TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-83/+3-329-1F		

4. Title and Subtitle	5. Report Date
Analysis of a Transportation Network Data Base	August, 1982
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Barbara A. Hilger	Research Report 329-1F
9. Performing Organization Name and Address	10. Work Unit No.
Texas Transportation Institute The Texas A&M University System College Station, Texas 77843	11. Contract or Grant No. 2-18-82-329
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas State Department of Highways and Public Transportation Transportation Planning Division P.O. Box 5051 Austin, Texas 78763	September 1981 Final - August 1982
	14. Sponsoring Agency Code

15. Supplementary Notes

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration. Study Title: "Analysis and Design of the Transportation Network Data Base"

16. Abstract

The State of Texas has over 75,000 miles of highway on the state-maintained highway system. Data collected and stored for these on-system highways consist of roadway environmental and traffic data, detailed accident data, bridge inventory and appraisal data, railroad crossing inventory data, and construction and maintenance history data. Processing of this volume of data for purposes of network and project planning activities is not a simple task. Over 1,000 distinct pieces of information have been identified in previous research studies for possible inclusion in a comprehensive data base structure.

The State Department of Highways and Public Transportation (SDHPT) consists of 24 districts that oversee the daily project activities of the Department. In addition, 14 centralized divisions and various administrative sections are involved daily in the network and project planning activities necessary for the monumental task of maintaining a highway system of this size.

The integration of all pertinent information on every point or segment of the 75,000 miles of roadway into a comprehensive data file structure is a priority issue, necessary to meet the current and future requirements for information by SDHPT, cooperating research agencies, and the Federal Highway Administration (FHWA). Other states have also recognized this need and are developing just such an integrated system to provide better access to information. Few states however, have to contend with the problems associated with storing and accessing such a large base of information.

The primary purpose of this study was to determine the best approach to meet the Department's requirement for current and readily accessible data by means of an integrated data file storage technique. Based upon the approach recommended herein, a transportation network data base can be designed which will incorporate the required information.

17. Key Words	18. Distribution Statement
Data Base Systems Analysis Pavement Management Systems Transportation Network	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	34	

Form DOT F 1700.7 (8-69)

ANALYSIS OF A TRANSPORTATION NETWORK DATA BASE

by

Barbara A. Hilger

Research Report Number 329-1F

Analysis and Design of the
Transportation Network Data Base

Research Project 2-18-82-329

conducted

in cooperation with the
U. S. Department of Transportation
Federal Highway Administration

by the

Transportation Planning Division
Texas State Department of Highways and Public Transportation

Texas Transportation Institute
The Texas A&M University System

August 1982

METRIC CONVERSION FACTORS

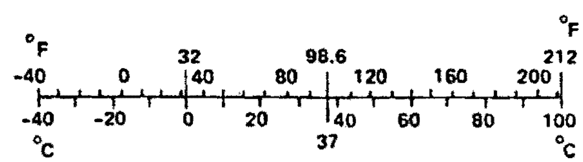
Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



ABSTRACT

The State of Texas has over 75,000 miles of highway on the state-maintained highway system. Data collected and stored for these on-system highways consist of roadway environmental and traffic data, detailed accident data, bridge inventory and appraisal data, railroad crossing inventory data, and construction and maintenance history data. Processing of this volume of data for purposes of network and project planning activities is not a simple task. Over 1,000 distinct pieces of information have been identified in previous research studies for possible inclusion in a comprehensive data base structure.

The State Department of Highways and Public Transportation (SDHPT) consists of 24 districts that oversee the daily project activities of the Department. In addition, 14 centralized divisions and various administrative sections are involved daily in the network and project planning activities necessary for the monumental task of maintaining a highway system of this size.

The integration of all pertinent information on every point or segment of the 75,000 miles of roadway into a comprehensive data file structure is a priority issue, necessary to meet the current and future requirements for information by SDHPT, cooperating research agencies, and the Federal Highway Administration (FHWA). Other states have also recognized this need and are developing just such an integrated system to provide better access to information. Few states however, have to contend with the problems associated with storing and accessing such a large base of information.

The primary purpose of this study was to determine the best approach to meet the Department's requirement for current and readily accessible data by means of an integrated data file storage technique. Based upon the approach recommended herein, a transportation network data base (TRANSNET) can be designed which will incorporate the required information.

SUMMARY

This report describes the existing automated system for roadway information in Texas, the problems inherent with a system developed over 20 years ago, and the need for access to current information for the daily planning activities of the Department. The report addresses the evolution of file access methods which are now at the stage of handling the coordination of data required and the flexibility to change as needs for information change.

IMPLEMENTATION STATEMENT

This study indicates that problems do exist in the existing automated system for storage of roadway information. Recommendations are made for additional research in the following areas:

- the feasibility of a statistical data base to handle the large volume of historical data, particularly in relation to techniques for projecting missing data.
- continuation of previous research performed to determine what variables should be retained, for what time period, and in what physical file structure.
- design of a file storage mechanism utilizing relational data base software techniques that will provide the flexibility to respond to changing needs for information.

DISCLAIMER

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

The State of Texas has over 75,000 miles of highway on the state-maintained highway system. Data collected and stored for these on-system highways consist of roadway environmental and traffic data, detailed accident data, bridge inventory and appraisal data, railroad crossing inventory data, and construction and maintenance history data. Processing of this volume of data for purposes of network and project planning activities is not a simple task. Over 1,000 distinct pieces of information have been identified in previous research studies for possible inclusion in a comprehensive data base structure.

The State Department of Highways and Public Transportation (SDHPT) consists of 24 districts that oversee the daily project activities of the Department. In addition, 14 centralized divisions and various administrative sections are involved daily in the network and project planning activities necessary for the monumental task of maintaining a highway system of this size.

The integration of all pertinent information on every point or segment of the 75,000 miles of roadway into a comprehensive data file structure is a priority issue, necessary to meet the current and future requirements for information by SDHPT, cooperating research agencies, and the Federal Highway Administration (FHWA). Other states have also recognized this need and are developing just such an integrated system to provide better access to information. Few states however, have to contend with the problems associated with storing and accessing such a large base of information.

A specific requirement for an integrated data system in Texas is related to current efforts to develop a comprehensive pavement management system. As outlined by Epps, et al. (1, p.2.2) one of the improvements which should be made for pavement management is

> ". . . the development of a comprehensive data bank of information which could be used to more objectively identify present and past history of design, construction, maintenance, and performance. At the present time most highway agencies lack a central, readily accessible data bank."

BACKGROUND

In 1937 SDHPT began maintaining information about the highway system using control/section and milepoints to identify each location. The Roadway Information System (RIS) was developed in 1964 to automate the roadway data which were previously maintained manually. The RIS evolved as a set of individual files for specific information as each application was identified. Primarily the applications were based on the need for reporting certain information to meet federal requirements. For this reason, each file was initially designed as a historical record of activities that had occurred during each year.

In subsequent years it became apparent that this stored information could provide valuable input to the administration, planning, design, construction, maintenance, and safety activities of the Department. This change in perspective, however, requires that the data must be stored in a manner that allows access in a much more timely fashion than has been possible in the past. Engineering and management decisions must be made based on current conditions and valid, comprehensive information. Data that are several months, and in some cases years, old will no longer suffice.

OBJECTIVE

The primary purpose of this study was to determine the best approach to meet the Department's requirement for current and readily accessible data by means of an integrated data file storage technique. Based upon the approach recommended herein, a transportation network data base (TRANSNET) can be designed which will incorporate the required information.

CHAPTER II

ROADWAY INFORMATION SYSTEM

The Roadway Information System (RIS) is composed of eight individual files of data and a mechanism for retrieving information from each of the files (2). The eight files are:

- Accident file
- Bridge Log file
- Railroad Grade Crossing file
- Road Life file
- Road Inventory and Traffic Log file
- Bridge Inventory and Inspection file
- Milepoint Equivalency file
- Milepoint/Milepost Equivalency file

Each of these files depicts information pertaining to the state-maintained highway system. Individual records on each of the files are keyed to a milepoint, or a range of homogenous milepoints, within a control/section. The specific beginning and ending milepoints of each range are coded to thousandths of a mile and ranges may vary from file to file.

The file structure of the RIS was initially designed in the late 1960's utilizing the type of software prevalant at the time. While random access of individual records is possible, most processing is done in a sequential manner. The files were generally designed and optimized for one application; that is, supplying historical data to meet federal reporting requirements.

RIS FILES

The Accident file contains information describing accidents that occurred during the year on the state-maintained highway system. Each record consists of 66 pieces of information describing an individual accident. These records are formatted according to the Texas Department of Public Safety coding instructions (3), with control/section and milepoint added to identify each location. On-system accident information from prior years, as well as any off-system accident information, is not readily accessible. Over half of the 458,017 accidents which occurred in Texas during 1981 were not on the state-maintained highway system (4, p. 2) Access to this information would be beneficial (e.g., over 60 percent of the Bridge Division activities is on off-system roadways).

The Bridge Log file contains records describing each bridge located on the state-maintained highway system, and on the federal-aid metropolitan and topics systems. Records contain 75 fields of data describing each structure, and the location of each bridge is generally identified by control/section, milepoint, job number, and bridge number. Most of the fields are coded from manual bridge records when bridge construction and reconstruction projects are completed. Other fields are duplicated from the Road Inventory and Traffic Log file also in the RIS. In addition, the descriptive information available for each structure on the Bridge Log file is duplicated to update the Bridge Inventory and Inspection file in the RIS.

The Railroad Grade Crossing file contains information for each railroad company at each crossing location. Records consist of 53 fields describing each crossing on the state-maintained highway system, city streets, and county roads; and the location of each crossing is identified by control/section and milepoint.

The Road Life file consists of both historical and active information pertaining to the construction and cost of state-maintained highways. Seventy-nine fields of data are coded from the final construction plans and cost folders after a project has been completed. The Road Life file generally uses control/section and job number for location identification. For records coded after 1976, beginning and ending milepoint are also available. Codes used to describe the roadway (e.g., surface and base materials) were developed many years ago, and do not reflect newer materials now in use.

The Road Inventory and Traffic Log file contains roadway characteristics and traffic data for homogenous segments of the state-maintained highway system. Roadway descriptive data are coded primarily from the manual roadway inventory strip maps which are updated when final plans are submitted upon the completion of a project. The records are keyed by control/section, as well as, the beginning and ending milepoints for the range of milepoints described. Information pertaining to maintenance projects is not always available from the district offices for processing, and construction data may not be available on the automated file for six to eight months after the project has been completed. Codes used to describe certain aspects of the roadway (e.g., pavement type) were developed in the 1930's and do not reflect newer materials now in use.

The Milepoint file and the Milepoint/Milepost Equivalency file are non-substantive files required to maintain the location keys on the other RIS files. Only after the road inventory strip maps are updated, and the milepoints are established, can mileposts be available to field personnel for location identification. In addition, when a roadway has been moved during reconstruction, a record of the previous milepoint is not retained.

RIS LOCATION IDENTIFICATION

The Texas State Department of Highways and Public Transportation uses two primary roadway reference methods:

1. County, Control/Section, Milepoint
2. Highway System, Highway Number, Milepost

The Milepoint Method is a document-oriented reference utilizing road inventory logs to determine mileage from the beginning of a Control/Section. Control/Sections are defined for homogenous segments of roadway, with discontinuities usually resulting in a new Control/Section. This reference method has been used in Texas since 1937. In 1964 these road inventory logs were automated as part of the Roadway Information System.

The Milepost Method is a sign-oriented reference used for locating the distance from the beginning of a route while actually in the field. The Department began placing mileposts in 1956, but did not complete the placement. This effort was renewed in 1968, and most roads are now posted.

There has been considerable debate within the Department about which reference method to use, with the result that either one or both methods have been used to locate particular data. The Milepoint/Milepost Equivalency File was developed to cross-reference data files which use different location schemes. Older systems tend to use the milepoint method of location identification, while the systems designed more recently generally use the milepost method of location identification.

In referring to various methods of location identification, NCHRP Synthesis 21 (5, p. 20) states:

". . . there is really not a great deal of fundamental difference between the several most commonly used methods ... all use a distance measurement from an 'incident' to a known point. The characteristics are the same whether the calculation for true milepoint is done in the field, accomplished manually in the office from a straight-line diagram, or performed by the computer."

Both the milepoint and the milepost reference methods, when used consistently, meet the three criteria of known point, direction, and distance. Table 1 modified from the NCHRP Synthesis 21 (5, p. 13) gives an overview of the advantages and disadvantages of both methods.

Since there really is not a "fundamental difference" between the two used methods, the critical question becomes "Should duplicate reference methods be maintained?" Many of the problems identified by users of the current Roadway Information System stem from the use of multiple identification references: Combining information from different files can be difficult; data are out-dated after being processed through all the steps required in the office; accuracy of location identification is not consistent. While combining information from a variety of files can be accomplished now with cross-reference files and experienced programming personnel, the data are not structured in such a way that this can be done by the average user on a routine basis.

Advantages and Disadvantages of Location Methods

MILEPOINT	MILEPOST
ADVANTAGES	
Special signs are not needed	Because signs reflect mileage, which is familiar to most people, this method can be easily learned
	There is fairly uniform spacing, so the user does not have to proceed more than some fixed distance to find a marker
	The numerical sequence provides easy orientation
	The signs apply to all concurrent routes
DISADVANTAGES	
When construction changes require revisions to diagram or log milepoints and street maps, steps must be taken to ensure that users of the method receive the revisions	Changes in the length of route after initial placement of signs result in numbers not reflecting true milepoints
There may be instances of misspelled names, street and road names that are similiar or identical, roads and streets with no names or numbers, or roads with more than one name that require special consideration	Where there are concurrent routes, the numbers and the signs reflect mileage, for only one of the routes.
	The placement of signs can create problems for maintenance forces.

Table 1. (5)

CHAPTER III

DATA BASE CONCEPTS

The existing files and systems used by the Department have been developed over the last twenty years to meet specific needs and applications as they arose. This development has been based on data processing capabilities and file structures available at the time.

The stage 1, or elementary data files shown in Figure 1 were predominant in the early 1960's. This file structure requires that data be stored essentially the same way as it is entered (e.g., if each record is an 80-character card, then 80-character images of those cards are stored). The description of each record and file has to be a part of every program. Since the data are usually restricted to one application, the same field (e.g., AADT) may be repeated on many different files. If the field is updated, all of the various files with that field must be updated. If the description of any file changes, all computer programs accessing that file must be updated as well. Many of the first files of roadway information used this file storage technique.

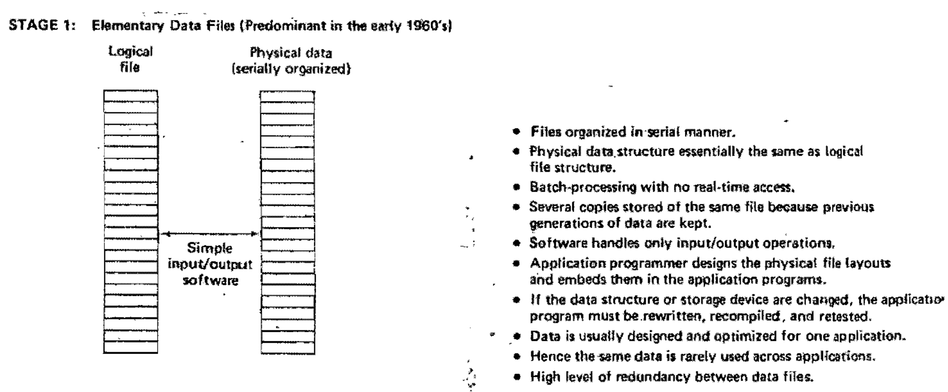


Figure 1. (6, p.24)

In the late 1960's, the Department developed the current Roadway Information System (RIS) using the Stage 2 file access method depicted in Figure 2 which was prevalent at the time. This method for file access and storage makes it possible to directly access a record without processing all records in the file. While access of each record was improved over the stage 1 sequential access, this technique does little to manage the data; data redundancy is often necessary across various application files. The merged Accident and Roadway data file developed by the University of North Carolina is another example of this access method (7).

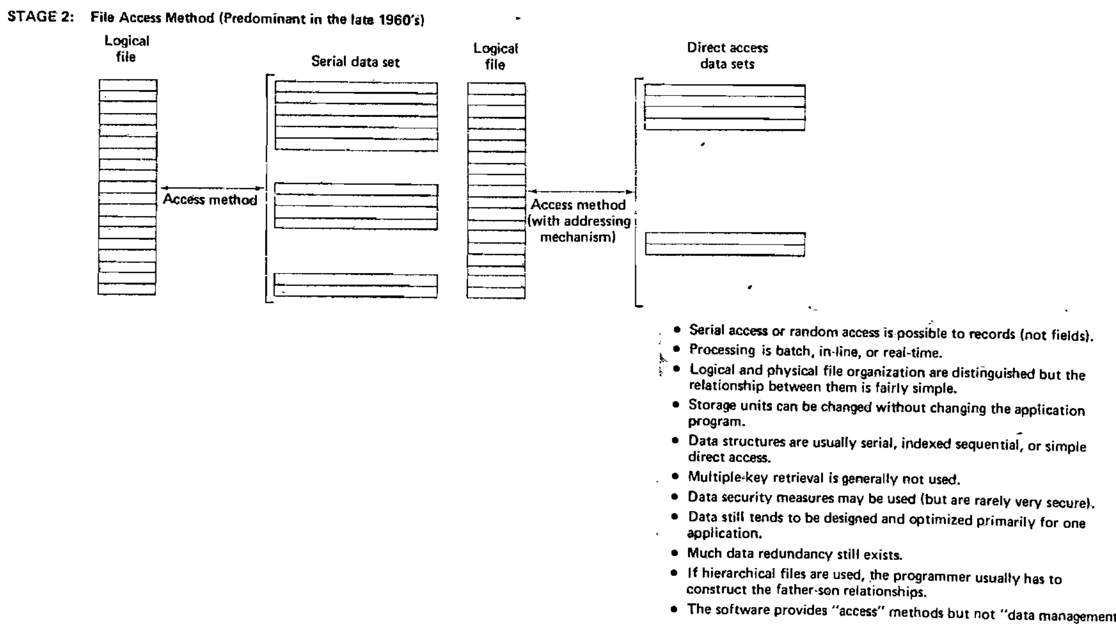


Figure 2. (6, p.24)

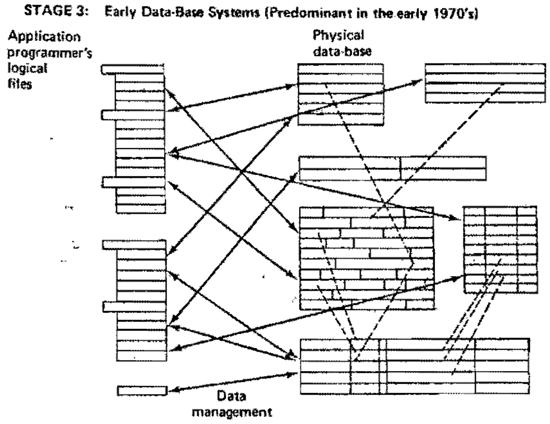
During the 1960's the term "data base" began to be used to replace "files of data" and "data sets". As often happens with a new term,

> ". . . many users promoted their files by changing their title to data base, without changing their nature to include nonre-

dundancy, data independence, interconnectedness, security protection, or in many cases, real-time accessibility." (6, p.22)

In the early 1970's, however, true data base systems became prevalent and better data management was possible. With this stage 3 software, depicted in Figure 3, data redundancy can be reduced or eliminated, and diverse applications can use the same information (e.g., if AADT changes, the field can be updated one time, and all programs which use that information will use the new value).

STAGE 3: Early Data-Base Systems (Predominant in the early 1970's)



- Multiple different logical files can be derived from the same physical data.
- The same data can be accessed in different ways by applications with different requirements.
- Software provides the means to lessen data redundancy.
- Data elements are shared between diverse applications.
- Absence of redundancy facilitates data integrity.
- The physical storage organization is independent of the application programs. It may be changed often in order to improve the data-base performance without application program modification.
- The data is addressable at the field or group level.
- Multiple-key retrieval can be used.
- Complex forms of data organization are used without complicating the application programs.

Figure 3. (6, p.25)

Many of the more advanced automated systems for pavement management activities use this stage 3 data base software. Arizona DOT designed a pavement management system using IMS data base software (8), which is a hierarchical file structure accessed by Arizona DOT users interactively. Figure 4 shows an example of the hierarchical file structure used by the U. S. Army Construction Engineering Laboratory for the PAVER system (9, p.142). This particular example of the stage 3 file access method uses System 2000 software.

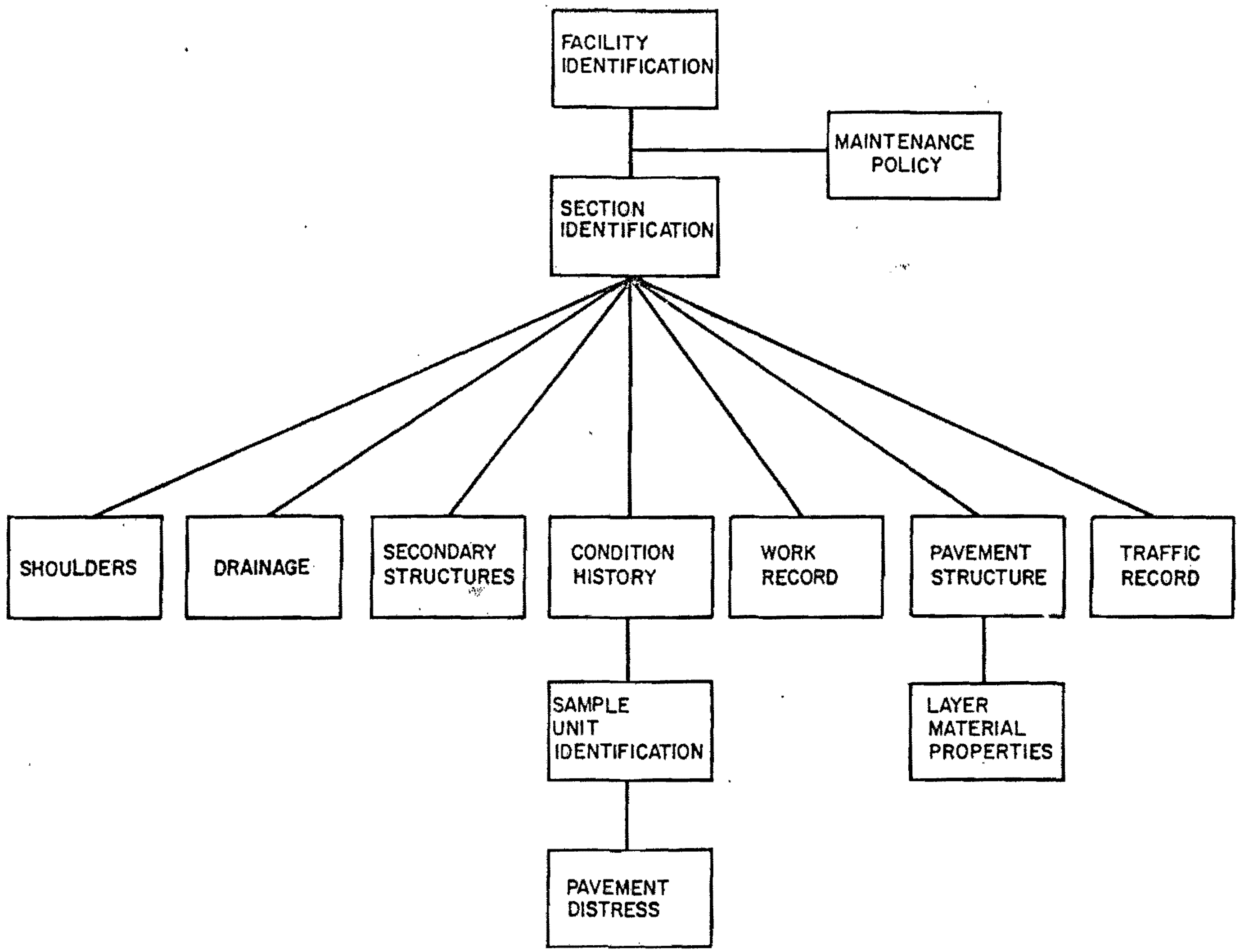


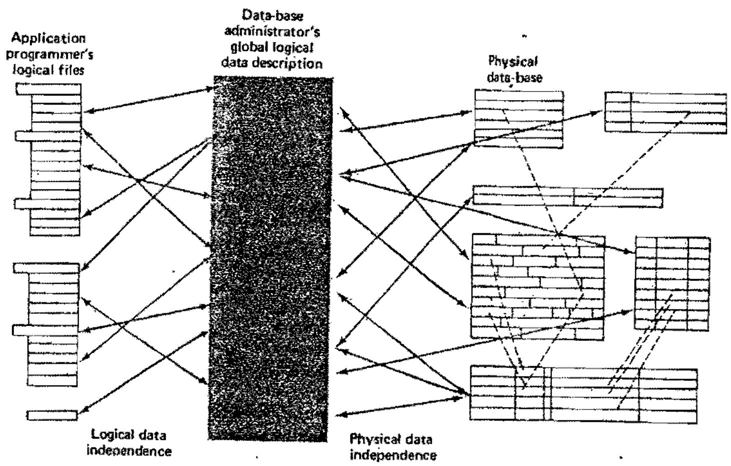
Figure 4. PAVER Data Base File Structure (9, p.142)

Use of hierarchical data base file structures has been a giant step forward for making information available to a variety of users and applications by allowing the user to access individual fields by name without any concern as to where the data were physically stored. However, not all applications are conducive to a top-down hierarchical approach. Using the PAVER data base file structure in figure 4 as an example, if a certain application wanted to report pavement distress on shoulders, the processing would have to proceed from the pavement distress record up to sample unit identification, condition history, and section identification, then down to shoulders and back again. Such searches for data consume vast amounts of computer time and in some complex hierarchies can be impossible to carry out.

The Stage 4 data base access method which came into prominence in the late 1970's, evolved to correct the awkward data relationships imposed on the user with the hierarchical structure. The Stage 4 structure is depicted in Figure 5. Network data base management systems allow the user greater flexibility in deciding how various data items will relate to each other. Changes in these relationships can be made fairly easily as the needs of the users change. The pointers used to establish these relationships require additional storage, however; and thus, increased flexibility is offset by an increase in the use of the computer resources.

The second type of stage 4 access methods is the relational data base management system. The Adaptable Data Base (ADABAS) software used by the Department for its Management Information Systems, is an example of the relational type of data base. In very simplified terms, a relational data base is a way to organize data items in tables made up of rows and columns. Access procedures based on relational calculus or relational algebra are used to map the various combinations of data items the user decides to relate to one another.

STAGE 4: Today's Requirement in Data-Base Systems



- Software provides logical as well as physical data independence, allowing a global logical view of the data to exist independently of certain changes in the application programs' views of data or the physical data layouts. The data independence capabilities listed in Fig. 3.2 are provided.
- The data base may evolve without incurring high maintenance costs.
- Facilities are provided for a Data-Base Administrator to act as controller and custodian of the data, and ensure that its organization is the best for the users as a whole.
- Effective procedures are provided for controlling the privacy, security, and integrity of the data.
- Inverted files are used on some systems to permit rapid data base searching.
- Data bases are designed to provide answers to unanticipated forms of information request.
- Data migration is facilitated.
- The software provides a *data description language* for the Data-Base Administrator, a *command language* for the application programmer, and sometimes a *data interrogation language* for the user.

Figure 5. (6, p.25)

One reason that data storage and access capabilities have evolved from sequential files is to allow greater flexibility and timeliness in responding to the changing needs of the users. Use of systems, whether manual or automated, is rarely static but, rather, evolutionary over time.

> "A data base in an organization is no more a static entity than are the contents of an organization's filing cabinets. The details of data stored, and the way it is stored, change continuously." (6, p.26)

While allowing flexibility for future changes, the implementation of any data base management system can be a very expensive and long term project. Indeed, "unless the implementation is very carefully conceived and executed, a DBMS will all too often bog down hopelessly when a certain critical level of data volume and/or complexity is achieved." (10, p.19)

CHAPTER IV

RECOMMENDATIONS FOR A TRANSPORTATION NETWORK DATA BASE

There are three major objectives to be met by a transportation network data base (TRANSNET) system:

- All pertinent information for individual points, as well as contiguous sections of roadway, must be accessible
- Data must be input in a timely fashion and easily retrieved
- Maintenance and enhancements to the system must be responsive to the growing needs for information by the Department, federal agencies, and cooperating research agencies.

Although each user has specific needs, these requirements represent the general concern of all users. Since the Department has been using the RIS for the past 20 years, there is a natural tendency to keep the existing programs and procedures that have been developed. It is more critical, however, that TRANSNET be able to interface with future needs. The desire for compatibility with the past has resulted in excessive "patching" of existing software and has pointed out more than ever the need for a file structure that can change with the Department.

The development of TRANSNET, a comprehensive data base of roadway information, using Stage 4 data base software such as ADABAS, will meet the three major objectives stated. In addition, the problems outlined in Chapter II can be alleviated. Since Mark IV software can interface with ADABAS, many of the existing programs which are still useful could be retained.

RECOMMENDATIONS FOR DATA INPUT

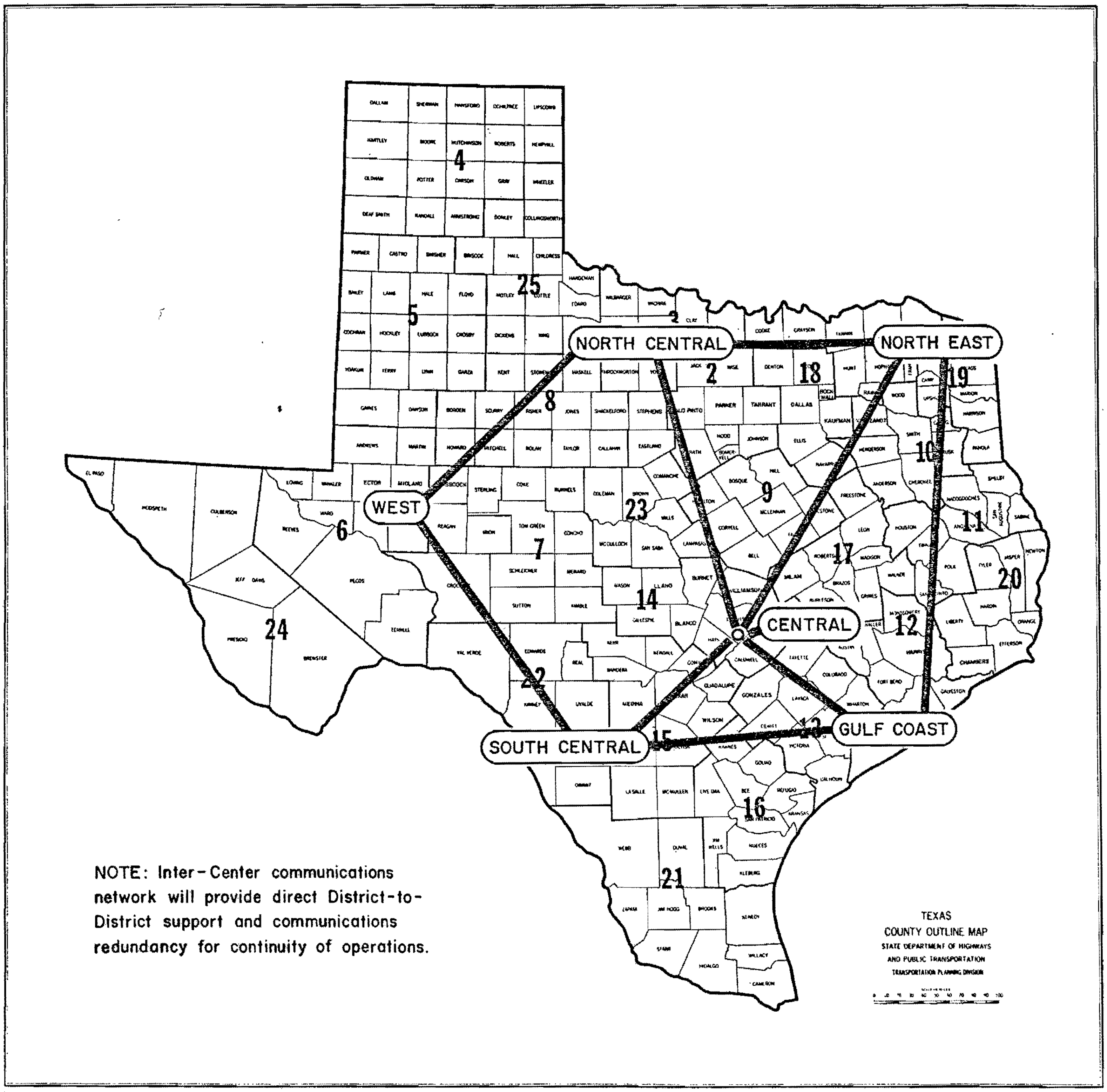
The steps taken to automate the roadway inventory data are cumbersome and allow many opportunities for errors to occur. A more appropriate method for automating this information would be to enter data from each district office. Data entry close to the source of the data collection will decrease the opportunity for random errors and will improve the timeliness of data availability by several months.

Automation of data from the district offices can be performed on a microcomputer, through the regional minicomputer, directly to the large computer mainframe, or any desired combination of the various network locations. (The Regional Network is illustrated in Figure 6). While the district offices should be responsible for their own roadway information, certain data constraints and edits can be established by the Data Base Administration (DBA) for consistency throughout the state.

During the design and construction phase of a project, preliminary data are automated for the Design and Construction Information System (DCIS). The DCIS was designed using ADABAS software (11). An automated interface between the DCIS and TRANSNET will provide skeletal data which can be expanded as the project is completed. Other existing systems should be examined to determine the feasibility of automated interfaces to provide data, thus reducing the necessity for repetitious data entry.

RECOMMENDATIONS FOR ACCESS KEY

The use of a single location reference key will simplify access to the stored information. Accuracy of data will improve when only one reference location method is used for all data, and Department personnel



SDHPT REGIONAL AUTOMATION CENTERS

Figure 6. (12)

are consistently using the same method. Improving the timeliness of the data can only be accomplished by reducing the number of steps required between the source of data collection and the source of data entry. Since the system is to consist of data related to the roadway, the logical source of input is from each district office.

If it is agreed that roadway data should be input from each district office and only one reference method is to be used, the decision as to which method should be used becomes clear. The method most effectively used in the field consists of Highway System, Highway Number, and Milepost. A recommended key to identify all locations is shown in Table 2 . There are many advantages to the Department of using this location method. The major expenses in any highway department are in the area of construction, rehabilitation, and maintenance. These project-level activities must be coordinated to reduce cost and inefficiency and must be performed at the right location. Using field reference identification will allow the quickest and most convenient way to locate job sites, and will improve the efficiency of the Department personnel who perform the most cost-intensive activities. Data for planning will be available in a more timely fashion. In addition, the data will be more accurate when fewer manual procedures are required to input the data to an automated system, and when the data are entered by personnel familiar with what is actually occurring in the field.

RECOMMENDATIONS FOR DATA ITEMS

Over 1,000 variables were identified by Department personnel using the Delphi process in research performed by Bush (13). By including

RECOMMENDED PRIMARY ACCESS KEY

POSTED ROADWAY	NUMBER OF CHARACTERS	NON-POSTED ROADWAY
FIPS County Code	3	FIPS County Code
Control/Section	6	FIPS City Code
Priority Highway System/Number	6	Primary Street Code
Beginning Milepost	6	Intersecting Street Code
Direction From Milepost	2	Direction From Intersection
Displacement From Milepost	6	Distance From Intersection

Note: Federal Information Processing (FIPS) codes are recommended for counties and cities (14).

Priority Highway System number is the highest classified highway when more than one highway is posted along the same route.

Table 2

only those variables identified by the Department users as moderately important to imperative in regard to pavement management activities, over 900 variables would be included in the TRANSNET data base for each location on 75,000 miles of roadway. While a data base of this magnitude could be built, it is not recommended. The Bush Study should be updated and carried forward to its logical conclusion. Several more iterations may be needed, as well as a study to verify the use and need (real rather than perceived) of these data. It is imperative that "Care should be taken to avoid collection of data which would be interesting as compared to data which is necessary" (1,p.2-10) for pavement management activities.

Many users expressed a desire for readily available historical data. Because of the volume of data involved, it is not recommended that historical data be maintained on-line for quick access. However, there is a way to compensate for the lack of this information. In relational data base management systems, such as ADABAS, the facility exists for relating two separate data bases, the current real-world and the "ideal-world" statistical data based on prior information. A request for information that does not exist on the current data base can be processed against data that do exist on the statistical data base as shown in figure 7.

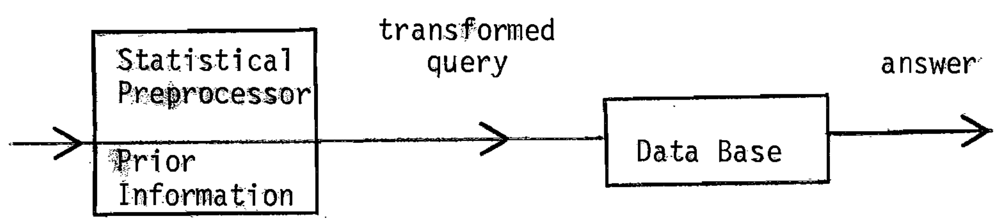


Figure 7. (15, p.475)

Not only can this method be used for obtaining historical data, but can also be used for generating values for missing data that correspond to existing data.

It is essential before creating this statistical data base that a study of pertinent variables be performed to ensure that appropriate statistical measurements are retained. The statistical data base can be periodically updated when the real-world data base values change. While it is recommended that each real-world data base be on a district level, the statistical data base should be centrally resident for access by all districts. By utilizing the concept of distributed data processing, the central processing facility can be used for summary data required for planning purposes.

REFERENCES

1. Epps, J.A., Finn, F.N., and Monismith, C.L. Pavement Management Including Recycling. ITS Extension Programs, University of California, Berkeley, March 1982.

2. State Department of Highways and Public Transportation. Roadway Information System User Manual. Transportation Planning Division and Automation Division. Revised September 1981.

3. Texas Department of Public Safety and State Department of Highways and Public Transportation. Motor Vehicle Traffic Accident Coding Instructions. January 1982.

4. Hilger, B.A. and Scott, E.S. Single Variable Tabulations of 1981 Accidents in the State of Texas (TARE 49). College Station: Texas A&M University, Texas Transportation Institute, August 1982.

5. Highway Research Board. Highway Locations Reference Methods. National Cooperative Highway Research Program, Synthesis of Highway Practice, Number 21, 1974.

6. Martin, James. Computer Data Base Organization. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1975.

7. Johnson, Ralph D. Jr., Fischer, William C., Hamilton, Elizabeth G., and Council, Forrest M. Highway Safety Improvements through Utilization of Merged Accident and Roadway Data (Volume 1). Chapel Hill: University of North Carolina, Highway Safety Research Center, August 1977.

8. Arizona Department of Transportation. Pavement Management System Data Base Design - Data Base Description Standards. Information Systems Group, Engineering Systems Section, March 1980.

9. Shahin, Mohamed Y., Darter, Michael I., and Chen, Thomas T. Development of a Pavement Maintenance Management System Volume VII: Maintenance and Repair Consequence Models and Management Information Requirements (ESL-TR-79-18). Champaign, Illinois: Construction Engineering Research Laboratory for the Engineering and Services Laboratory, Air Force Engineering and Services Center, Tyndall Air Force Base, Florida, December 1979.

10. Hamlin, Christopher L. An Introduction to Data Base Management Systems. Microsystems, May/June 1982, 16-19.

11. Texas Department of Highways and Public Transportation. Design and Construction Information System - General Design Specification. October 1977.

12. Texas Department of Highways and Public Transportation. Remote Terminal Networks and Central Computer Center Operations. November 1981.

13. Bush, Ronald R. Influence of Cognitive Style in a Methodology for Data Base Design (Research Report Number 123-27). Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration, by the Texas Highway Department, Texas Transportation Institute, and the University of Texas at Austin, February 1975.

14. General Services Administration - Office of Finance. Worldwide Geographic Location Codes. November 1976.

15. Wong, Eugene. A Statistical Approach to Incomplete Information in Database Systems. ACM Transactions on Database Systems, September 1982, 7 (3), 470-488.