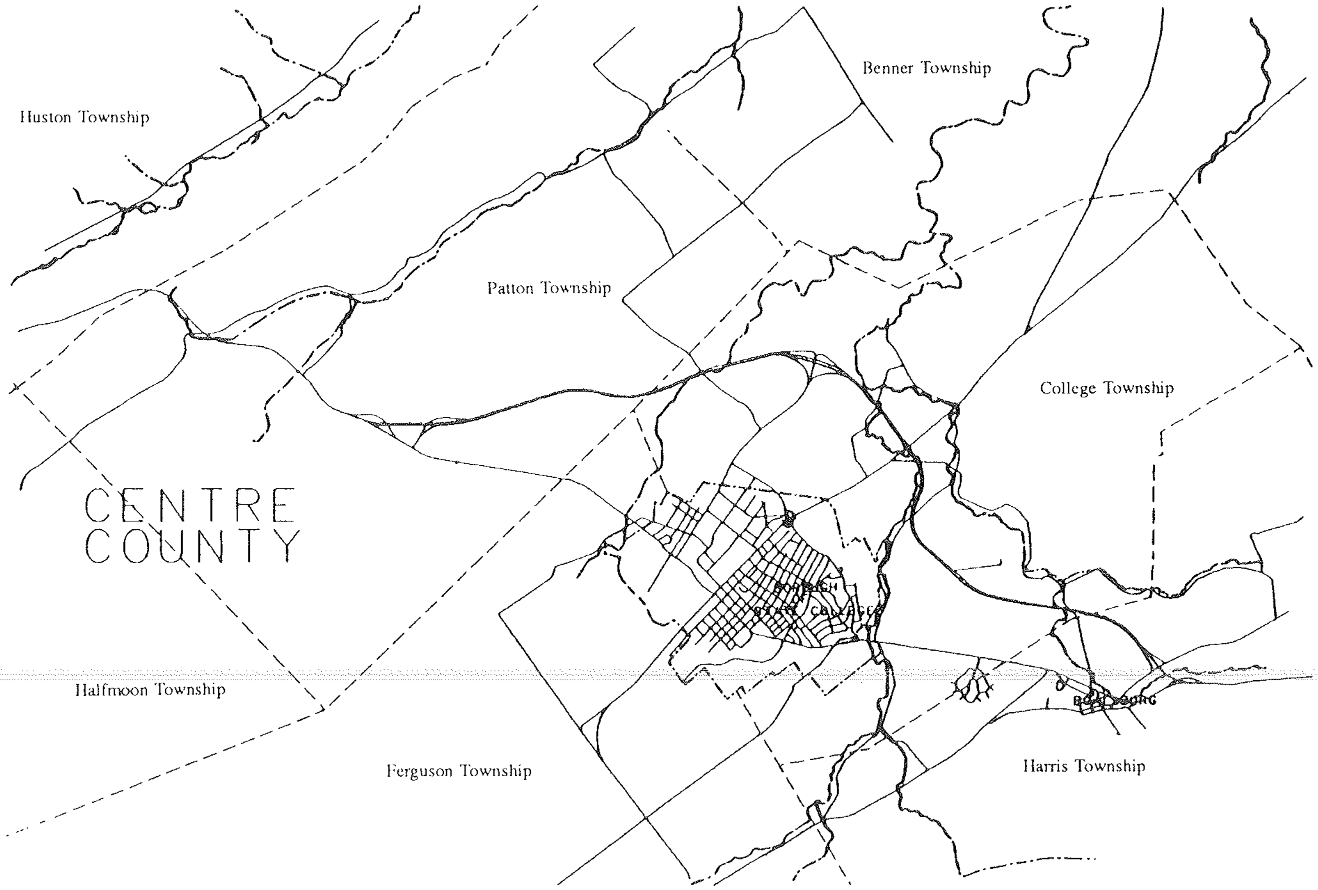

FIGURE 5 Centre County TFIS sample application areas.

Digitized Highway Sections Versus Roadway-Management Sections

Information on roadway-management sections in the Centre County application was derived from a number of sources provided by the Pennsylvania Department of Transportation (PennDOT). These included STAMPP microcomputer data files, county maps showing legislative route (LR) numbers, and straight-line diagrams depicting stationing of intersections and various other objects and points of interest.

Using these available information sources, the RMSs were scaled on the map from locations of known station (such as an intersection or stream crossing), and their ends marked and noted on the map along with such information as the beginning and ending station of the RMS. Stations and physical (map) locations of items such as bridges and culverts were also noted on the map documents (Figure 6). Having been digitized into the work space, the facilities were then edited using GPG's full-screen editor. Descriptive data derived from the STAMPP data files and the straight-line diagrams were added to the attribute fields in this process. An example of the attribute editing screen is shown in Figure 7.

At this point in the digitizing process, each Type 2 span facility digitized corresponds to a roadway-management section. In other words, there are no breaks in the facility between the end points of the RMS, and each is continuous. An example of a roadway management section is shown in Figure 8. The RMS is the section of South Atherton Street stretching between the College Avenue and Hamilton Avenue intersections indicated by shading. It begins at station 0 + 00 and ends at 26 + 74 and is 2,674 ft long.

Difficulty with the data model begins to arise, though, when side streets and highways that connect to the RMS are digitized into the work space. The GPG pointing rules used to add highway sections are written so that if a highway facility already exists at the point where the new facility is to be added, the software splits the existing facility and adds a new absolute *X*, *Y* point at that spot. The reason for this is to maintain the connectivity of the network, so that if a highway network trace is desired at some time in the future, there will be connectivity in the data base between the RMS and the side street, as there is in reality. As soon as these splits begin to occur, the RMS becomes partitioned and fails to be the single, continuous entity it was when originally digitized.

Thus, when side streets connecting South Atherton Street were digitized, in our example splits occurred at each new intersection that was created. This means that splits occurred at the intersections with Beaver, Foster, Nittany, Fairmount, and Prospect Avenues, and the original single RMS section was therefore partitioned into six separate and distinct highway sections. The only thing that bonds them together is commonality of the attribute data fields; when the original RMS was split up, all of the subsections were given data fields that were carbon copies of the original. As the situation stands at this point, each of the six highway sections that make up the example RMS can have its attribute data edited independently. This means that even though it may not be intentional on the part of the user, different facilities making up the RMS may have different attributes, a situation that violates the assumption of homogeneity of the RMS.

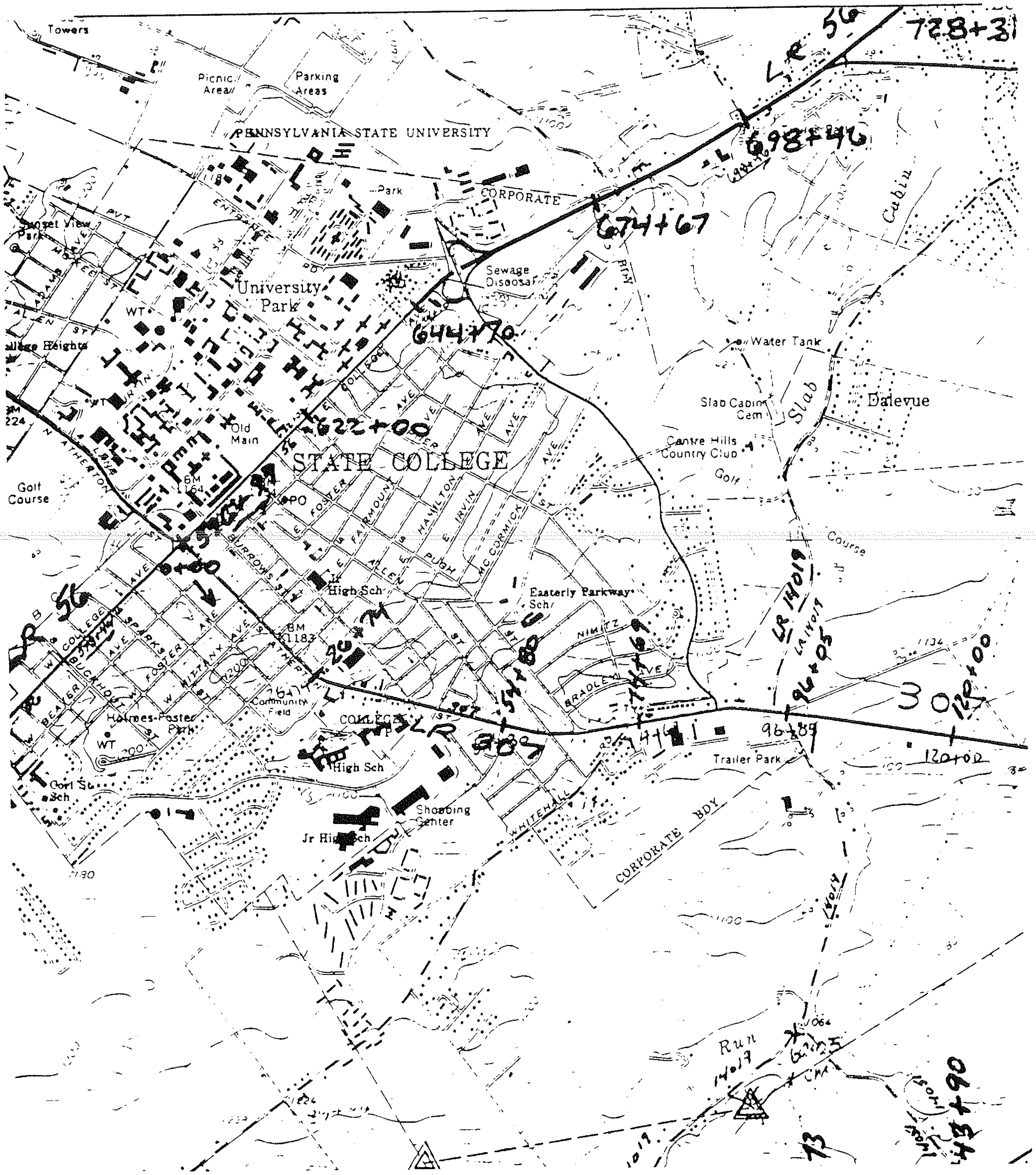

FIGURE 6 Section of map document with scaled RMSs and annotation.

From the foregoing discussion, it is apparent that some thought must be given by an agency implementing TFIS to the method of modeling a roadway-management section. Some method needs to be devised by which a series of highway facilities in a TFIS work space can be logically glued together into an RMS.

It was seen fairly early in the research project that perhaps the glue might best manifest itself in the form of an RMS identifier code that would be included as part of the attribute data of each highway facility. For this reason, the facility definition for highways, which defines all of the attribute information associated with a highway, contains the field "ID", an 8-byte character data field.

This ID field was defined for character data in case the organization implementing TFIS uses some form other than an integer numbering system to define section identifiers. The use of character data allows some imagination to be used in designing RMS identifier codes. In some sections of the Centre

```
EDIT HIGHWAY     210   TYPE: 2   LAYER: E KEY: SYS      11    FIELD 1    MAX 24
 PT1:    1968436      227357       PC1    1
 PT2:    1968855      226935       PC2:   1                CASE  UPPER
----------------------FIELD DATA-------------------------------------
 ID      NAME      SUB  TYPE        CHG?        VALUE
 1     USERID           CHAR      E  -     BRENDA
 2     DATE             CHAR      E  -     05/14/86
 3     TIME             CHAR         -     13:07:27
 4     NAME             CHAR         -     SOUTH ATHERTON ST.
 5     WIDTH            REAL         -     24.0
 6     ID               CHAR         -     001LR307
 7     START            REAL         -     0.0
 8     END              REAL         -     26.74
 9     LENGTH           REAL         -     2674.0
 10    DIRN             CHAR         -     EAST
 11    JURSDCTN         ALIAS        -     STATE
 12    NBRLANES         HALF         -     3
 13    PVMTTYPE         CHAR         -     CONCRETE
 14    CURPSI           REAL         -     4.1
 15    LASTMAIN         CHAR         -       /  /
ENTER COMMAND OR PFKEY (PF1 OR HELP FOR INFORMATION)
====> COPY 1
```

FIGURE 7 Example of data fields for a highway section.

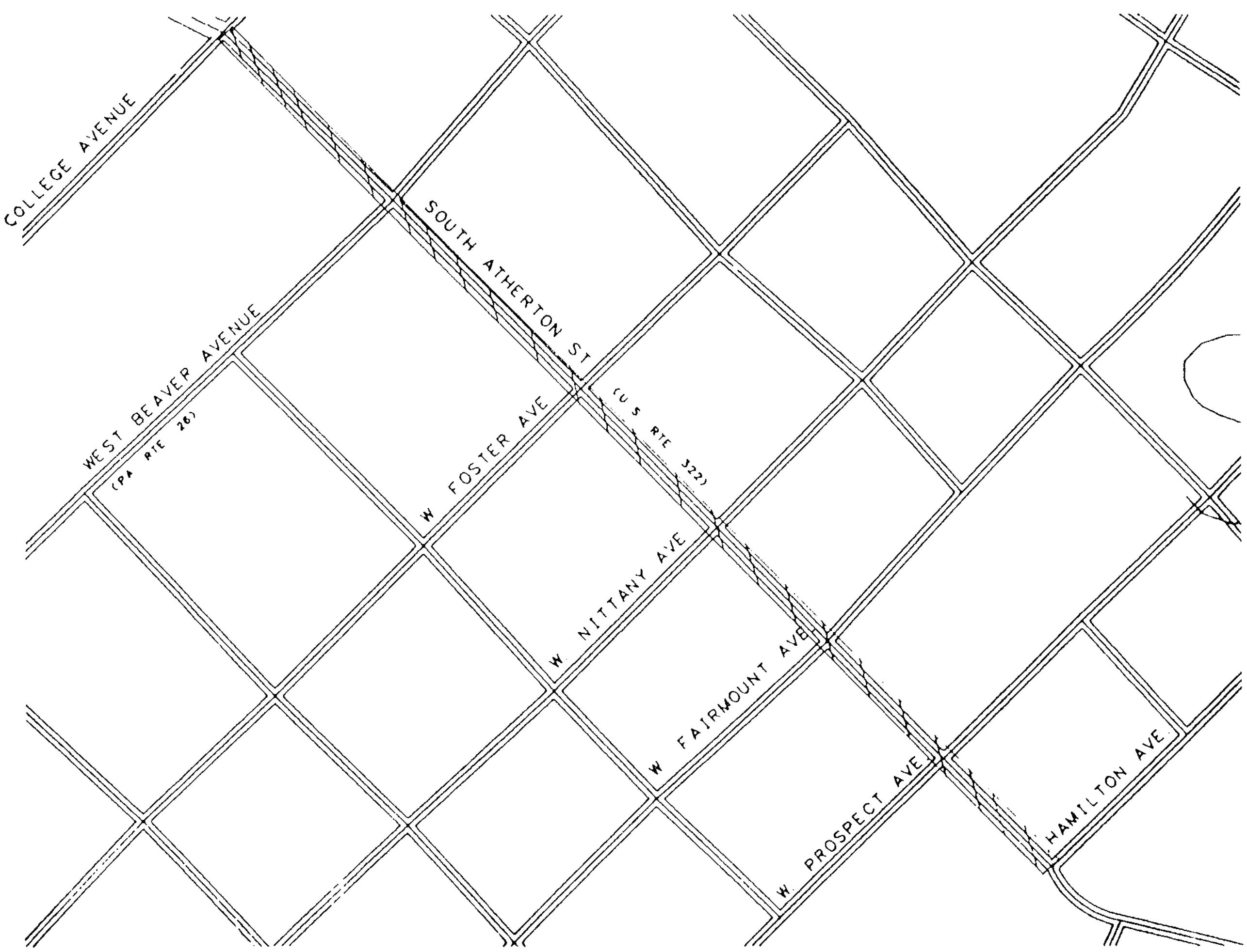

FIGURE 8 Roadway management section after digitizing process.

County sample application, for instance, RMS were given codes such as '001LR307', '002LR307', and so on. Thus, information can be captured in the code to indicate the LR number and the section number within that LR. In the foregoing example, the codes would designate the first and second roadway-management sections within LR-307, which is South Atherton Street (US-322).

The software control files written to deal with roadway management sections use the fact that all highway facilities making up the RMS will necessarily have the same ID code. When a facility is found whose ID matches the one sought, then that facility is flagged, generally by setting one of the bits in the 32-bit status word assigned to each facility.

Once this flagging process has taken place, the facilities composing the RMS are defined, and subsequent processing is generally in the form of a mass update of the attribute data associated with facilities in the RMS. However, the user may merely want a graphic depiction of the extent of a particular RMS. In this case all those sections are made to blink on the graphics screen (although other graphic report options exist).

OTHER CONSIDERATIONS FOR DATA MODELING

The Argument for Digitizing Largest to Smallest

Considerable time and effort can be saved during the digitizing process by giving some thought to the order in which highway facilities are digitized. The Centre County sample application showed that it was most efficient to digitize from the top down. That is, important arterials and Interstate highways were digitized first, then less important LRs for which STAMPP data were available, and last, collector routes and streets.

This digitizing hierarchy stems from the idea of partitioning roadway management sections. A major objective in a data conversion exercise such as digitizing a map is to convert the most data with the least time and effort and to make digitized sections as long as possible. For example, when a 2,500-ft-long roadway-management section is digitized as a single facility, its attribute data need only be edited once in order to transfer the information from a map or the STAMPP data file into the TFIS work space. If, on the other hand, roads of lesser magnitude that intersect the RMS are digitized first, then the same RMS will have to be digitized as a series of facilities and individually edited.

Instability of Map Documents

It has been known for many years that paper map documents are subject to expansion and contraction as a result of changes in temperature and humidity. For most map applications, the degree of change in the dimensions of the document is not noticeable or detrimental. When maps are digitized into a computer data base, however, even small deviations in the scale of the map are quite noticeable given the exceptional precision of modern electromagnetic digitizing equipment.

This fact became readily apparent during the course of the TFIS sample application. Part of the highway network was digitized into the work space each day. With weather and humidity changes from day to day and throughout the day, the paper map documents shrank and swelled, making it difficult to correlate on the parts of the network already digitized into the work space. At one point, correlation on an already-digitized facility was missed by nearly 100 data base units (feet, in this case) when part of the work space was digitized on one day, and then the map was reregistered with the coordinate system on the following day. It is important to note that a change in any dimension of the map of only 0.025 in. is sufficient to cause an error of 50 ft in the work space at this scale (1:24,000).

An attempt to alleviate this problem was made by laminating the map sheets between layers of polyethylene plastic. The idea was that by bonding the paper between layers of a material with different thermophysical properties, the shrinkage and swelling problems would be avoided. While this approach did indeed reduce their severity, it did not totally alleviate them. It is recommended that an organization implementing TFIS copy its map documents onto some stable base material, such as Mylar drafting film, using a photographic process before digitizing from them. Another approach would be to use a larger-scale map so that the size of the errors caused by shrinking and swelling would be smaller in relation to the data base unit.

Precision of Registration Coordinates

Part of the problem in obtaining proper correlation on a paper map document was also found to be the manner in which the state plane coordinates for the registration points on the map corners were originally obtained. Initially, the coordinates for the map corners were scaled from the 10,000-ft grid ticks that the USGS supplied on the map sheets.

It was later realized that some of the registration error could be alleviated by mathematically calculating the coordinates from the latitude and longitude at the corners of the sheet, which have even numerical values. As these are exact values, precise values of the state plane coordinates for the corners can be computed. These values can then be rounded to the nearest foot and input to GPG using the X, Y map registration process. The computation was facilitated by writing an interactive FORTRAN program that allows input of latitude and longitude and computes values for the state plane coordinates. This program makes use of a data file containing 11 parameters for the Lambert polyconic projection for the particular zone (in this case, the Pennsylvania North Zone).

Time Component of Digitizing Maps

Graphic data input into TFIS in the sample application was by hand-digitizing of maps through the use of an electronic digitizing tablet and cross-hair cursor. Attribute data entry was carried out via keyboard, and required individual editing of hand-digitized transportation facility sections. This proved to be an extremely time-consuming and labor-intensive process. The sample application area, consisting of two USGS 7.5-min topographic quadrangles, took a total of approximately 200 man-hours to be digitized, have attribute data transferred from other sources, and be checked for errors. If it is assumed that the typical map sheet will require an average of 100 man-hours to be entered into the work space (graphic data, attributes, and checking for errors), and if a man-hour costs $10 to the implementing organization, then total cost in labor alone to develop a

transportation data base for a state the size of Pennsylvania will be on the order of $1 million.

As most organizations concerned with the management of transportation facilities information already have some type of machine-readable data files, it would probably be more cost effective to put the data files directly into the TFIS data base. The conversion program itself would consist largely of input and output operations: reading the data records from the in-house file and reformatting the data under the conventions of Interface Format. It is quite probable that certain clean-up operations, particularly on the graphics, would need to be performed once the data had been transferred to the GPG work space, but it is expected that the total savings in data conversion time and effort would far outweigh the additional burden of these clean-up operations.

FINDINGS

The modeling of a roadway-management system can be based on previous work done in the utility area. Facilities with attributes distributed over a geographic area are similar in concept. Electric transmission lines and roadway segments can be similarly modeled. A TFIS model can be made up of several types of facilities. Type 1 represents point facilities, whereas Type 2 represents span facilities. The Type 4 and 5 facilities represent the data elements associated with the roadway network.

The application of TFIS to an existing highway and bridge network made apparent some of the problems that may occur when an organization responsible for the maintenance of information on transportation facilities implements TFIS in a production environment. Specifically, these problems include the instability of paper map documents, the precision of coordinates used when registering map documents, and the time and cost involved in digitizing maps.

Based on the problems encountered in the Centre County application of TFIS, which were previously discussed, the following recommendations are made to any organization implementing a TFIS.

1. Map documents to be used in the digitizing process should be copied onto a stable medium that is not as subject to the problems of expansion and contraction as paper.
2. Coordinates of points used for the registration of the map document on the digitizing tablet should be determined with a precision of not less than the data base unit (typically 1 ft). In most cases, this will preclude the method of scaling coordinates from the map document itself. Coordinates of sufficient precision may be computed mathematically from true spherical coordinates noted on the map document (such as latitude and longitude values noted on USGS maps).
3. The possibility of converting machine-readable data files already existing in the organization into a format usable with the TFIS software should be considered. It is quite possible that the programming effort involved in the conversation will be far more cost effective than hand-digitizing of map documents and the subsequent manual entry of attribute data. Moreover, this method offers less chance for human error to cause problems with data integrity.

The research team has determined that if sufficient forethought and planning are allocated to these issues, TFIS can indeed be successfully applied to manage information on geographically dispersed transportation facilities. Further research and development efforts on this highly flexible system are likely to yield a system that is more efficient, cost effective, and easy to use, and that will in the long run greatly reduce the information-management costs of organizations charged with responsibility for transportation facilities.

ACKNOWLEDGMENT

The authors would like to express their appreciation to IBM Corporation, which provided funding for the project. Appreciation is also extended to the Pennsylvania Department of Transportation, which provided highway information data from their STAMPP program.

REFERENCES

1. W. R. Hudson, R. Haas, and R. D. Pedigo. Pavement Management System Development. *NCHRP Report 215*. TRB, National Research Council, Washington, D.C., 1979.
2. *A Program of Study in Pavement Management*. Center for Transportation Research, University of Texas at Austin, Aug. 1984.
3. *Distribution Facilities Information System Overview*. IBM Publication No. GE20-0636, IBM Corporation, White Plains, N.Y., Aug. 1980.
4. W. P. Kilareski, F. L. Mannering, D. R. Luhr, and S. A. Kutz. Transportation Facilities Information System for Pavement Management. In *Transportation Research Record 1048*, TRB, National Research Council, Washington, D.C., 1985.
5. *GIS Concepts Class Notes*. IBM Corporation, Poughkeepsie, N.Y., 1985.

Publication of this paper sponsored by Committee on Pavement Management Systems.