COO-4830-3
Distribution Category UC-97b

DEVELOPMENT OF ADVANCED METHODS FOR PLANNING ELECTRIC ENERGY DISTRIBUTION SYSTEMS

FINAL REPORT

By

Turan Gönen

Bobbie L. Foote

John C. Thompson

School of Electrical Engineering and Computing Sciences
in cooperation with
School of Industrial Engineering
University of Oklahoma
Norman, Oklahoma 73019

October 1979

Prepared for
THE UNITED STATES DEPARTMENT OF ENERGY
Office of the Assistant Secretary for Energy Technology
Division of Electric Energy Systems
Work Performed Under Contract No. EM-78-S-02-4830

DEVELOPMENT OF ADVANCED METHODS FOR PLANNING ELECTRIC ENERGY DISTRIBUTION SYSTEMS

ABSTRACT

An extensive search has been made for the identification and collection of reports published in the open literature which describes distribution planning methods and techniques. In addition, a questionnaire has been prepared and sent to a large number of electric power utility companies. Also, a large number of these companies were visited and/or their distribution planners interviewed for the identification and description of distribution system planning methods and techniques used by these electric power utility companies and other commercial entities.

As a first step, it was necessary to determine the present scope of computer applications in distribution system planning within the power utility industry. Through the Oklahoma Gas and Electric Company and consultants on this project, many computer programs that could be used as a part of the system planning were gathered. The need for large scale analysis as well as information retrieval and display in choosing the best alternative required an interactive problem solving environment, based on a data base management system and a library of application programs, with an analytical capability that is far beyond the conventional algorithms. This interactive environment provides a necessary man-machine interface to initiate any program that exists in the Library.

However, the developed interactive design seeks to do more than just achieve a superficial compatibility among the individual programs. Rather, what is sought is truly a system of harmoniously functioning parts which assist the human planner. The system consists of three major elements: (1) a collection of planning programs with capabilities similar to those described previously, (2) a database supporting the input-output requirements of these programs, and (3) a generalized network.

The information basis for the system is the database. The database is logically organized as a relational database in which the attributes are identifiers taken from a finite set $A_1, A_2, \ldots, A_n$. Each A_i has associated with it a set of values called domain, written as $\mathrm{dom}(A_i)$. A relation on the set of attributes, $R(A_1, A_2 \ldots, A_n)$ is a subset of the Cartesian product

$$\mathrm{dom}(A_1) \times \mathrm{dom}(A_2) \times \ldots \times \mathrm{dom}(A_n).$$

An element of this subset $(a_1, a_2, \ldots a_n)$ is called a tuple. A relation may be simple visualized as a table of rows and columns. The rows represent tuples while columns represent the values of a particular attribute contained in the relation. A database consists of one or more relations. Keys are attribute values which uniquely specify tuples.

The permissible operations on relations are the operations on sets familiar from set theory plus several additional ones. Operations on the data base are performed by the Data Base Management System (DBMS). These operations are directed as the user level by a language similar to the relational algebra of Codd. The operations of importance found in the relational algebra which are not normally part of set theory are selection, projection, and join commands.

The collection of algorithms suitable for implementation on a digital computer, which maintain the network model and allow a designer to work interactively with it, are divided into three classes. An overseer called the shell, which processes the planner's demands, providing his access to a number of problem solving programs, a network editor, which allows the planner to create and modify networks using concepts and commands natural to the subject, and finally, a data base management system which maintains the data base. While the end results of the research is conceptual in nature, selected features of the Distribution System Planning Model have been implemented. This implementation provides a validity check of the notions and abstractions comprising the DSPM.

In addition, the distribution systems planning models have been reviewed and a set of new mixed-integer programming models have been developed for the optimal expansion of the distribution systems. The models help the planner to select: (1) optimum substation locations; (2) optimum substation expansions; (3) optimum substation transformer sizes; (4) optimum load transfers between substations; (5) optimum feeder routes and sizes subject to a set of specified constraints. The models permit following existing right-of-ways and avoid areas where feeders and substations cannot be constructed. The results of computer runs were analyzed for adequacy in serving projected loads within regulation limits for both normal and emergency operation.

USDOE Contract No.

EM-78-S-02-4830

OU Account No.

156-615 (2438)

DEVELOPMENT OF ADVANCED METHODS FOR PLANNING ELECTRICAL ENERGY DISTRIBUTION SYSTEMS

PROJECT STAFF

Principal Investigators: M. Earl Council
Turan Gönen

Co-Investigators: John E. Fagan
Bobbie L. Foote
John C. Thompson

Subcontractors: Fred Glover
Darwin Klingman
Analysis, Research and Computation, Co.
Bradford E. Smith
C.H. Guernsey & Co.

Special Consultants: John Freed
Lester A. Burris
Oklahoma Gas & Electric Co.

Project Technical Manager: Turan Gönen

Project Monitor: M. Earl Council

USDOE Project Manager: Tobias Trygar

ACKNOWLEDGMENT

The authors wish to thank Dr. A.A. Aly of the University of Oklahoma for his efforts at the beginning of this research and for various helpful discussions. Also, the authors are indebted to Dr. S. Lakshmivarahan and the following graduate students for their valuable assistance in this project:

P. Enouen
D.C. Yu
R.M. Cubert
S. Dasananda
N. Kebriaei
A. Özduman
F. Banaii

A special thank you is extended to Oklahoma Gas and Electric, and Public Service of Oklahoma for the continuous cooperation and encouragement. Several distribution system planning engineers, of the following companies, have consented to be available for personal interviews and/or have made available to the research team specialized materials. Their help is also gratefull acknowledged.

Oklahoma Gas and Electric Co.	Oklahoma City, OK
Public Service of Oklahoma Co.	Tulsa, OK
Kansas Gas and Electric Co.	Wichita, KS
Texas Electric Co.	Ft. Worth, TX
Texas Power and Light Co.	Dallas, TX
Dallas Power and Light Co.	Dallas, TX
C.H. Guernsey Co.	Oklahoma City, OK
Arizona Public Service Co.	Phoenix, AZ
San Diego Gas and Electric Co.	San Diego, CA
Pacific Gas and Electric Co.	San Francisco, CA
Northeast Utilities Service Co.	Hartford, CT
Georgia Power Co.	Atlanta, GA
Northern States Power Co.	Minneapolis, MN
Public Service Co. of New Mexico	Albuquerque, NM
Consolidated Edison of New York, Inc.	New York, NY
Dayton Power and Light Co.	Dayton, OH
Pacific Power and Light Co.	Portland, OR
Duquesne Light Co.	Pittsburgh, PA
Gulf States Utilities Co.	Beaumont, TX
El Paso Electric Co.	El Paso, TX
Sierra Pacific Power Co.	Reno, NV

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The overall goal of this research project was the conceptual design of an interactive distribution planning model for an electric power supply system. In order to reach this overall goal, the contractor was required to perform the following tasks:

TASK A

The contractor should assess the present state-of-the-art, in theory and practice, of electrical energy distribution system planning and analysis methods. This assessment shall include:

1. The identification and collection of reports published in the literature which describe distribution planning methods and techniques. For those techniques which are judged by the contractor to be of an advanced or unique nature and which have been coded for use of digital computer programs and their operating and data requirements.

2. To the extent possible, the identification and description of distribution planning methods and techniques in use by electric utilities and other commercial entitites.

3. The development of a classification system for distribution planning and analysis techniques and models which recognize wide range of functions performed by these techniques and models, data requirements and where digital computer codes exist their input/output requirements. Using this system, the contractor shall classify those techniques and models identified in tasks A1 and A2.

TASK B

The contractor must develop specifications for models which are required in the distribution system planning and analysis process and shall develop specifications for and a prototype of a data base sufficient to support such models. This task shall include:

1. The development of criteria for assessing the degree to which existing models meet the specifications developed;

2. The analysis of models identified in tasks A1 and A2.

3. A characterization of the conditions under which the models identified in tasks A1 and A2 are solvable.

4. The identification of distribution design techniques used in practice.

5. The identification and cataloging of planning objectives, concepts and constraints accepted in practice.

6. The selection of planning and analysis concepts and methods which form the core of efficient distribution planning methodologies.

7. The determination of the extent to which existing digital computer programs implementing the concepts and methods identified in B6 are software compatible.

8. The selection of a set of existing programs which best meet the need of the distribution planning methodologies and are sufficiently software compatible to permit use on this project.

9. The analysis of the data requirements of all phases of the distribution planning process.

10. The conceptual design of a data base sufficient to meet the data requirements of the distribution planning methodologies developed;

11. A detailed design of the data base characterized in B10.

12. The implementation of data base on a modern digital computer in a manner which maximized the transportability of the data and related software and is compatible with other software used in the planning methodologies;

13. The evaluation of the distribution planning methodologies developed by the contractor including the effects of major assumptions in the models and techniques and the amount of effort and cost required to expand the implemented software package into a production grade program.

Analysis of the literature, interviews with utility engineers and discussions by the research team have determined that the following functions must be performed by an integrated interactive distribution planning system based on deterministic data. The functional components of a complete interactive distribution planning system, as shown in Figure 1, are:

1. A distribution planning data base with an organization induced by models and calculations which use the data.
2. A data base management system to organize the data, add to it, take data out, perform quality checks, and retrieve data for operating modules.
3. A network editor module to create electrical network representations at the command of the planner from the data base or from the plans generated by the plan generator.

4. A parameter construction module which will calculate the parameters for analysis and optimization programs from the data contained in the data base.

5. An analysis module which will evaluate the planned network according to standard analyses such as reliability, voltage performance, etc. as well as subjective evaluations by the planner.

6. An efficiency evaluation module which will take analytical results from the analysis module and compute the efficiency of plans according to each objective defined in the design of the system.

7. A plan generator module. This module will either intake plans generated by the planner or create plans by means of an optimization model.

8. An evaluation model which contains a general objective function taking into account all of the objectives which will evaluate plans generated by the planner and optimization model if more than one plan is created.

9. A master program called the shell which will

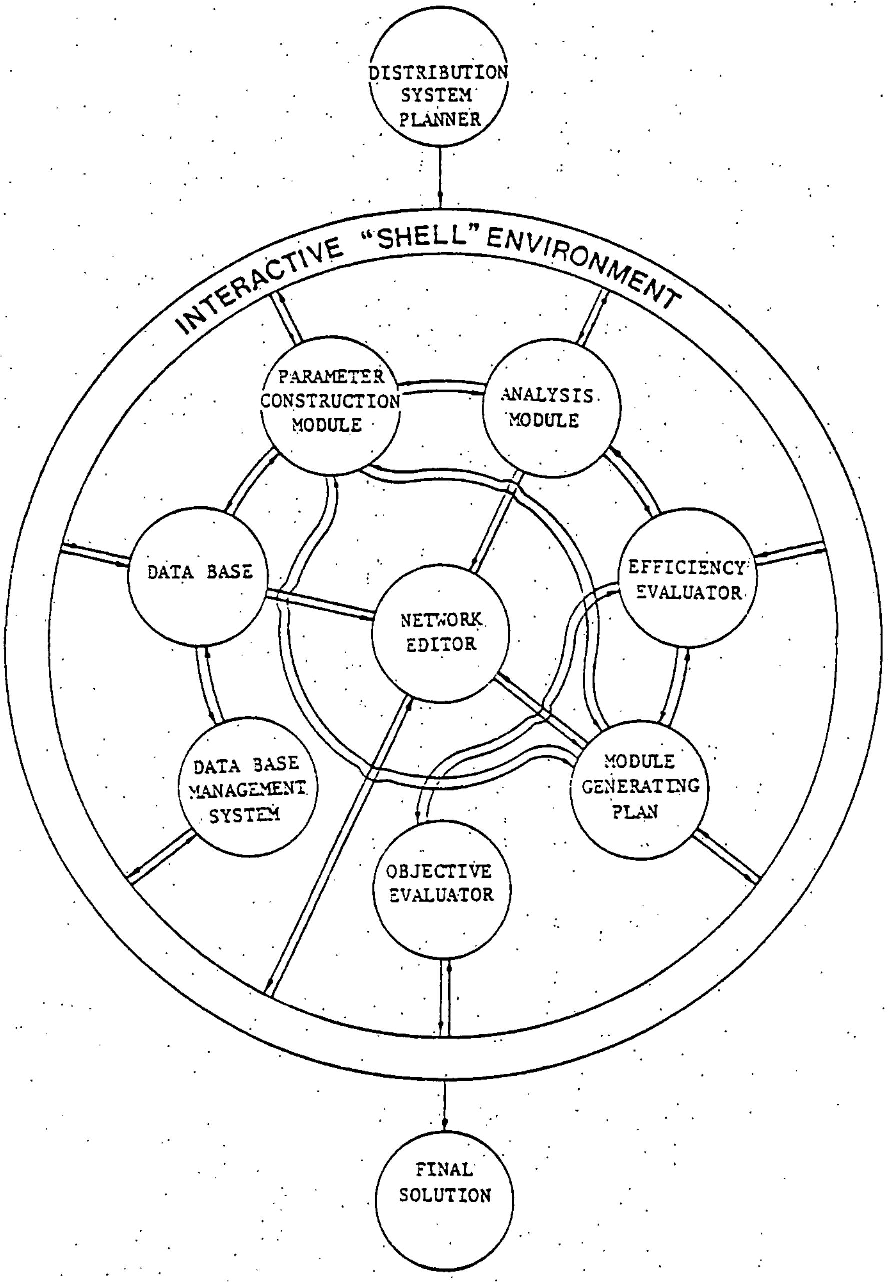


Figure 1

coordinate the above eight modules at the behest of the planner and display the results at any stage of the operation.

The preceding is a list of the nine functions which comprise the design of an interactive distribution planning system. The operation of this system is the cognitive function of the "human" distribution planner. A function is defined here as a system element which "generates" a specific set of outputs, for a given specific set of inputs, according to a given set of rules.

Some of these functions are explained in greater details in the individual task reports. However, some of them are not included in the given tasks, and, therefore, must be included here.

Right at the beginning of the research it became apparent that the data base connot contain direct data elements for the models, since these elements change as models change and cannot be updated automatically with changes in raw data such as cost of conductors without additional computations. The best design then for the total system is to have a parameter computation function which uses the base data in the data base. The data base needs to be kept updated and thus when a model is to run, a parameter construction module is necessary to compute the required parameters which are updated.

The following examples can be given to explain this point. For example, feeder power loss curves are among the very important parameters in the general optimization models. The data base must contain as basic data the present costs of power, carrying charge rates, the present line impedances of existing network arcs, taking into the account conductor sizes and loads, and possible voltage levels. Another example is the cost of incremental capacity additions. They are also important parameters and constructed from the present basic data on transformers, labor charges; cost of conductors, additional equipment, and hardware required; present transformer sizes of those transformers in place and/or available in inventory; other locations in the system where a "traded out" transformer can be used. A subprogram using this present raw data can compute the incremental capacity costs for each substation. Further, the power loss curves can be computed at a given voltage level by a subprogram using engineering economy formulas which outputs the results in a format which can be directly usable in the optimization models. For example, in the case of convex approximation, the results will be a series of slopes based on a preset number of grid approximations. As a further example, demand values must be aggregated based on a given grid approximation of the problem from the raw data set of demands at each coordinate. Of course, the fixed costs of conductor installations must be computed from the length of the feeder (network arc), conductor size, present cost of conductor per unit distance, costs of pole type selected and other hardware, and present labor costs with present estimates of installation time.

Function six and eight evolved from interviews with the distribution system planners. In general, the lowest cost plan is not always chosen and therefore the cost itself is not a prime determinant. The interviews with distribution system planners provided the detailed information in terms of the planning objectives and constraints, which are given in task B.5, that must be taken into account. These objectives and constraints force a multi-criteria evaluation to be a part of the functions of the integrated planning system.

A survey of the current literature indicates that the basic concept formulated by Churchman, et al[1], is still the basis for solving multi-objective decision problems. Goal programming, pareto optimality, utility theory, and the basic procedure suggested by Churchman is currently in use today.

Both goal programming and pareto optimality approaches were rejected because it would require the planner to reformulate linear programming problems if his objectives changed and would violate simplicity requirements on optimization approaches. The user is not expected to be an expert in mathematical programming. In addition, these approaches do not offer any valuable gain in accuracy of decision in the type of problem addressed. The basic procedure as suggested by Churchman is:

1. State clearly independent and mutually exclusive objectives.
2. Order the objectives.
3. Weight the objectives.
4. State clear measures of attainment of the objective by means of efficiency curves.

The third step can be done by a decision tree approach suggested by Churchman or by a more elegant approach suggested by Saaty[2] in 1977. The Saaty approach, although it was attractive, was not used because subjective preference assessments must still be made which require some additional subprograms to compute eigen values and again there is no demonstrated superiority in decision accuracy. The Churchman method, on the other hand, is simple and gives the user a control over the process. The Saaty process was tested by the research staff to check its performance.

The efficiency curves are developed by the Churchman method also rather than following a utility approach. The utility curve approach is demonstrated by Crawford[3], et al in 1978 but no advantage to using this more complex method was shown. Bammi[4], in 1979, provided a typical example that simple approaches are still very successful in both weighting and efficiency calculations. A summary of the details of the multi-objective evaluation process suggested by the research team is given in Appendix E along with the results of a trial run at OG & E.

The functions six and eight require subprograms to allow the planner to construct the efficiency curve for a given measure using either exponential, parabolic or straight line curves and then subprograms to compute efficiency from given values of the measure which are either computed by other programs or input by the planner. Other subprograms are required to allow the planner to determine the weights.

This leads to function seven. In tasks B.6-B.13, a detailed description of the subprogram function which will represent a present or proposed distribution system network with appropriately defined nodes and arcs is described. This function will allow a planner to construct a proposed distribution plan of his own design which can be analyzed by the analysis module and then an overall evaluation computed by functions six and eight.

Function seven would also require a subprogram which will construct the mathematical optimization model, obtain its parameters from the parameter module and then allow the planner to input starting solutions and advanced trial solutions to speed the proof of the optimality. The key to the design here is the planners control of the representation. The planner must be able to indicate possible feeder routes, substation locations, the number of conductor size possibilities, limit the number of incremental capacities considered, and set routes where reconductoring might be beneficial. The optimization model must be constructed off the subset of the total number of possibilities indicated by the planner. This would impose a strong requirement on function nine. The planner must be able to interject at any step of the process and control the action step by step.

Implementation of this system would require a large number of subprograms to be written with several large master programs with the capacity to fit in new subprograms at will. However, each individual program called for by the nine functions can be done with present computer technology and each computation can be accomplished with methods known today. The only limit in some cases is the size of the planning problem except in one model. These limits are dealt with in task B.3. The appropriate optimization models are described in Task B.6 and a model which requires new solution techniques is explained.

Conclusions

The preliminary results of a research effort to apply the latest technology to the problem of energy distribution system planning has been described. The approach is a systems approach, characterized by an attempt to see the system of a network or directed graph, the vertices of which are described by a number of parameters which bind each vertex to a physical component or collection of physical components constituting the distribution system. Just as the network model

---

[1] Churchman, C.W., R.L. Ackoff, and E.L. Arnoff, Introduction to Operations Research, New York: John Wiley & Sons, Inc., 1957

[2] Saaty, L.L., "A Scaling Method for Priorities in Hierarchical Structures", Journal of Mathematical Psychology, Vol. 15, June 1977, pp. 234-281.

[3] Crawford, D.M., B.C. Huntzinger, and C.W. Kirchwood, "Multi-objective Decision Analysis for Transmission Conductor Selection", The Institute of Management Sciences, Vol. 24, No. 16, Dec. 1978, pp. 1700-1710.

[4] Bammi, D. and D. Bammi, "Development of a Comprehensive Land Use Plan by Means of a Multiple Objective Mathematical Programming Model", Interfaces, Vol. 9, No. 2, Part 2, Feb. 1979, pp. 50-64.

unifies the conceptual view of the distribution system, a relational data base model unifies presentations of all pertinent data, including that data representing the network model.

The collection of algorithms suitable for implementation on a digital computer, which maintain the network model and allow a designer to work interactively with it, are divided into three classes. An overseer called the shell, which processes the planner's demands, providing his access to a number of problem solving programs, a network editor, which allows the planner to create and modify networks using concepts and commands natural to the subject, and finally, a data base management system which maintains the data base. While the end result of the research is conceptual in nature, selected features of the Distribution System Planning Model have been implemented. This implementation provides a validity check of the notions and abstractions comprising the DSPM.

Figure 2 and 3 display graphically the tasks which the contractor agreed to perform. These tasks are shown as a part of an overall set of tasks that will result in an operating interactive distribution planning system. Task blocks encased in solid lines represent the tasks of this project. The crosshatching represents the completion of the task.

In addition to the tasks required in the work statement, considerable progress has been made toward the prototype distribution planning system as is indicated by the dashed blocks partially crosshatched in

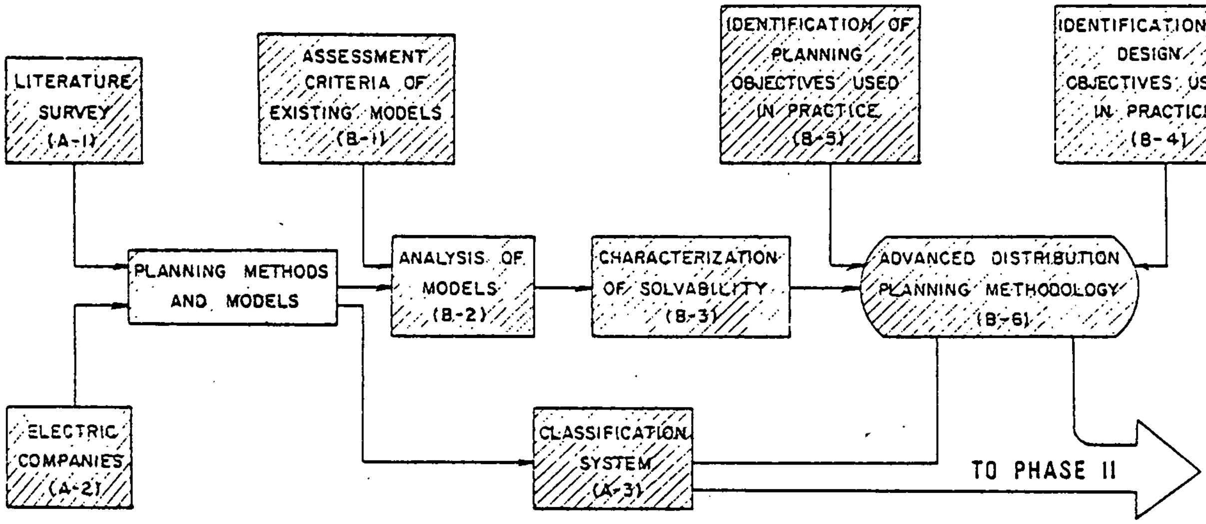


Figure 2. Phase I

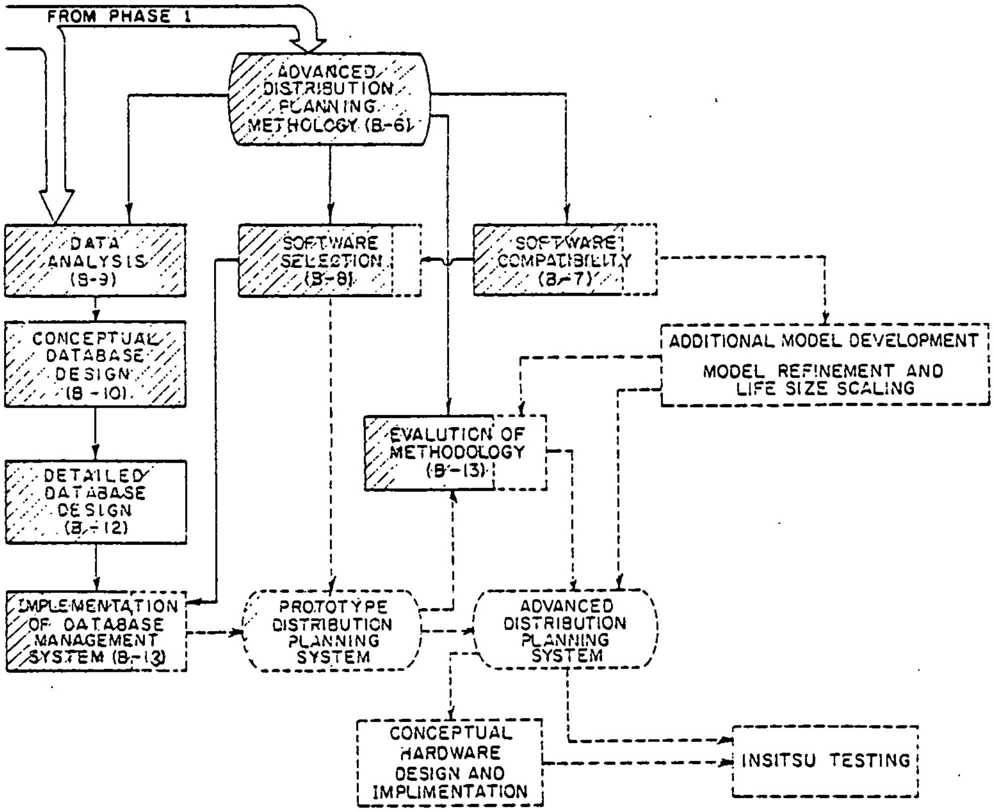


Figure 3. Phase II

Figure 3. This prototype system is described by the nine functional elements presented in the executive summary plus task reports B.6, B.12, and B.13.

As a result of the research performed under tasks B.12 and B.13, some progress has been made toward the implementation of a prototype distribution planning system. The blocks entirely closed by the dotted lines represent the research efforts which require the results of the present project but which currently lie outside the scope of this project.

TASK A.1: THE SURVEY OF LITERATURE AND EXISTING COMPUTER PROGRAMS

TASK A.1(a): LITERATURE SURVEY

An extensive search has been made for the identification of papers and reports published in the open literature which describe the distribution planning methods and techniques. The bibliography is included at the end of this report. References have been selected which deal predominantly with the distribution system. Emphasis is placed on references which illustrate practical as well as theoretical applications of distribution system planning techniques. The listing of the titles is subdivided into three sections, namely; (1) analyses, (2) models, (3) techniques, depending upon the general substance of each article. However, a title may be listed in more than one section if the paper covers material that can be included in various sections.

The entries in each section are listed in alphabetical order. The last name of the first order author determines the alphabetical position. Only the more readily available foreign publications are included. A list of the periodicals which have been cited and their place of publication is given following the bibliography.

TASK A.1(b): THE COLLECTION AND EXAMINATION OF SELECTED COMPUTER PROGRAMS

Introduction

A questionnaire has been prepared and sent to a large number of electric power utility companies. Also, a large number of these companies were visited and their distribution planners interviewed by the authors for the identification and description of distribution system planning methods and techniques used by these electric power utility companies and other commercial entities. The information gathered from the interviews will be reported in the future.

The Computer Programs Used for Distribution System Planning

Today, many electric distribution system planners in the industry utilize computer programs, usually based on ad hoc techniques, such as load flow programs, radial or loop load flow programs, short circuit and fault current calculation programs, voltage drop calculation programs, and total system impedance calculation programs, as well as other tools such as load forecasting, voltage regulation, regulator setting, capacitor planning, reliability and optimal siting and sizing algorithms, etc. However, in general, the overall concept of using the output of each program as input for the next program is generally not in use. Of course, the computers do perform calculations more expeditiously than other methods and free the distribution engineer from detailed work. The engineer can then spend his time reviewing results of the calculations, rather than actually making them. Nevertheless, there is no substitute for engineering judgment, based on adequate planning at every stage of the development of power systems, regardless of how calculations are made. In general, the use of these tools and their bearing on the system design is based purely on the discretion of the planner and overall company operating policy.

Collection and Examination of Selected Computer Program

In order to determine the present scope of computer applications in distribution system planning within the power industry, a large number of electrical power utility companies as well as other sources have been contacted. In addition, the 1976 and 1977 editions of the "Computer Program Availability Reports" of the Edison Electric Institute (EEI) have been screened. The EEI reports list computer programs possessed by each of the member companies and which are readily available to any other member company on request.

Through the Oklahoma Gas and Electric (OG&E) Company and consultants of this project, many computer programs that could be used in distribution system planning were gathered. Some additional information about the use of each computer program was also obtained. This ranged from simple input/output requirements to detailed user's manuals. Of the 28 programs thus gathered, 16 performed circuit analysis. Although analysis itself is not planning, these programs are still of interest since any proposed system alternative must be analyzed to assure that requirements are satisfied. However, some of the analysis programs received offer a limited planning capability in the form of capacitor application or load growth. The analysis programs are compared in Table 1. Also, some additional information and the input/output requirements of the selected programs are given in Sections A.2 and A.3(b). Even though most of the entries are self-explanatory, the following additional comments are pertinent:

a) Most programs assumed balanced three phase conditions; only three programs treated unbalanced cases.

b) Two programs provided the option of inserting impedances in the fault path before calculating fault currents.

c) Six programs provided losses for each line and the total system while three programs provided only for some form of total system.

d) Obviously, capacitor planning could be done in a primitive way with any analysis program simply by altering the input data to include, exclude or move capacitors. The difficulty of the task would depend on the method of providing input data. Those programs noted as providing capacitor planning "by modifying input data" are the ones which claimed capacitor planning capability in their documentation and altered the input data to accomplish it.

e) One program provided a very elegant and comprehensive treatment of load growth. It allowed specification of different growth rates at four different levels of the system, namely: substation, feeder, any lines, and any nodes. These rates could then themselves be changed from year to year.

f) The programs all make some use of direct access storage, mostly to preserve cable and device characteristics and system description files. However, some of the programs build arrays on a disk for processing; thus the problem size in this case is limited only by the amount of disk space provided.

g) Finally, there is only one program which can handle a loop or network system, the rest of the programs are restricted to only radial systems.

The other eleven computer programs collected are so diverse in application that comparison would be meaningless. Instead, the following brief descriptions are provided:

P5

This program uses load density on a grid basis to optimize location of a new substation among existing substations. It assigns load to new and existing substations. IBM 370/145, 100k memory is required. It is

Table 1. A Performance Summary of Collected Distribution System Analysis Programs

PROGRAM OUTPUT	P1	P2	P3	P4	P6	P7	P12	P14	P15	P16	P18	P19	P20	P22	P23	P25
Voltage Profile	By-Phase	3φ	3φ	3φ	3φ	3φ	3φ	By-Phase	3φ	3φ	No	Substation Bus V Only	No	3φ	No	By-Phase
Power Flow	No	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	No	No	No	Yes	No	Yes
Current Flow	By-Phase	3φ	3φ	3φ	3φ	3φ	3φ	No	3φ	3φ	No	No	No	3φ	No	No
Fault Current	3φ, φ-φ φ-G inserted impedance	3φ φ-G	No	3φ, φ-φ, φ-N, φ-G	3φ, φ-φ, φ-G	3φ, φ-φ, φ-G	No	3φ, φ-φ, φ-G inserted impedance	3φ, φ-φ, φ-G	Fault MVA	3φ, φ-G	3φ, φ-G	3φ, φ-φ, φ-G	3φ, φ-G	3φ, φ-φ, φ-G	3φ, φ-φ, φ-G
Motor Start Voltage Dip	Two Loads Simultaneously	No	No	No	Yes	No	No	Two Loads Simultaneously, Load Nodes Only	No	Volts Dip Per KVA Inrush	No	No	No	No	No	No
Total Impedance To Node	Yes	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	Yes	No	No
Losses	Yes	Yes	Yes	Yes	Estimated Annual KWHR & $	Totals	No	Yes	No	Yes	No	No	No	No	No	Total KW
Regulator Settings	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	Yes
Planning Aids Available																
Capacitor Planning	Add, Change, Delete	No	No	No	By Modifying Input Data	No	No	By Modifying Input Data	By Modifying Input Data	No	No	No	No	No	No	Yes
Load Growth	System Growth Factor, 20 Periods or Rates	System Growth Factor, 1 Period	No	No	By Changing Demand Factor	System Growth Factor, 1 Period	System Growth Factor, 1 Period	Very Versatile	No	No	No	No	No	No	No	System Growth Factor, 5 Periods
Program Characteristics																
Problem Size	2000 Nodes 1000 Trans 20 Cables	75 Nodes 20 Branches	200 Nodes 99 Cables	150 Nodes	500 Nodes	3000 Nodes	200 Nodes	1000 Nodes	1440 Nodes	100 Nodes 200 Lines	200 Lines	Substation Bus	Limited By Disk Space Only	400 Nodes	150 Nodes	500 Nodes
Memory Requirements	508 K	130 K	33 K	64 K	238 K	832 K	110 K	214 K	29K+180K Disk	210 K	100 K	30 K	100 K	24 K	125 K	150 K
Computer Type	IBM 370/158	Univac Spectra 70/45	CDC 6000	IBM 370/145	IBM 370/158	IBM 370/158	PDP 10	IBM 370/155	IBM 370/145	IBM 370/158	IBM 370/165	IBM 360/65	IBM 370/145	IBM 1800	IBM 370/185	Honeywell 66/22
Interactive	Available	Available	No	No	Yes	No	Yes	No	No	No	Yes	Yes	No	No	No	No
Remarks						Uses TLM Data as Input				Not Restricted to Radial System		Fault Current At Substation Bus	Data Base Maintenance Tool	Can Connect to "Backed-Up" System of Same Size	Calcomp Plotter: Tree Diagram of Fault Currents for Fuse Coordination	

simple but could be very useful in planning.

P8

It can be used as a distribution transformer load management (TLM) tool. It extracts demand data from customer account record through account number. This program provides input data for analysis program P7. IBM 370/158, 100K memory is required.

P9

This program calculates overloadability of large pad-mount transformers as a function of ambient temperature and allowable loss of life. IBM 370/158, 100K memory is required. It could have limited application in planning.

P10

It generates loading and unbalance tables for different combinations of single phase and three-phase loads. It then selects transformer sizes for grounded Y-ungrounded Y transformers. IBM 370/158, 100K memory is required.

P11

It is the same as P10 except it is for open Y-open Δ and open Δ - open Δ transformers.

P13

This program calculates the economic loading of pole-mounted transformers. IBM 370/158, 192K memory is required.

P17

It is for distribution reliability analysis. The program quantitatively compares the performance of distribution circuits for various alternatives. IBM 370/165, 96K memory is required. It could be very helpful in evaluating a proposed circuit configuration.

P20

This program selects the most economical conductor for a distribution feeder, considering initial costs, maintenance costs, capitalization costs and losses throughout the expected life of the feeder. IBM 370/145. It appears to be a very useful program.

P21

It calculates and predicts loads on a substation transformer and up to six associated circuits. IBM 1800, 24K memory is required. It has limited usefulness. The same functions could be accomplished within other analysis programs.

P24

It calculates the size and number of substations required for a rural system. Its output includes feeder lengths, and loading and total costs.

P28

This program uses system analysis techniques to optimize expansion of an interconnected system of distribution substations. IBM 1130, 150K memory is required. The present data input technique of the program is very much company oriented. It would be a valuable planning tool if modified for more general use.

TASK A.2: THE IDENTIFICATION AND DESCRIPTION OF DISTRIBUTION PLANNING METHODS AND TECHNIQUES IN USE BY ELECTRIC UTILITIES AND OTHER COMMERCIAL ENTITIES

Introduction

In 1953, an Edison Electric Institute subcommittee made the following definition of systems planning:

"System planning is the preparation of a rational program for the development of an electric power system, so that it can evolve in an orderly and economic manner. It includes forecasting and analyzing loads, rationalizing standards of service, anticipating trends in equipment design and coordinating the various elements of the system into a well-designed whole; it is particularly concerned with plans for changes and additions to generation, transmission, substations and distribution facilities. It is not concerned with the problems of day-to-day operation or design except to the extent that those problems affect future system development. Briefly, electric system planning is the process of determining when, in order to assure adequate electric service at minimum average annual cost to the community."[1]

This definition has been widely accepted within the planning community. Long-range planning provides a conceptual framework to support each phase of future system expansion. Thus an orderly system development may be realized with each addition being an integral part of the future system. Effective long-range planning cannot be performed independently on separate parts of the system since changes in one part can influence requirements of another part. In short, long-range planning helps to:

*Identify the future system requirements.
*Define the system constraints.
*Define the possible planning alternatives.
*Evaluate each planning alternative.
*Define the optimum long-range system configuration.

Dillard and Sels[2] concluded that system planning is necessary because systems grow. In any system, there comes a time when existing capacity is not sufficient to meet the future demand with an adequate margin for emergency situations. Therefore, the system planner must anticipate the future in such a way that service reliability is maintained with minimum revenue requirements. Planning in the present which does not lead to a near optimum trade-off between these competing factors will create highly undesirable situations in the future.

Thus, system planning is essential to assure that the growing demand for electricity can be satisfied by distribution system additions which are both technically adequate and reasonably economical. Even though considerable work has been done in the past on the application of some type of automation concept to generation and transmission system planning, its application to distribution system planning has unfortunately been somewhat neglected. In the future, more so than in the past, electric utilities will need a fast and economical planning tool to evaluate the consequences of different proposed alternatives and their impact on the rest of the system to provide the necessary economical, reliable and safe electric energy to consumers.

The objective of distribution system planning is to assure that the growing demand for electricity, in terms of increasing growth rates and high load densities, can be satisfied in an optimum way by additional distribution systems, from the secondary conductors through the bulk power substations, which are both technically adequate and reasonably economical. All these factors and others, e.g., the scarcity of available land in urban areas and ecological considerations, can put the problem of optimal distribution systems planning beyond the resolving power of the unaided human mind. Distribution system planners must determine the load magnitude and its geographic location. Then the distribution substations must be placed and sized in such a way as to serve the load at maximum cost effectiveness by minimizing feeder losses and construction costs, while considering the constraints of service reliability.

In the past, the planning for the other portions of the electric power supply system and distribution system frequently has been authorized at the company division level without review of or coordination with long range plans. As a result of the increasing costs of energy, equipment and labor, better system planning through use of efficient planning methods and techniques is inevitable and important. The distribution system is particularly important to an electric utility for two reasons: (1) its close proximity to the ultimate customer, and (2) its high investment cost. Since the distribution system of a power supply system is the closest one to the customer, its failures affect customer service more directly than, for example, failures on the transmission and generating systems, which usually do not cause customer service interruptions.

Therefore, distribution system planning starts at the customer level. The demand, type, load factor and other customer load characteristics dictate the type of distribution system required. Once the customer loads are determined, they are grouped for service from secondary lines connected to distribution transformers that step down from primary voltage. The distribution transformer loads are then combined to determine the demands on the primary distribution system. The primary distribution system loads are then assigned to substations that step down from transmission voltage. The distribution system loads, in turn, determine the size and location, or siting, of the substations as well as the routing and capacity of the associated transmission lines. In other words, each step in the process provides input for the step that follows.

Table 2 shows some of the power supply system components and factors which need to be considered in distribution system planning. Distribution system planning must not only take into consideration substation siting, sizing, number of feeders to be served, voltage levels, and type and size of the service area, but also the coordination of overall subtransmission, and even transmission planning efforts, in order to insure the most reliable and cost effective system design. The subtransmission system includes the major supply of bulk stations, subtransmission lines from the stations to distribution substations and the high voltage portion of the distribution substations. Of course, the expected reliability of the system design that supplies the customer, has an immense effect on the cost of serving the customer.

In current practice, used by most of the utility systems planning departments, the planning engineer partitions the total distribution system planning problem into a set of subproblems which can be handled by using available, usually ad hoc, methods and techniques.

---

[1] AIEE Committee Report, "System Planning Practices," AIEE Transactions, Pt. III (PAS), Vol. 74, Oct. 1955, pp. 896-900.

[2] J.K. Dillard and H.K. Sels, "An Introduction to the Study of System Planning by Operational Gaming Models," AIEE Transactions, Pt. III (PAS), Vol. 78, December 1959, pp. 1284-1290.

Table 2. Power Supply System Components and Factors Affecting the Distribution System Planning

POWER SYSTEM COMPONENTS		FACTORS
Bulk Power Supply System		*Number of Bulk Stations *Location of Bulk Stations *Size of Bulk Stations
Subtransmission System		*Subtransmission Voltage *Radial or Loop System *Economical Conductor Size *Number of Subs Per Group *Line Layout to Serve Various Groups
Distribution System	Substations	*Service Area *Size of Substations *Number of Substations *Location of Substations
	Primary System	PRIMARY MAIN FEEDERS *Conductor Size *Length *kVA Rating *Voltage Level EXPRESS FEEDERS *Conductor Size *Length *kVA Rating *Voltage Level PRIMARY LATERALS *Conductor Size *Length *kVA Rating *Voltage Level *Number of Distribution Transformers Per Lateral CAPACITORS *kVA Per Feeder
	Secondary System	*Conductor Size *Number of Customers *Transformer Size

The planner, in the absence of accepted planning techniques, may restate the problem as an attempt to minimize the cost of subtransmission, substations, feeders, laterals, etc., and the cost of losses. In this process, however, he is usually restricted by permissable voltage values, voltage dips, etc., as well as service continuity and reliability. In pursuing these objectives, the planner ultimately has a significant influence on additions to and/or modifications of the subtransmission network, locations and sizes of substations, service areas of substations, location of breakers and switches, sizes of feeders and laterals, voltage levels and voltage drops in the system, the location of capacitors and voltage regulators, and the loading of transformers and feeders. There are, of course, some factors that need to be considered such as transformer impedances, insulation levels, availability of spare transformers and mobile substations, dispatch of generation, and rates charged to customers. Further information is given in Section B.4.

The Present Computer Applications

Figure 4 shows a basic functional block diagram of a typical distribution system planning process that is followed currently by the utilities. The process can be repeated for each year of a long-range, e.g., 5-10 year, planning period. However, in the development of this diagram, no attempt has been made to represent the planning procedure of any one of the companies specifically, but rather to outline, roughly, a typical planning process. Further information is presented in Section B.4.

Today, many electric distribution system planners in the industry utilize computer programs, usually based on ad hoc techniques, such as load flow programs, radial or loop load flow diagrams, short circuit and fault current calculation programs, voltage drop calculation programs, and total system impedance calculation programs, as well as other tools such as load forecasting, voltage regulation, regulator setting, capacitor planning, reliability and optimal siting and sizing algorithms, etc. However, in general, the overall concept of using the output of each program as input for the next program is generally not in use. Of course, the computers do perform calculations more expeditiously than other methods and free the distribution engineer from detailed work. The engineer can then spend his time reviewing results of the calculations, rather than actually making them.

In order to give a better idea about the usage of

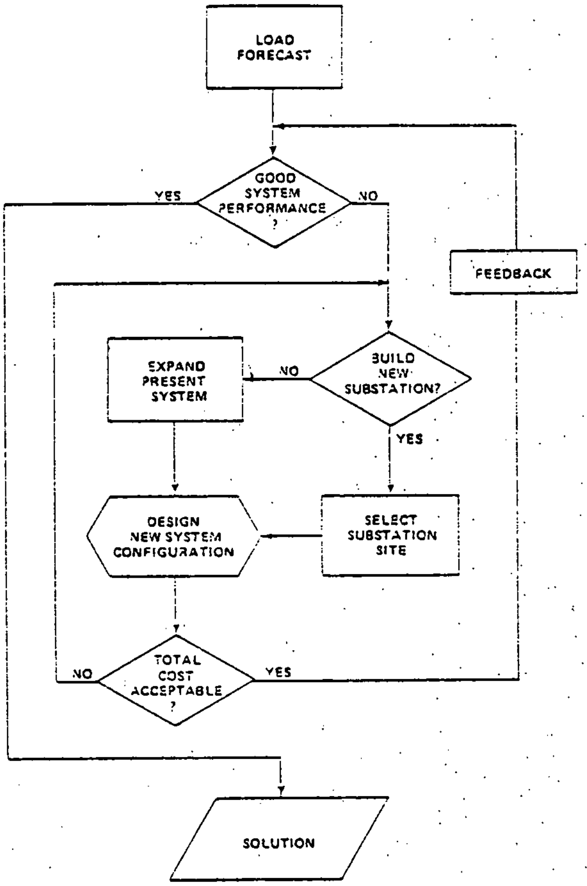


Figure 4. Block diagram of a typical distribution system planning process.

the computer programs by the power industry, the following additional information is given for the selected and examined 28 programs following an index, as shown in Table 3, which associates the program codes, given in this report for the purpose of comparison, i.e., P_i's, with the EEI catalog number and the source company.

P1 NORTHEAST UTILITIES SERVICE CO. IBM 370/158

DATA IN: Conductor impedances; connected kVA; PF; DF; exact loads if known; phase connections; number of customers; line lengths, transformer impedances and ratings; transformer connections and ratios; fault impedances; motor start kVA (or amps) and PF; load growth rate and period; capacitor changes.

DATA OUT: Fault current at every node, using impedance supplied above, or zero impedance if none supplied. Flicker voltage--two motors simultaneous start, any nodes.

Load flow: ϕ-G voltage each phase, % drop, kW loss, kVAR loss, kW load, kVAR load, amps, PF, for every node.

Load growth: same as above for any 20 future years or growth rates.

Summary of loads, losses, caps, max. voltage, min. voltage.

COMMENTS: Interactive version available.
Uses about 900K.
2000 nodes, 1000 tranf./reg.

P2 ATLANTIC CITY ELECTRIC CO. Univac Spectra 70/45

DATA IN: Transformer impedance; system impedance; section length; wire types; number of phases; connected kVA; capacitor switch type; cap or reg. control limits; PF; DF; current limit; kW demand and PF at load centers.

DATA OUT: Voltage levels, current, kW loss, kVAR load; (all 3ϕ only); fault currents.

Table 3. Index of the Selected and Examined Computer Programs

OU Code No.	EEI Catalog No.	Company	Function
P1	061BH06	Northeast Utilities Service Company (NUSCO)	Analysis
P2	084CD01	Atlantic City Electric Company (ACE)	Analysis
P3	100HD02	Consolidated Edison Company (Con Ed)	Analysis
P4	126CH01	Central Hudson Gas & Electric Corp. (CHG&E)	Analysis
P5	126CH02	Central Hudson Gas & Electric Corp. (CHG&E)	Substation Location
P6	132HC06	Niagara Mohawk Power Corp. (NMP)	Analysis
P7	139BI03	New York State Electric & Gas Corp. (NYSE&G)	Analysis
P8	139BI04	New York State Electric & Gas Corp. (NYSE&G)	Transformer Load Mgmt.
P9	139FD01	New York State Electric & Gas Corp. (NYSE&G)	Transformer Capacity
P10	139FD03	New York State Electric & Gas Corp. (NYSE&G)	Transformer Loading
P11	139FD04	New York State Electric & Gas Corp. (NYSE&G)	Transformer Loading
P12	146HC02	Rochester Gas & Electric Corp. (RG&E)	Analysis
P13	196DB03	GPU Services Corp. (GPU)	Transformer Loading
P14	303FD05	Southern Company Services, Inc. (SCS)	Analysis
P15	443DB06	Ohio Edison Company (OE)	Analysis
P16	482AB18	Detroit Edison Company (DE)	Analysis
P17	549GC03	Wisconsin Electric Power Company (WEP)	Reliability
P18	549GC05	Wisconsin Electric Power Company (WEP)	Analysis
P19	554BG01	Northern States Power Company (NSP)	Analysis
P20	731DB01	Oklahoma Gas & Electric Company (OG&E)	Substation Expansion
P21	None	Oklahoma Gas & Electric Company (OG&E)	Analysis
P22	741GE03	Public Service Company of Oklahoma (PSO)	Analysis
P23	None	Public Service Company of Oklahoma (PSO)	Conductor Size
P24	968BK04	Hawaiian Electric Company, Inc. (HE)	Substation Load
P25	968BK07	Hawaiian Electric Company, Inc. (HE)	Analysis
P26	972DL05	Pacific Power & Light Company (PP&L)	Analysis
P27	None	C.H. Guernsey Company	Economic Secondary System
P28	None	Iowa State University	

COMMENTS: Limited to 20 branches, 75 loads, 3φ only.

P3 CONSOLIDATED EDISON CO. CDC 6000
DATA IN: Cable impedances; branch lengths; load data; cap data (fixed, switched); reactor data.
DATA OUT: Transformer kW, kVAR, kVA and PF; static kVAR; primary and secondary voltage and angle; node kW, kVAR, kVA; net loss; branch amps and % loading.
COMMENTS: Interactive TSO.
Balanced 3φ assumed.
200 nodes, 99 cable codes.

P4 CENTRAL HUDSON GAS & ELECTRIC CO. IBM 370/145
DATA IN: Conductor data; span lengths; peak kW; peak kVAR; regulator and transformer data; substation R and X.
DATA OUT: Voltage; short circuit currents; % thermal loading; losses; regulator settings; peak loads; off-peak loads.
COMMENTS: Assumes balanced 3φ for load, voltage and SC calcs, 150 nodes, 3 branches/node.

P5 CENTRAL HUDSON GAS & ELECTRIC CO. IBM 370/145
DATA IN: Load density by grid coordinates; location and kVA size of existing substations; size of each new substation; initial guess for new location.
DATA OUT: Location of all substations; load assigned each substation; load moment each substation; convergence report.

P6 NIAGARA MOHAWK POWER CO. IBM 370/158
DATA IN: Metered current; full load voltage; light load voltage; demand factor; span lengths; wire data; distributed kVA; point kVA; point CkVA; motor code or starting kVA or current.
DATA OUT: Voltage profile; load kVA; DF; PF; total losses (MWHR and $); full load regulated volts; full load unregulated drop; short circuit currents 3φ, L-L, L-N faults; positive and negative sequence R and X; motor start dip; motor run dip; max. start current; voltages with compensation.
COMMENTS: Assumes balanced 3φ.
Load batch, run TSO.
Can alter system by TSO to find best configuration.

P7 NEW YORK STATE ELECTRIC & GAS CO. IBM S/370/158
DATA IN: Substation metered load; wire data; on/off peak kW, kVAR at nodes (can be extracted from TLM file); cap kVAR; number of customers; connection; regulator data; booster data; transformer data; substation data.
DATA OUT: On peak voltages; off peak voltages; on peak load; off peak load; 1φ assignment for best balance; balance conditions; 3φ, L-L, and L-G fault current; thermal loading
COMMENTS: 3000 nodes, 10 regulators.
Assumes balanced 3φ for load and voltage calcs. Method of data entry seems improper.

P8 NEW YORK STATE ELECTRIC & GAS CO. IBM S/370/158
DATA IN: Line number; pole number; transf. code; substation code.
DATA OUT: Individual customer list--demand; Summary list (each line is one complete secondary); diversified and non-diversified demands in kW and kVAR; percent loading on peak, off peak; statistical summary.
COMMENTS: Extracts demand data from customer account record through account number.

P9 NEW YORK STATE ELECTRIC & GAS CO. IBM S/370/158
DATA IN: Temperature tables--include hottest temperature permitted (hot-spot); transformer thermal data (comprehensive description).
DATA OUT: Overloadability of large pad-mounted transformer.

P10 NEW YORK STATE ELECTRIC & GAS CO. IBM S/370/158
DATA IN: Secondary voltage; service conductor length; 1φ and 3φ load power factors.
DATA OUT: Generates loading and unbalance table for different combinations of 1φ and 3φ load. Transformer sizes for above.
COMMENTS: Grounded Y-ungrounded Y transformers.

P11 NEW YORK STATE ELECTRIC & GAS CO. IBM S/370/158
DATA IN: Secondary voltage; service conductor length; 1φ and 3φ load power factors.
DATA OUT: Generator loading and unbalance chart for different combinations of 1φ and 3φ load. Transformer sizes for above. Gives real and reactive load.
COMMENTS: Open Y-open Δ or open Δ - open Δ transformers.

P12 ROCHESTER GAS & ELECTRIC CO. PDP 10
DATA IN: (Details unknown.)
DATA OUT: 1φ, 2φ and 3φ voltage drop.
COMMENTS: Includes regulators, capacitors, ratio banks.
Table look-up for wire impedances.
Data may be edited interactively.
Small program (336 statements).

P13 GPU SERVICES CO. IBM 370/158
DATA IN: Peakload duration; lifetime cost of transformer; transformer rating and loss data; oil temperature data; financial data; load peak data (hourly).
DATA OUT: Economic peak loading summaries for transformer under study.

P14 SOUTHERN COMPANY SERVICES, INC. IBM 370/155
DATA IN: Wire data; span length; wire type; substation voltages; transformer impedances; feeder voltages; load type; load power factor; load kW, kVAR; demand factor; estimated loss; branch fixed or estimated loads; load connection by phase; PF; DF; line transformer/regulator impedances; capacitor ratings; load growth data; inserted fault impedance; motor start kVA or amps and PF.
DATA OUT: Fault currents; voltage dip; voltage by phase; load flow; losses; PF; various summaries and totals.
COMMENTS: Very flexible treatment of load growth.
Handles unbalanced systems.

P15 OHIO EDISON CO. IBM 370/145
DATA IN: Conductor size, type, number of phases; connected distribution transformer kVA; connected capacitor kVAR; primary metered demands.
Also: substation transformer impedance; system impedance; regulator impedance; feeder demand and PF, all supplied by a separate program, also available.
DATA OUT: Short circuit current; load; voltage levels; total impedances.
COMMENTS: For planning and protection studies.

P16 DETROIT EDISON CO. IBM 370/158

DATA IN: Circuit load; circuit PF; system R and X; bus load; bus load PF; capacitor kVAR; bus P and Q; line code; line length; regulator rating.
DATA OUT: Line kW, kVAR, amps and losses; line voltages; load summary; total R and X, fault MVA, and volts drop per 1000 kVA for each bus; Z bus matrix.
COMMENTS: Allows for jumper connections to another circuit.
200 lines, 100 buses, 10 regulators.

P17 WISCONSIN ELECTRIC POWER CO. IBM 370/165
DATA IN: Outage rate per mile, percent outages sustained and average repair time for 20 different system types (i.e., OH vs. UG, number of phases, voltage); branch lengths; number of customers; kW loads; protective devices; alternate sources.
DATA OUT: Frequency and duration outage indices, by customer and load.
COMMENTS: Can be run consecutively with circuit changes to permit comparison of options. Handles several types of fuses, breakers and reclosers.

P18 WISCONSIN ELECTRIC POWER CO. IBM 370/165
DATA IN: System impedance; substation transformer impedance and rating; line code and lengths.
DATA OUT: Symmetric and asymmetric currents for 3Φ and L-G faults; X/R ratios.
COMMENTS: 200 line sections.

P19 NORTHERN STATES POWER CO. IBM 360/65
DATA IN: Type of feed to substation; transformer connections; impedances of feed lines; 3Φ and L-G fault MVA available at source buses.
DATA OUT: Fault current and MVA at both high and low low sides, for 3Φ and L-G faults.

P20 PUBLIC SERVICE COMPANY OF OKLAHOMA IBM 370/145
DATA IN: Line configurations; phase and neutral conductors; device sizes and types.
DATA OUT: Short circuit report for fuse coordination.
COMMENTS: This data base is a versatile model of distribution feeder lines.

P21 HAWAIIAN ELECTRIC CO. IBM 1800
DATA IN: Substation transformer voltages and ratings; substation power factor; growth rate; peak demand kW; circuit and breaker ratings in amps; circuit power factor and growth rate; circuit currents by phase; amount of change in load; compound growth rate.
DATA OUT: Substation and circuit loads up to five future years.
COMMENTS: Limited to six circuits per substation transformer.

P22 HAWAIIAN ELECTRIC CO. IBM 1800
DATA IN: Substation transformer impedance and ratings; system impedances; current ratings of secondary bus, disconnect, circuit breaker, current transformer, and voltage regulator; demand in kVA; current ratings of conductors; load power factors; span lengths; wire impedances; connected load; metered load; circuit configuration.
DATA OUT: Voltage levels; line loads; short circuit currents.

P23 PACIFIC POWER AND LIGHT CO. IBM 370/185
DATA IN: Source impedance; transformer impedance; circuit MVA; circuit configuration; feeder impedances; line lengths.
DATA OUT: 3Φ, L-L and L-G fault currents.
COMMENTS: Uses Cal-Comp plotter to produce graph of fault currents for use in fuse coordination.

P24 C.H. GUERNSEY CO.
DATA IN: Present kW; growth rate; existing area size; total system load levels; costs; impedances; system data; power factors for areas.
DATA OUT: Substations required; kVA per substation; feeder lengths and loading; feeder construction and loss costs; new substation costs.
COMMENTS: A planning tool for a rural distribuiton system; includes up to five transition points.

P25 OKLAHOMA GAS AND ELECTRIC CO. IBM 1130
DATA IN: Substation transformer voltages, rating and impedance; power factor; conductor data; conductor configuration; connected capacitors, transformer, reactor and breaker data; line lengths; growth rates; system data.
DATA OUT: Voltage profile; load analysis; capacitor placement; regulator placement; thermal overload flag; fault currents.
COMMENTS: 3Φ or individual phase output.
500 line sections.
Variable growth rates.

P26 PUBLIC SERVICE COMPANY OF OKLAHOMA IBM 370/145
DATA IN: Cable data; circuit configuration; projected demand; cable installation and maintenance costs; company financial data.
DATA OUT: Most economical conductor selection.
COMMENTS: Considers initial costs, maintenance costs, capitalization costs and costs of losses for entire life of conductor.

P27 OKLAHOMA GAS AND ELECTRIC CO. IBM 1130
DATA IN: Transformer Load Management data; load growth multiplier; substation expansion options; substation data.
DATA OUT: Substation load forecast; optimum system transformer capacity; load transfer for optimum utilization.
COMMENTS: Uses linear analysis and integer programming techniques. Data input is company peculiar.

P28 IOWA STATE UNIVERSITY IBM 370/158
DATA IN: Load data; secondary line patterns; transformer data; installed costs of transformers and its hardwares; installed costs of secondary line conductors; installed costs of service drop; secondary line electrical data; service drop electrical data; fixed charge rate; cost of pole, hardware, or secondary pedestal; cost of unswitched primary voltage shunt capacitors; distribution transformer exciting current; power system investment cost; etc.
DATA OUT: Most economical secondary systems.
COMMENTS: The program checks all designs against user-furnished criteria for conductor ampacity, voltage drop, motor starting voltage dip, and maximum allowable overloading of distribution transformers. Only those designs which are close to the minimum total annual cost and which also meet all design criteria are outputed.

TASK A.3: THE DEVELOPMENT OF A CLASSIFICATION SYSTEM FOR DISTRIBUTION PLANNING AND ANALYSIS TECHNIQUES AND MODELS WHICH RECOGNIZE WIDE RANGE OF FUNCTIONS PERFORMED BY THESE TECHNIQUES AND MODELS, DATA REQUIREMENTS AND WHERE DIGITAL COMPUTER CODES EXIST THEIR INPUT/OUTPUT REQUIREMENTS. THE CLASSIFICATION OF THE TECHNIQUES AND MODELS IDENTIFIED IN TASKS A.1 AND A.2, USING THE AFOREMENTIONED CLASSIFICATION SYSTEM

TASK A.3(a): THE DEVELOPMENT OF A CLASSIFICATION SYSTEM FOR DISTRIBUTION PLANNING AND ANALYSIS TECHNIQUES AND MODELS WHICH RECOGNIZE WIDE RANGE OF FUNCTIONS PERFORMED BY THESE TECHNIQUES AND MODELS

For these tasks a classification system was to be developed in order to examine the state-of-the-art of the current literature addressing electric power distribution planning models and techniques. Since the conceptual design being addressed in this report focused on the total set of decisions that need to be integrated, it was clear that the models of the current literature should be classified by the subset of planning decisions addressed. This would provide to capture, at a glance, the focus of a particular model or technique. This does not, of course, give any insight about the approach in terms of the criteria, modeling approach, solution method, or success in reaching the stated goals, as they have been described in the relavant literature. These aspects have been addressed, to a certain extent, in TASKS B.1 and B.2.

In general, the developed classification system has been presented by a 16 x m matrix. Each row of the matrix represents a specific model or technique identified by its reference number from the bibliography which is included at the end of this report. An X in a given cell of the matrix denotes that the decision of that row of the matrix is addressed by the paper.

Analysis models such as reliability and profile computations are not required to be addressed in this task by definition. The following is a list of the decisions needed to be made in a given distribution system, as defined by the group:

1. The decision of whether a given system's capacity is or is not exceeded.
2. The decision of whether the present configuration can or cannot serve the demand.
3. The decision of whether reconductoring is needed and if needed when it should be done.
4. The decision of whether or not to change the primary distribution voltage and if it is needed when it should be done.
5. The decision of whether or not to add new circuits connecting demands and if it is needed when it should be done.
6. The decision of if, where, and when to transfer loads,
   a) under normal operating conditions,
   b) under emergency operating conditions.
7. The decision of if and when to upgrade the rating of a substation by exchanging the existing transformer with a larger transformer.
8. The decision of when and where to locate a new substation.
9. The decision of when and where to add power factor compensation and how much.
10. The decision of when, where and how many voltage regulators to place.
11. The decision of conductor size in circuit added.
12. The decision of when, where and at what voltage to add a feeder line.
13. The decision of what route to select for a given new feeder.
14. The decision of the effects of demand increase on transmission system.
15. The decision of whether underground service is or is not needed.
16. The decision of whether power losses are or are not achieved.

TASK A.3(b): THE IDENTIFICATION AND TABULATION OF INPUT/OUTPUT DATA REQUIREMENTS OF THE EXISTING OR OBTAINABLE COMPUTER PROGRAMS

The identification and tabulation of input/output data requirements of the 28 computer programs are presented in Tables 4-19.

TASK A.3(c): THE CLASSIFICATION OF THE TECHNIQUES AND MODELS WHICH ARE IDENTIFIED IN TASKS A.1 AND A.3

The classification of the techniques and models which are identified in the previous tasks, in relation to the list of the fundamental decisions that has been presented in the section of TASK A.3(a) is given in Table 20 .

Table 4

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Source																											
Base MVA	MVA												MVA					MVA									
Base Voltage			KV												KV												
Substa. Bus Voltage	KV	KV	KV	KV															KV								
Source Pos.Seq. Res.		%													%		Ω or %		Ω		%	Ω					
Source Pos. Seq. Reac.		%													%		Ω or %		Ω		%	Ω					
Source Zero Seq. Res.																			Ω		%						
Source Zero Seq. Reac.																			Ω		%						
Load Power Factor			✓										✓														
Estimated Losses	%												%														
Supply Circuit No.															✓												
Source Type Code																		✓									
3 φ Fault MVA Avail																		✓									
L-G Fault MVA Avail																		✓									
Substation and Feeder																											
I.D. Number		✓					✓	✓						✓						✓	✓		✓				
Substation Name		✓		✓			✓						✓	✓	✓		✓		✓	✓	✓			✓		✓	
Substation Type		✓																									
Substation Size					KVA																		MVA				
Substation Area		✓																									
Substation Coordinates					✓																						
New Substation Location Guess					✓																						
Number of Substations					✓																		✓				
Predicted Sub-sta. Capacities																							MVA				
Numbers of Connecting Substations																							✓				

Table 5

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Substation and Feeder (cont.)																											
Possible Expanded Substation Capacities																							MVA				
Number of Possibilities																							✓				
New Substation Number																											✓
New Substation Name																											✓
Substation Area Reallocation																											✓
Max Allowable Load Transfer																							MVA				
Costs of Alternative Plans																							✓				
Substation Protective Device																			✓								
Substation Capacitor Size													KVAR														
Substation Fixed Costs																									%		
Substation Variable Costs																									$/KW		
Relay Type		✓																									
Relay Top Setting		✓																									
Relay Lever Setting		✓																									
Current Transformer Ratio		✓																									
Substation Comments													✓	✓										✓			
Transformer Type Code																								✓			
Transformer Size																	MVA			KVA	KVA	MVA		KVA			
Transformer Impedence		%					Ω							Ω			%	%			%	Ω		%			
$\frac{2R_1+R_0}{3}$														Ω													
$\frac{2x_1+x_0}{3}$														Ω													
Transformer Force-cooled Rating																				KVA	KVA						

Table 6

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Substation and Feeder (cont.)																											
Transformer Zer % Loss of Life Rating																				KVA	KVA						
Transformer On % Loss of Life Rating																				KVA	KVA						
Normal High Side Voltage													KV				KV	KV		KV	KV		KV	KV			
Actual High Side Voltage													KV	KV													
Nominal Low Side Voltage																	KV	KV		KV				KV		KV	
Actual Low Side Voltage														KV	V												
Actual ϕ-G Voltage														KV													
Connection Code																		✓				✓					
Neutral Reactance		%																									
Secondary Bus Rating																					AMP						
Secondary Disconnect Rating																					AMP						
Secondary Current Transformer Rating																				AMP	AMP						
Power Target				MW																							
Off-Peak Ratio				✓																							
On Peak Voltage				KV		V	V																				
Off Peak Voltage				KV		V	V																				
Peak Demand																				KW	KVA		MVA	KVA by PHASE	MVA	KW	
Metered Load						AMP	KVA																				
Feeder Number			✓	✓		✓	✓	✓					✓	✓	✓				✓		✓			✓			
Feeder Name		✓											✓	✓						✓	✓						
Uses Feeder Fil			✓																✓								
Feeder Data Record Number																			✓								
System File Name																										✓	
Data File Name																										✓	
Feeder Rating																				AMP	AMP		MVA				
Nominal Feeder Voltage													KV								KV						

Table 7

INPUT \ PROGRAM	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Substation and Feeder (cont.)																											
Actual Feeder Voltage													KV														
Starting Voltage																								V			
Feeder Load													KW,KVA or AMP		KVA						KVA						
Feeder Currents by Phase																				AMP							
Load Type Code													✓		✓												
Number of Loads		✓																									
Load Level		✓																							✓		
Min Load Level		✓																									
Area Size																										mi^2	
Area Load																										KW	
Feeder Load Factor																									✓		
Feeder Power Factor													✓														
Feeder Demand Factor													✓														
Supply Power Factor																								Dec			
Preferred Power Factor																								Dec			
Load Power Factor														%						%	%			Dec		%	
Number of Nodes							✓																				
Number of Feeders																				✓							
Tie Node													✓														
Bus Number														✓													
Jumper Busses															✓												
Jumper Circuit															✓												
Jumper Substation Name															✓												
Jumper Wire Code															✓												
Jumper Length															Ft												
Breaker Rating																				AMP	AMP						
Reactor Rating																					KV						
Reactor Impedance																					%						
Cost of Losses													$/KW-YR														

Table 9

PROGRAM INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Substation and Feeder (cont.)																											
TLM Peak Codes							✓																				
Map Scale Code						✓																					
Case Descriptio		✓																									
Tape Number													✓														
District Number																			✓								
Date of Data																					✓						
Conductor																											
Uses Cable File	✓	✓	✓	✓		✓	✓						✓		✓				✓			✓		✓			
Code Number	✓	✓	✓	✓		✓	✓						✓		✓				✓			✓		✓			
Conductor Name																									✓	✓	
Conductor Size																					✓						
Conductor Type																					✓						
Pos. Seq. Res.	Ω/mi		%/kft	Ω/kft									Ω/mi	Ω				%			%	Ω/mi	Ω/mi	Ω/kft	Ω/mi	Ω/mi	
Pos. Seq. Reac.	Ω/mi		%/kft	Ω/kft									Ω/mi	Ω				%			%	Ω/mi				Ω/mi	
Zero Seq. Res.	Ω/mi												Ω/mi	Ω				%			%	Ω/mi					
Zero Seq. Reac	Ω/mi												Ω/mi	Ω				%			%	Ω/mi					
Ampere Capacity	AMP		AMP	AMP									AMP								AMP						
Installation Costs																									$/mi	$/mi	
Conductor Fixed Costs																									%		
Replacement Cod													✓														
Charging			KVAR																								
Cable Construction																								✓			
Thickness of Sheaf																								mils			
Diameter over Jacket																								In			
Cable Diameter																								In			
Conduit Diamete																								In			
Number Strands																								✓			
Insulation Thickness																								mils			
Number of Neutrals																								✓			
Thickness of Jacket																								mils			
Number Neutral Strands																								✓			
Number of Reconductor Cards							✓																				

Table 9

PROGRAM INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Node																											
Node Number	✓	✓	✓				✓								✓						✓	✓		✓			
Branch Number		✓																									
Nominal Voltage	KV			KV		KV	KV						KV														
Exact Load	KV, KVAR	KW											KW, KVA	KVA													
Connected Load	KVA	KVA	KVA			KVA							KW, KVA	KVA		KW					KVA			KVA			
Metered Load																					KVA			KVA			
Constant Load																				✓							
Peak Load				KW, KVAR			KW, KVAR																				
Off Peak Load							KW KVAR																				
First Load Cent		✓																									
Last Load Center		✓																									
Source Load Center		✓																									
Power Factor	✓	✓											✓		✓												
Demand Factor	✓	✓				✓							✓														
Demand Capacity Factor		✓																									
Loss Factor		✓																									
Load Location Factor																									✓		
Device Type		✓					✓												✓					✓			
Load Type Code																								✓			
Transformer Bank Size		✓																									
Capacitor Size		KVAR	KVAR	KVAR		KVAR	KVAR								KVAR				KVAR	KVAR				KVAR			
Capacitor Type																			✓								
Capacitor Block Size													KVAR														
# Blocks "on"													✓														
# Blocks Availab													✓														
Cost of Fixed Capacitors												$/ KVAR															
Cost of Switched Capacitors												$/ KVAR															
Capacitor Manufacturer Code																			✓								
Capacitor Phase Code													✓						✓								
Control Limits		✓	✓																✓								
Reactor Size			KVA																		AMP			KVAR, AMP			
Reactor Impedanc																								Ω			
Recloser Size																			AMP		AMP			AMP			

Table 10

PROGRAM INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Node (cont.)																											
Recloser Type																			✓								
Fuse Size																			Code		Amp			Amp			
Switch Size																					Amp						
Switch Configuration													✓						✓								
Load Action Cod													✓														
Change Identifier		✓																		✓							
Date of Change																					✓						
Desired Load fo Load Flow		✓																									
Estimated Power Factor for Load Flow		✓																									
Line Section For Load Flow		✓																									
Final Point Indicator																					✓						
Phase Connectio	✓					✓	✓						✓														
Number Customer	✓						✓									✓											
Voltage Code																					✓						
Line																											
Connected Nodes	✓		✓	✓		✓	✓						✓		✓		✓							✓			
Initial Node																✓			✓								
Final Node																✓			✓								
Node Name													✓														
Cable Code	✓	✓	✓	✓		✓	✓						✓		✓		✓		✓			✓		✓			
Wire Descriptio				✓									✓														
Length	Ft	k Ft	Ft	Ft		K Ft/ Map In.							Ft/ Map In.	100 Ft	Ft	Mi	Mi		Ft		Ft	Ft		Ft			
Type Construction		✓														✓			✓		✓			✓			
Number Phases		✓		✓									✓				✓				✓	✓		✓			
GMD																								Ft			
Neutral Indicator																					✓						
Ground Resistance																	Ω		Ω								
Exposure														100 Ft													
Desired Voltage															V												
Voltage Drop Pe KVA-Mile $X10^6$																							V				
Voltage Limit Code																								✓			
File Line Number			✓																								

Table 11

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Line (cont.)																											
TLM Line and Pol #'s							✓	✓																			
Configuration Code																			✓								
Number of Branches		✓																			✓			✓			
Branching Codes														✓								✓		✓			
Customer ID								✓																			
New Customer Cod								✓																			
Distributed Load						KVA																					
Outage AVG/mile																✓											
% Sustained																✓											
Repair Time																Hr											
Protective Device Code																✓			✓								
Branch Containin Instantaneous Device																✓											
Alternate Source																✓											
Branch Containin Auto Isolating Switch																✓											
Lateral Spacing																										Mi	
Action Code													✓														
Capacitor/Regulator Treatment Code																								✓			
Additional Data Code																								✓			
Series Capacitor																											
Number in Series																								✓			
Number in Parallel																								✓			
Voltage Rating																								KV			
Size																								KVA			
Years of Life																									✓		
Number of Load Factors																									✓		
Number of Conductors																									✓		
Location Factor																									✓		
Peak Responsibility																									✓		

Table 12

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Line (cont.)																											
Reserve Factor																									✓		
Map Scale						✓							✓														
Stepup/Down Transformer or Regulator																											
Transformer ID									✓																		
Rating	KVA		KVA	KVA					KVA			KVA												KVA			
Transformer Impedance																								%			
Pos. Seq. Res.	%			Ω			Ω						%											%			
Pos. Seq. Reac.	%			Ω			Ω						%											%			
Zero Seq. Res.	%												%														
Zero Seq. Reac.	%												%														
Connection	✓												✓	✓													
Fixed Ratio	✓																										
Uses Transformer File						✓																					
Transformer Code								✓																			
Loading			✓																								
Secondary Voltage Code																								✓			
Trans. Pr. Volt				KV									KV														
Trans. Sec.Volt				KV			KV						KV														
Regulator Rating		KVA		AMP			AMP												AMP		AMP			KVA			
Max. Boost	%	%		%			%						%		%				%		%			%			
Max. Buck	%	%		%			%						%		%				%		%			%			
Booster				%			%																				
Number and Size Transformers							KVA																				
Number and Size Regulators							KVA																				
Temperature Tables									✓																		
Cooling Type Code									✓																		
Weight of Core and Coils									Lbs																		
Core Losses									W			KW															
Copper Losses									W			KW															
Weight of Tank and Fittings									Lbs																		
Oil Quantity									Gal.																		

Table 13

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Stepup/Down Transformer or Regulator(Cont.)																											
Full Load Winding Rise									°C																		
Full Load Top Oil Temp.									°C			°C															
Hot Spot Conductor Rise									°C			°C															
Pump Losses									W																		
Tank Thickness									In.																		
Area of Top									In^2																		
Area of 4 Sides									In^2																		
Service Voltage										V	V																
Service Conductor Length										Ft	Ft																
PeakLoad Duration(365 Values)												%															
Case Title												✓															
Purchase Price												$															
Labor to Install												$															
Labor to Remove												$															
Salvage												$															
Book Life												Yrs															
Tax Life												Yrs															
Exciting Current												%															
Full Load Reactive Loss												KVAR															
Cost of Core Loss												$/KWHR															
Cost of Copper Loss												$/KWHR															
Temp Rise Coefficient												✓															
Thermal Time Constant												Hrs															
Hourly Peaks(24)												%															
Regulated Node													✓														
Regulator Type																			✓					✓			
Full Load Loss																								Watts			
No Load Loss																								Watts			

Table 14

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Financial																											
State Tax												%															
Federal Tax												%															
Rate of Return												%	%												DEC		
Allocation of General Plant												%															
OPS & Maint. Expense												%															
Adm & Gen Expense												%															
Cap in Bonds												$															
Cap in Stocks												$															
Bond Interest Rate												%															
Production Fixed Costs																									%		
Transmission Fixed Costs																									%		
Production Variable Costs																									Mills/KWHR		
Transmission Variable Costs																									$/KW		
Energy Costs																									Mills/KWHR		
Load Growth																											
Starting Year	✓												✓							✓			✓		✓		
Growth Rates	%						Factor						%							% + 100			%	DEC		%	%
Next Year Load Level		✓																									
Final Desired Load							✓																				
Growth Codes							✓																				
Year Desired													✓										✓				✓
Change in Load																				KVA							
Number Growth Periods																								✓			
Number Years/Period																								✓			
Fault Impedance																											
Node Number	✓												✓														
Inserted Resistance	Ω												Ω														
Inserted Reactance	Ω												Ω														

Table 15

PROGRAM / INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Motor Start																											
Node Number	✓												✓														
Starting KVA	KVA												KVA														
Starting Current	AMP												AMP														
Starting Power Factor	✓												✓		✓						✓						
Phase Connection	✓												✓														
Miscellaneous																											
New Min Volts	V												%														
New Max Volts	V												%														
Min Secondary Volt		V																									
Present Date	✓			✓			✓							✓										✓			
Solutions Desired	✓	✓											✓		✓						✓		✓	✓		✓	
Season/Year		✓																									
New Demand Factor	✓																										
New Power Factor	✓																										
New Capacitor Value	KVAR																										
Case Code		✓											✓														
# Cases,												✓															
Change Case #															✓												
Change Type															✓				✓								
Grid Density Map					KVA																						
Array Size					✓																						
Block Dimensions					✓																						
Max Iterations					✓																						
Fixed/movable Code					✓																						
Print Option Codes					✓								✓		✓					✓			✓			✓	
Card Type															✓												
Punch Output Code																				✓							
Page Spacing Indicator																				✓							
Date of Load Reading																				✓							
Atlas Sheet Numbers																							✓				
Sheet Quadrant Code																							✓				
Area Load																							KVA				
Load Multiplier																							✓				
Last Year of Study																							✓				
First Year of Density Data																							✓				

Table 16

PROGRAM INPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Miscellaneous (cont.)																											
Last Year of Density Data																							✓				
System Peaks for Later Years																							KVA				
Time Limit																							min.				
Town																								✓			
Number Service Areas																										✓	
Years Between Transitions																										✓	
Saturation Weight Factors																											✓
Total Region Load																											KW
% Residential																											✓
% Commercial																											✓
% Industrial																											✓
Commercial Block Loads by Coordinates																											KW
Industrial Block Loads by Coordinates																											KW
#New Houses Per Planning District																											✓
Mapping Options																											✓
Load Class Codes																											✓
# Houses by 1/4 Section																											✓
# Apartments by 1/4 Section																											✓
# Commercial Loads by 1/4 Section																											✓
#Industrial Loads by 1/4 Section																											✓
# Lighting Loads by 1/4 Sect.																											✓
#Public Authority Loads by ¼ Section																											✓
Amount of Load Ea. for above 6 entries																											✓ ✓
Saturation Level																											✓
Planning District#																											✓
Substation Area																											

Table 17

PROGRAM / OUTPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Voltage Profile	By-Phase	3φ	3φ	3φ		3φ	3φ					3φ		By Phase	3φ	3φ			Substa. Bus V Only			3φ			By-Phase			
Power Flow	✓		✓	✓			✓							✓	✓	✓						✓			✓			
Current Flow	By-Phase	3φ	3φ	3φ		3φ	3φ					3φ			3φ	3φ						3φ						
Fault Current	3φ,φ-φ φ-G Ins. Imp.	3φ, φ-G		3φ, φ-φ φ-N, φ-G		3φ, φ-φ, φ-G	3φ, φ-φ, φ-G					No		3φ, φ-φ, φ-G Ins. Imp.	3φ, φ-φ, φ-G	Fault MVA		3φ, φ-G	3φ, φ-G	3φ, φ-φ φ-G		3φ, φ-G	3φ, φ-φ φ-G		3φ, φ-φ φ-G			
Motor Start Voltage Dip	Two Loads					✓								Two Loads		V/KVA												
Total Impedance to Node	✓			✓		✓	✓							✓		✓		✓		✓		✓						
Losses	✓	✓	✓	✓		Est. Ann. Totals KWHR & $								✓		✓									Total KW			
Regulator Settings	✓			✓		✓	✓					✓		✓											✓			
Capacitor Planning	✓					✓								✓	✓										✓			
Load Growth	✓	✓				✓	✓					✓		✓											✓			
Substation Location					✓																							
Total Load Each Substation					KVA																							
Load-Moment Each Substation					✓																							
Load Allocation					✓																			✓				
Convergence Report					✓																							
Customer List								✓																				
Transformer Load Summary List								✓																				
Transformer Overload List								✓																				
Load and Customer Statis. Summary								✓																				
Transformer Size Error Listing								✓																				
% Transformer Overloadability									✓																			
Outside Tank Temp.									✓																			
Transformer Total Load										KVA	KVA																	
Transformer Real Load											KW																	

Table 18

PROGRAM OUTPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P2
Transformer Reactive Load											KVAR																	
Transformer Sizes Selected										KVA	KVA																	
Voltage Unbalance										P.U.	P.U.																	
Transformer Economic Summary													✓															
Life Loss and Costs Summary													✓															
Customer Interruption Frequency Indices																	✓											
Customer Interruption Duration Indices																	✓											
Kilowatt Interruption Frequency Indices																	✓											
Kilowatt Interruption Duration Indices																	✓											
Forecast Substation Load																								KVA				
Optimum System Transformer Capacity																								KVA				
Circuit Loads																	✓											
Substation Loads																	✓											
Forecast Circuit Loads																	✓											
Forecast Feeder Loads																	✓											
Substation Load Summary																	✓											
Yearly Costs for Different Conductors																									$/mi.			
Lifetime Levelized Annual Costs for Different Conductors																									$/mi.			
Number of Substations Required																										✓		
Feeder Lengths																										mi.		
Feeder Loads																										KW		
Feeder, Loss and Substation Costs																										$/mi^2		
Existing Loads by Type																											KW	
Load Increases by Type																											KW	

Table 19

PROGRAM / OUTPUT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
Predicted Loads by Type														
Substation Service														
Program Characteristics														
Problem Size	2000 nds. 1000 trans 20 tables	75 nds. 20 branches	200 nds. 99c.	150 nds.		500 nds.	3000 nds.					200 nds.	nine multiple cases	1000 nds.
Language	Fort	Fort	Fort	Fort	Fort	Fort	Fort PL/I	PL/I	Fort	Fort	Fort	Basic	Fort	COBOL Fort
Memory Requirements	508K	130K	33K	64K	100K	238K	832K	100K	100K	100K	100K	110K	192K	214K
Computer Type	IBM 370/158	Univac Spectra 70/45	CDC 6000	IBM 370/145	IBM 370/145	IBM 370/158	IBM 370/158	IBM 370/158	IBM 370/158	IBM 370/158	IBM 370/158	PDP 10	IBM 370/158.	IBM 370/155
Interactive	Available	Available	✓			✓						✓		

PROGRAM / OUTPUT	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
Predicted Loads by Type														KK
Substation Service														✓
Program Characteristics														
Problem Size	1440 nds.	100 nds. 200 lines	20 const types	200 lines	Substa bus	limited by Disk space	six feeders	400 nds.	150 nds.		500 nds.			
Language	Fort	Fort	PL/I	PL/I	Fort	Fort COBOL	Fort	Fort	Fort	Fort	Fort	Fort	Fort	For
Memory Requirements	29K+ 180K Disk	210K	96K	100K	30K	100K	24K	24K	125K	65K	150K		8K	
Computer Type	IBM 370/145	IBM 370/158	IBM 370/165	IBM 370/165	IBM 360/65	IBM 370/145	IBM 1800	IBM 1800	IBM 370/185	IBM 1130	Honeywell 66/22	IBM 370/145		IBM 370/145
Interactive				✓	✓								✓	

TABLE 20

DECISIONS (Reached by Implementing the Models and Techniques)		MODELS AND TECHNIQUES (By Their Bibliography Numbers)																			
No.	Description of the Decision	2-51	2-40	2-70	2-71	2-79	2-24	2-61	3-3	2-8	2-52	2-31	2-16	2-38	2-2	2-41	2-37	2-44	2-60	2-55	2-59
1	System Capacity is or is Not Exceeded	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X
2	The Present Configuration Can or Cannot Serve the Demand		X	X	X	X	X	X	X	X	X	X			X	X		X	X	X	X
3	If and When to Reconduct					X						X	X			X	X		X	X	X
4	If and When to Change Distribution Voltage											X		X							
5	If and When to Add Circuits Connecting Demands		X	X	X	X	X			X	X	X			X	X	X	X	X	X	X
6	If, Where, and When to Transfer Loads under: a) Normal Operation b) Emergency Operation	X		X	X	X	X		X	X	X	X	X	X	X	X	X	X			
7	If and When to Upgrade Substation Rating by a) Transformer Change out b) Transformer Addition	X		X	X	X					X	X		X				X			
8	When and Where to Build New Substation			X	X	X				X	X	X	X	X	X			X			
9	Power Factor Correction: a) When b) Where c) How Much																				
10	Voltage Regulation a) When b) Where c) How Much							X	X		X		X	X		X	X	X			
11	Optimization of New Conductor Size					X						X	X	X	X				X		
12	Voltage Level Selection for Primary Systems		X									X									
13	Selection of Feeder Route			X	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X
14	Effects of Demand Increase on Transmission System									X								X			
15	Underground Service is or is Not Needed																				
16	Minimum Power Losses are or are Not Achieved							X		X		X	X	X	X		X	X			

TABLE 20 (Cont'd)

DECISIONS (Reached by Implementing the Models and Techniques)		MODELS AND TECHNIQUES (By Their Bibliography Numbers)																			
No.	Description of the Decision	2-9	2-21	3-46	3-4	3-19	3-18	2-42	3-48	2-45	2-81	2-71	2-19								
1	System Capacity is or is Not Exceeded	X	X	X	X	X			X	X	X	X	X								
2	The Present Configuration Can or Cannot Serve the Demand	X	X	X		X			X	X	X		X								
3	If and When to Reconduct					X	X			X											
4	If and When to Change Distribution Voltage									X											
5	If and When to Add Circuits Connecting Demands				X			X	X	X											
6	If, Where, and When to Transfer Loads under: a) Normal Operation b) Emergency Operation					X	X	X		X											
7	If and When to Upgrade Substation Rating by a) Transformer Change Out b) Transformer Addition									X											
8	When and Where to Build New Substation					X															
9	Power Factor Correction: a) When b) Where c) How Much									X	X		X								
10	Voltage Regulation a) When b) Where c) How Much		X	X					X	X	X		X								
11	Optimization of New Conductor Size					X				X	X										
12	Voltage Level Selection for Primary Systems									X											
13	Selection of Feeder Route				X	X	X	X		X											
14	Effects of Demand Increase or Transmission System									X											
15	Underground Service is or is Not Needed																				
16	Minimum Power Losses are or are Not Achieved			X						X	X		X								

TASK B.1: THE DEVELOPMENT OF CRITERIA FOR ASSESSING THE DEGREE TO WHICH EXISTING MODELS MEET THE SPECIFICATIONS DEVELOPED

The following criteria were developed to assess the utility of existing models. It is necessary, however, to clarify that two words have special meanings when used with respect to a model criterion: "must" means that a model which fails to meet the criterion as specified is totally unacceptable, and is rejected irrespective of its performance relative to other criteria; "should" means that a measure is developed to express the degree of acceptability which a model manifests with respect to the criterion, and that overall model utility is expressed in terms of individual criterion measures.

The criteria fall into two categories: those which relate to the model itself, and those which relate to the computer program which is the implementation of the model.

a. Model-Related Criteria

1. The model must functionally fit within the scope of the total distribution planning model being proposed under the DOE Contract.
2. The model must address relevant planning decisions or support other analysis leading to planning decision processes included within the distribution planning tool "shell".
3. The criteria within the model which form the basis for decisions made by the model should be acceptable (e.g., a cost criterion which uses only capital investment may be incomplete).
4. The model should address in full the complexity of the decision. A model which addresses only part of the complexity of the decision should be readily expandable in scope to include a larger portion of the decision-making process.
5. Models should obtain optimal or exact solutions; however, models which produce near-optimal (heuristically-obtained) solutions or approximate solutions may be acceptable (at a lower value of performance measure).
6. All data required by the model must reside within a data base administered by the data base management system defined by the contract, and must be readily available and cost-effective to obtain within the utility industry today.

b. Implementation-Related Criteria

1. Model implementation should run on at least a class of machines roughly equivalent to the IBM System/370 Model 148; that is, the implementation should not require greater capability of an operating system than that found on this class of machines.
2. Model implementation should be amenable to operation in an interactive environment within the machine class defined in criterion 1 above.
3. Model implementation must be restricted in core storage utilization to a maximum of 128K bytes, in order that models be portable across a sufficiently wide variety of machines (implementations which exceed 128K in size may still satisfy this criterion via overlay techniques).
4. Model implementation should be efficient in terms of usage of peripheral devices, in two respects: space efficiency and access efficiency. Space efficiency means that record and block sizes are chosen in a manner which minimizes wastage of the storage capability of the peripheral device. Access efficiency means that block size is chosen large enough, consistent with other constraints, such that the number if I/O requests is minimized; and that frequently-used information is retained by the program in core instead of obtained from peripheral storage each time it is needed.
5. The implementation language should be commonly available and transportable (e.g., FORTRAN). Assembly language programs are therefore rated very poorly on this criterion.
6. Implementation must provide answers in "real time", by which is meant that the implementation, if operating interactively, should have a reasonable response time at the terminal; and, if operating in a background or batch mode, must provide its answers in time to contribute to the information available to the planner at or prior to the time at which he makes his decision.
7. Information display should be acceptable from a human factors perspective; it should communicate using the vocabulary of the distribution planner; it should display only that information which is relevant; and it should operate so that information is easily located by and clearly displayed for the planner.

TASK B.2: THE ANALYSIS OF MODELS IDENTIFIED IN TASKS A.1 AND A.2

In TASK B.1, two different groups of criteria were set up to evaluate models proposed in the literature of the distribution systems planning. Using the developed criteria, the models are evaluated for this task. The evaluation is represented by a matrix with thirteen rows and m columns. Each row represents one of the criteria derived in TASK B.1. Each column represents a specific model or technique identified by its reference number from the bibliography. A symbol in each given cell evaluated the model or technique, described in a specific paper, relative to the criteria. The following list describes the symbols used in the matrix.

Evaluators of the Model-Related Criteria

FM ≡ satisfies criteria 1
NFM ≡ does not satisfy criteria 1

RPD ≡ satisfies criteria 2
NRPD ≡ does not satisfy criteria 2

MC ≡ model criteria expresses total system measures
MCP ≡ model criteria is acceptable
MCU ≡ model criteria is unacceptable

OPTSLN ≡ optimal solutions obtained
HEUSLN ≡ good solutions obtained
APPSLN ≡ only approximate solutions are obtained

DA ≡ data abailable
DCE ≡ data cost effective
DACE ≡ data available and cost effective to obtain

Evaluators of Implementation-Related Criteria:

I1S ≡ satisfies criteria
I1U ≡ does not satisfy criteria

I2S ≡ satisfies criteria
I2U ≡ does not satisfy criteria

I3S ≡ satisfies criteria
I3U ≡ does not satisfy criteria

I4SA ≡ model is efficient in space and access
I4S ≡ model is efficient in space utilization only
I4A ≡ model is efficient in access utilization only
I4 ≡ model not efficient

I5S ≡ satisfies criteria
I5U ≡ does not satisfy criteria

I6S ≡ satisfies criteria
I6U ≡ does not satisfy criteria

I7VRC ≡ display uses correct vocabulary, uses only relevant information and clearly displays it
I7VR ≡ vocabulary good and information relevant
I7VC ≡ vocabulary good and clear display
I7CR ≡ relevant information only and clear display
I7V ≡ vocabulary good
I7C ≡ clear display
I7R ≡ relevant information
I7U ≡ unsatisfactory on all counts

The following tables present the evaluation of the models and techniques according to the given criteria. In addition to the tables, comment illuminating research needs are included.

Some Comments on the Selected Models:

Enver Masud (2-51)

Masud has developed an innovative approach to planning substation capacities. The model optimises the substation capacities subject to a variety of constraints on cost, load, voltage and reserve requirements. It has been successfully applied to 1600 square mile urban area served by 70 distribution substations.

M. Juricek, et al (2-40)

This paper presents a transportation model for feeder modification to alleviate undervoltage conditions and substation expansion plans. It provides a quick but not too accurate guideline planning and as the authors admit, perhaps it could be used to generate good initial or starting plans for better and final proposals.

Crawford and Holt (2-24)

This paper discusses the optimal location and sizes of substations and optimises service boundaries, given the alternative locations for substations and other contraints such as reliability. It uses the well known shortest path algorithm and transportation model. However the model is currently being used on small computers in batch mode only.

Shelton and Mahmoud (2-71) (2-70)

A mixed integer programming approach to distribution planning is given. The example given in the paper is confined to a specially contrived case in which there are only three substations catering to four neighboring areas. A more realistic model with few tens of substations should be considered to really evaluate the capability of such models in real planning situations. The present implementation needs more core space and it takes about 20 minutes to get feasible solutions.

Carson and Cornfield (2-16)

A route finding procedure whose theoretical basis is derived from calculus of variations is used for the design of networks supplying housing estates. A practical design of tapered radial network for 500 housing estates is given.

Hindi, et al (2-38)

A branch and bound / capacitated transhipment model for determining the optimal layout of a radial distribution network is given. The authors have developed a special purpose package with emphasis on the computational efficiency and modest storage requirements so that the method can be used on medium size computing machines.

Garrett, et al (2-27)

It discusses only a radial sub-transmission system which is only a part of the rural electrical poverty system. Hence, it may not be of great use for the overall design objective.

Hindi, et al (2-37)

This paper only considers methods for the determination of the optimum profile for the cables for low-voltage distribution system.

Adams, et al (2-2)
It presents a mixed-integer linear programming approach and elaborates fixed-cost-transportation-type models. A variety of illustrative examples are given.

Okada and Uenosono (2-60) (2-59)
Using a shortest path analysis, this paper presents a procedure for optimal design of distribution systems in the context of long range development planning.

Boardman (2-10)
It discusses a heuristic simulation model for long range planning of sub-transmission and distribution systems. However the details of the implementation are not available.

Lawrence, et al (2-44)
It describes a heuristic model but implementation details are not given.

Goldfield and Lang (2-61)
They apply dynamic programming to long range planning decisions but implementation details are not given.

Converti, et al (2-19)
They have developed a sequence of computer programs for optimum design of distribution systems. However no details of the mathematics involved in the optionisation process are given.

Munasinghe, et al (2-55)
It presents a heuristic approach for long range distribution system planning in the context of capital scarce developing countries. The approach is essentially heuristic and the details of the implementation are not given but good examples are given.

Kujszczyk (2-42)
It is not a full scale model for distribution planning though the methods have been used for medium size housing estate distribution networks.

Table 21

CRITERIA	MODELS AND TECHNIQUES (By Their Bibliography Numbers)										
	2-79	2-51	2-40	2-24	2-3	2-71	2-70	2-31	2-16	2-38	2-27
M1*	FM	FM	FM	FM	FM	FM	FM	FM	FM	FM	NFM
M2	RPD	RPD	RPD	RPD	RPD	RPD	RPD	RPD	RPD	RPD	-
M3	MCP	MCP	MCP	MCP	MCP	MCP	MCP	MCP	MCP	MCD	-
M4	MPC	MFC	MPC	MPC	MPC	MPC	MPC	MPC	MPC	MPC	-
M5	OPTSLN	OPTSLN	OPTSLN	OPTSLN	OPTSLN	APPSLN	APPSLN	OPTSLN	APPSLN	OPTSLN	-
M6	DA	DA	DCE	DCE	DCE	DCE	DCE	DCE	DCE	DCE	-
IR1**	I1S	I1S	I1S	I1S	I1S	I1S	I1S	I1S	I1S	I1S	-
IR2	I2S	I2S	I2S	I2S	I2S	I2S	I2S	I2U	I2S	I2S	-
IR3	I31S	I3S	I3S	I3S	I3S	I3S	I3S	I3U	I3S	I3S	-
IR4	I4SA	I4SA	I4SA	I4SA	I4SA	I4	I4	I4SA	I4SA	I4SA	-
IR5	I5S	I5S	I5SA	I5S	I5S	I5S	I5S	I5S	I5S	I5S	-
IR6	I6S	I6S	-	I6S	I6S	I6U	I6U	I6S	I6S	I6S	-
IR7	-	I7VRC	-	-	-	-	-	-	I7VRC	-	-

*M1 denotes the first criterion of the model-related criteria.
**IR1 denotes the first criterion of the implementation-related criteria.

Table 21 (cont'd)

CRITERIA		MODELS AND TECHNIQUES (By Their Bibliography Numbers)									
	2-37	2-2	2-60	2-59	2-10	2-8	2-44	2-61	2-19	2-55	2-42
M1*	NFM	FM	FM	FM	NFM	FM	FM	FM	FM	FM	NFM
M2	-	RPD	RPD	RPD	-	RPD	RPD	NRPD	RPD	RPD	-
M3	-	MCP	MCP	MCP	-	MCP	MCP	MCP	MCP	MCP	-
M4	-	MPC	MPC	MPC	-	MPC	MPC	MUC	MPC	MPC	-
M5	-	APPSLN	OPTSLN	OPTSLN	-	HEUSLN	HEUSLN	HEUSLN	HEUSLN	HEUSLN	-
M6	-	DCE	DCE	DCE	-	DCE	DCE	DCE	DCE	DCE	-
IR1**	-	I1S	I1S	I1S	-	-	-	-	I1S	I1S	-
IR2	-	I2S	I2S	I2S	-	-	-	-	I2S	I2S	-
IR3	-	I3S	I3S	I3S	-	-	-	-	I3S	I3S	-
IR4	-	I4SA	I4SA	I4SA	-	-	-	-	I4SA	I4SA	-
IR5	-	I5S	I5S	I5S	-	-	-	-	I5S	I5S	-
IR6	-	I6S	I6S	I6S	-	-	-	-	I6S	I6S	-
IR7	-	-	-	-	-	-	-	-	-	-	-

*M1 denotes the first criterion of the model-related criteria.
**IR1 denotes the first criterion of the implementation-related criteria.

TASK B-3: THE SOLVABILITY OF MATHEMATICAL MODELS USED TO DESCRIBE DISTRIBUTION PLANNING MODELS

B.3.1. Introduction

One of the basic tasks of this research project was to assess the solvability of the planning models, especially the two basic models proposed by Masud in 1974 (2.41-43). One of the models proposed was a siting-sizing model which required solution of a (0-1) integer program. As part of this research project, Analysis, Research and Computation (ARC) Company has developed an algorithm which, if implemented, should allow a 20-30 substation problem to be solvable. A 30 substation problem with 10 plans for each substation would result in 300 integer variables (0-1) with 90 constraints. As it can be seen later that some of the current literature shows that problems with 1/3 this number of (0-1) variables could not be solved. However, the algorithm proposed should solve the problems described by Masud in less than 30 CPU minutes. A solution time of less than 10 CPU minutes is not uncommon. A 10 substation problem with 10 plans should solve in less than 7 minutes. This is based on a run time of a 113 variable MIP of 28 CPU minutes to optimality. The ARC report, which is given in Appendix C, envisions approximately a four time decrease in this time.

The transportation model proposed by Masud was expanded to a transshipment model by Wall, et al. (2.64). They have solved a problem with 981 demand points, 22 substations in 0.662 seconds on a CDC 7600. This is clearly within interactive time constraints.

Running these problems separately, of course, is a suboptimal approach. As part of this research project, some general approaches have been formulated. They have a much higher probability of being optimal.

The basic general model with a one period planning base would require in a worst case estimate for a 10 substation, 100 demand point problem, and straight line approximation of loss curves, 3,341 rows and 11,600 continuous variables and 11,620 (0-1) variables. A similar type problem solved by Mairs, et al.* with 2,500 plus rows, 3,700 (0-1) variables and 13,000 plus continuous variables was solved in 90 CPU minutes with an additional 22 CPU minutes to prove optimality on an IBM 370/168 using MPSX-MIP/370.

However, the problem size used in this research project is based on the worst case. If an interactive program is used to build the model, arcs between demand points which would have a very small likelihood of being connected for transshipment purposes could be eliminated as a potential network arc. Similarly, many possible reconductoring possibilities could be eliminated, as well as arcs between substations. The model was written with variables distinguishing these cases so that this possibility could be easily exercised. An actual case of the size proposed would be less than the problem of Mairs and is clearly solvable.

If the cost curves are not approximated by linear functions of power transmitted, but instead by a piecewise linear approximation with grid size equal to four, the number of rows rises to 97,151 and the number of continuous variables rises to 95,920 and the number of (0-1) variables rises to 60,150 for a ten site, 100 demand point problem. This problem is not solvable without new work on data handling techniques and approaches related to the structure of this particular problem. Again, however, an interactive approach which builds the model arc at a time using the planner's knowledge would insure a problem of this size would not have to be solved. In this case, for example, removal of a transshipment arc removes eight constraints plus 5 continuous variables and five (0-1) variables. Thus the interactive model building approach, allowing the planner to construct a potential network from which the model selects a realized network, will pay powerful dividends in solvability.

As an example, a three substation, eight demand point centers problem with transshipment and with straight line approximation of the loss function was solved. The problem had 310 continuous variables, 113 (0-1) variables and 111 rows. It was solved in 28 CPU minutes on an IBM 370/158 using MPSX-MIP and needed an additional 14 minutes to prove optimality. A problem with 110 (0-1) variables, 66 continuous variables and 140 rows was attempted by Shelton and Mahmoud (2.56) on an IBM 370/168. They achieved feasible solutions after five hours of CPU time, but did not find an optimal solution, even though they were using a faster machine. The solution time has been enhanced by submitting the planner's solution as a starting solution. An interactive capability clearly enhanced solvability here. (Note: The package we used was a slower version than used by Mairs. MPSX-MIP vs. MPSX-MIP/370. Mairs found a four fold difference in run time.)

The dynamic version will be approximately of a size T_h times the size of the one period model where T_h is the length of the planning period. At this stage a problem of this size would not be solvable.

The new algorithm proposed by ARC is in Appendix C as well as their report on its effect on solvability. The following table gives the maximum problem size for the general planning model, one period, and piecewise linear approximation with no arcs eliminated.

The size of the problem is approximately given by

$$\text{Rows} \cong k_1(7m^2) + k_2(7nm) + k_3(n+m+k_4)$$

$$\text{0-1 variables} \cong c_1(4m^2) + c_2(4nm) + c_3(n+m+c_4)$$

$$\text{continuous} \cong d_1(7m^2) + d_2(7nm) + d_3(n+m+d_4)$$

where $n = NS$.

Thus it can be seen that increases in n and m are the primary determinants of size. Increasing the number of demand points will require less complexity in power loss approximation. Going beyond twenty demand points will require a simpler loss structure, i.e., losses will be represented by power loss on route i,j = constant x power transmitted x # of miles of route = $(TVC)_{ij}X_{ij}$.

The following table demonstrates the sizes of some reasonably large distribution planning problems and compares them in size with the Mairs' model which has been solved.

Table 22. A comparison of the problem sizes of the models

	OU Model			Mairs' Model
	NS=4 m=10 G=4 R=8	NS=5 m=20 G=4 R=8	NS=10 m=100 G=4 R=8	
# of 0-1 variables	950	3,050	60,150	3,700+
# of continuous variables	910	3,000	95,920	13,000+
# of rows	2,507	5,481	97,151	2,500+

B.3.2. Convex Approximation

Some further improvements in solvability can be obtained if convex representation of the cost curves can be assumed. The envelop curve is not a convex

---

*Mairs, T.G., G.W. Wakefield, E.L. Johnson, and K. Spielberg, "On a Production Allocation and Distribution Problem," TIMS, Vol. 24, No. 15, Nov. 1978, pp. 1622-1631.

curve. It seems, however, that a convex approximation is not out of line since the envelop curve is an envelop of convex curves, as shown in Figure 5.

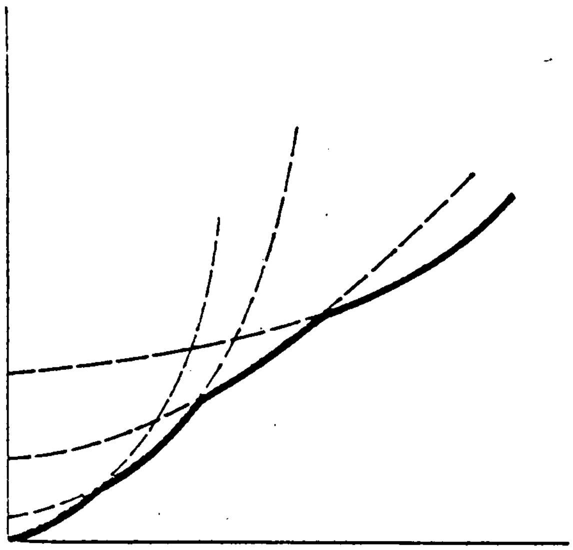


Figure 5. The convex approximation

If all curves are convex, a tremendous reduction in constraints and (0-1) variables can be obtained. Of course (0-1) variables present the most difficult barriers to solution. The following small program illustrates the approach:

$$\text{Min } X_0 = \sum_{i=1}^{4} C_i X_1^{\,i} = X_1^{\,1} + 3X_1^{\,2} + 5X_2^{\,3} + 10X_3^{\,4}$$

$$\text{Subject to } \sum_{i=1}^{4} X_1^{\,i} \geq 5$$

$$X_1^{\,1} \leq 2,\ X_1^{\,2} \leq 2,\ X_1^{\,3} \leq 2,\ X_1^{\,4} \leq 2$$

$$X_1^{\,i} \geq 0$$

Since the simplex seeks the cheapest costs, $X_1^{\,1}$ will come into solution first, then $X_1^{\,2}$, then $X_1^{\,3}$. The solution will be

$$X_1^{\,1} = 2$$

$$X_1^{\,2} = 2$$

$$X_1^{\,3} = 1$$

and $X_0 = 1(2) + 3(2) + 5(1) = 13$

The objective function is convex as shown in Figure 6. The slopes of the piecewise linear terms are 1, 3, 5, 10.

Let us consider $X_0 = \frac{1}{2}X_1^{\,2}$, $X_1 \leq 7$. Let us use a grid size of 5. Since there is an exponent, let us use subscripts for the grid indexing. X_{11}, X_{12}, X_{13}, X_{15} are the grid points. Let $X_{11} \leq 2$, $X_{12} \leq 2$, $X_{13} \leq 2$, $X_{14} \leq 2$, $X_{15} \leq 2$. The curve and its approximation are shown in Figure 7.

The grid sizes can be non-uniform. The approximation would result in a negligible error. In our approximation the slope of the line segments are 1, 3, 5, 7, 9. The uniformity is due to, of course, the nature of a parabola. The curve is then represented for the purpose of linear programming by

$$1X_{11} + 3X_{12} + 5X_{13} + 7X_{14} + 9X_{15}.$$

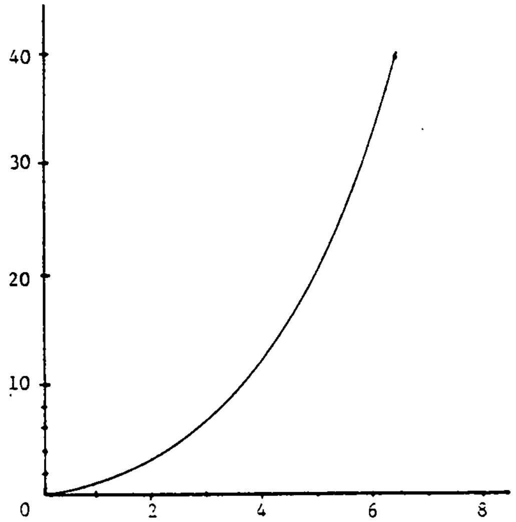


Figure 6. The objective function

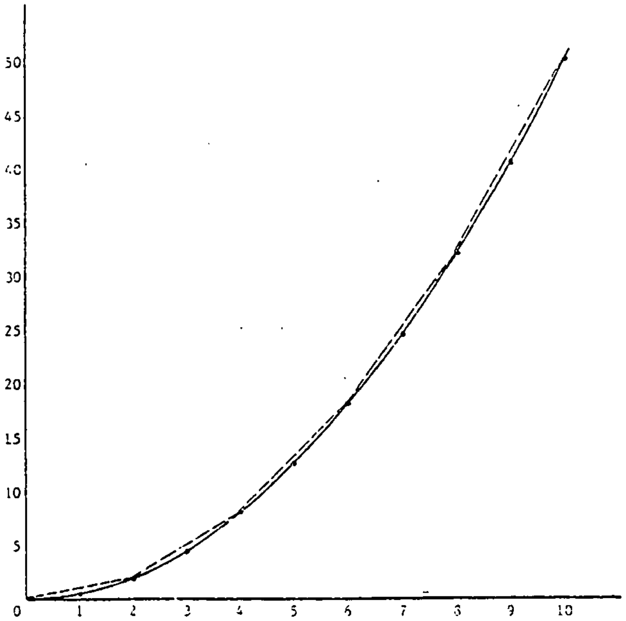


Figure 7.

Thus, if X_1 is power transmitted, then X_1 can be replaced by the above summation. The number of variables have been multiplied by five, but no new constraints and no (0-1) variables have been added. To judge the impact, let us recount the rows and variables for the case NS = 10, m = 100, G = 4, R = 8. In this approximation variables X_{ij}, Z_{ij}, Y_{ij}, $(RS)_{ij}$ would be replaced by

$$\sum_{g=1}^{4} X_{ijg},\ \sum_{g=1}^{4} Z_{ijg},\ \sum_{g=1}^{4} Y_{ijg},\ \sum_{g=1}^{4} (RS)_{ijg}$$ with upper

bounds on each of the variables. Table 23 illustrates some specific examples of using the convex approximation on the model developed in this research.

Table 23. The Effects of the Convex Approximation

	Non Convex Approximation	Convex Approximation
Continuous Variables	95,920	47,960
0-1 Variables	60,150	11,090
Rows	97,151	13,221

Size of Base Model (NS = 10, m = 100, G = 4, R = 8)

Now, as it can be remembered that an interactive approach which would eliminate arcs from the distribution network would result in a much smaller problem. In this case, removing a transshipment arc would remove four continuous variables, a zero-one variable and a constraint. Removing a reconductoring arc would remove four continuous variables, a zero-one variable and two constraints. In an interactive approach many reconductoring arcs would be removed since many of the feeder routes would have no conductors in place. For 100 demand centers many points would have no connections as it would be obvious to the planner that no connections should be made. Most demand centers would have transshipment possibilities to four or five other demand points, not the full one hundred. If each demand center could ship only to five others then the number of variables would be reduced by (4)(9,900) - 4(500) = 37,600 and 9,400 constraints would be removed. The new problem would have 10,360 continuous variables, 1,960 zero-one variables, 3,821 rows. The limit on rows in MPSX-MIP is 13,000. Problems of this type with 3,700 zero-one variables have been solved. Thus this problem is solvable in an interactive model building mode with the planner making decisions throughout the process. The planner's solution will also enhance computation time.'

B.3.3. A New Solution Method

The new solution method developed by ARC (See Appendix C) for LP/GUB knapsack problems improves the ability to solve the least cost substation capacity expansion problem several ways:

(1) The substation capacity expansion problem, formulated as a 0-1 integer programming problem is a "multiple choice" integer program whose choices are identified by GUB sets. It is possible to incorporate all of these GUB sets and any selected capacity constraint of the original IP problem into an LP/GUB relaxation of the IP problem. This relaxation yields: (a) new trial solutions, (b) new fathoming tests, (c) new choice rule information. These are important components of an effective solution approach for the substation capacity expansion problem.

(2) The LP/GUB knapsack relaxation becomes still more powerful by means of surrogate constraint solution strategies. These strategies generate a strong linear combination of the capacity constraints to take the role of the knapsack constraint. This increases the effectiveness of the fathoming tests, as well as providing better trial solutions and improved choice rule information.

(3) The utilization of subgradient optimization techniques makes it possible to generate sharpened surrogate knapsacks at intermediate solution stages, as well as at the outset of the overall solution effort. The modification of the surrogate knapsack, however, can be exploited by using an advanced start for the LP/GUB knapsack problem based on the preceding solution to this problem, following the methodology of the new algorithm.

(4) The solution of the 0-1 substation capacity expansion problem, incorporating the strategies indicated, requires the solution of the LP/GUB knapsack problem many hundreds or thousands of times. Thus, it is especially important to be able to solve this problem efficiently, as the new algorithm makes it possible to do. The efficiency of the new method, furthermore, does not rest simply on intuitive evaluation, but is confirmed by the derivation of computational bounds that strictly dominate the best previously known bounds for this problem. The new bounds are increasingly attractive relative to previous bounds as the size of the problem increases, therefore providing the IP solution method the increased ability to solve substation capacity expansion problems that were too large to handle efficiently in the past.

This new development is expected to result in a several-fold improvement in the speed of solving problems in the range of 10-15 substations, with 5-10 plans each. In addition, it provides the ability to solve problems larger than feasibly possible with previous methods. It is anticipated that problems involving roughly twice as many substations as the 10-15 substation problems should be solvable within entirely reasonable solution times, although the effort to solve these larger problems with previous methods is <u>exponentially</u> greater than the effort to solve the 10-15 substation problems.

Reports indicate that the 10-15 substation problems are the largest feasibly solvable to date. The new method should be able to solve these problems five to ten times faster than before, and also be able to solve problems with 20-30 substations in about the same amount of time currently required to solve problems with 10-15 substations.

TASK B.4: THE IDENTIFICATION OF DISTRIBUTION DESIGN TECHNIQUES USED IN PRACTICE

B.4.1. Introduction

In order to identify the distribution design techniques used in practice by the utility industry a questionnaire has been prepared and sent to a large number of electric power utility companies. Also, a large number of these companies were visited and their distribution planners interviewed by the authors for the identification and description of distribution system planning methods and techniques used by these electric power utility companies and other commercial entities. Some of the information gathered from the interviews and the questionnaires are presented in this section. A sample copy of the questionnaire is included in Appendix B. The companies contacted are:

Oklahoma Gas and Electric Co.	Oklahoma City, OK
Public Service of Oklahoma Co.	Tulsa, OK
Kansas Gas and Electric Co.	Wichita, KS
Texas Electric Co.	Fort Worth, TX
Texas Power and Light Co.	Dallas, TX
Dallas Power and Light Co.	Dallas, TX
C.H. Guernsey Co.	Oklahoma City, OK
Arizona Public Service Co.	Phoenix, AZ
San Diego Gas and Electric Co.	San Diego, CA
Pacific Gas and Electric Co.	San Fransisco, CA
Northeast Utilities Service Co.	Hartford, CT
Georgia Power Co.	Atlanta, GA
Northern States Power Co.	Minneapolis, MN
Public Service Co. of New Mexico	Albuquerque, NM
Consolidated Edison of New York, Inc.	New York, NY
Dayton Power and Light Co.	Dayton, OH
Pacific Power and Light Co.	Portland, OR
Duquesne Light Co.	Pittsburgh, PA
Gulf States Utilities Co.	Beaumont, TX
El Paso Electric Co.	El Paso, TX
Sierra Pacific Power Co.	Reno, NV

As a result of the information gathered from these sources as well as others, a flow diagram of the distribution system planning was developed as can be seen in Figure 4 or 17. However, in developing this flow diagram, no attempt has been made to represent any one of these companies specifically, but rather to outline the thinking process involved in the distribution system planning in general. Several companies used some sort of computerized historical file as a base for load projection. Each company had a system analysis program, one of which sized and placed capacitors and regulators, provided for conductor changing and the load growth. It was evident that there is room for improvement in distribution planning.

The distribution systems planner partitions the total distribution system planning problem into a set of subproblems which can be handled by using available, usually ad hoc, methods and techniques. The planner, in the absence of accepted planning techniques, may restate the problem as an attempt to minimize the cost of subtransmission, substations, feeders, laterals, etc., and the cost of losses. In this process, however, he is usually restricted by permissable voltage values, voltage dips, flicker, etc., as well as service continuity and reliability. In pursuing these objectives, the planner ultimately has a significant influence on additions to and/or modifications of the subtransmission network, locations and sizes of substations, service areas of substations, location of breakers and switches, sizes of feeders and laterals, voltage levels and voltage drops in the system, the location of capacitors and voltage regulators, and the loading of transformers and feeders.

There are, of course, some other factors that need to be considered such as transformer impedance, insulation levels, availability of spare transformers and mobile substations, dispatch of generation, and the rates that are charged to the customers. Furthermore, there are factors over which the distribution system planner has no influence, but, nevertheless, have to be considered in good long-range distribution systems planning, e.g., the timing and location of energy demands, the duration and frequency of outages, the cost of equipment, labor and money, increasing fuels costs, increasing or decreasing prices of alternative energy sources, changing socioeconomic conditions and trends such as the growing demand for goods and services, unexpected local population growth or decline, changing public behavior as a result of technological changes, energy conservation, changing environmental concerns of the public, changing economic conditions such as a decrease or increase in gross national products (GNP) projections, inflation and/or recession, and regulations of federal, state and local governments.

B.4.2. Factors Affecting System Planning

The number and complexity of the considerations affecting system planning appears initially, to be staggering. Demands for ever-increasing power capacity, higher distribution voltages, more automation and greater control sophistication constitute only the beginning of a list of such factors. The constraints which circumscribe the designer have also become more onerous. These include a scarcity of available land in urban areas, ecological considerations, limitations on fuel choices, the undesirability of rate increases and the necessity to minimize investments, carrying charges and production charges.

Succintly, the planning problem is an attempt to minimize the cost of subtransmission, substations, feeders, laterals, etc., as well as the cost of losses. Indeed, this collection of requirements and constraints has put the problem of optimal distribution systems planning beyond the resolving power of the unaided human mind.

B.4.2.1. Load Forecasting

The load growth of the geographical area served by a utility company is the most important factor influencing the expansion of the distribution system. Therefore, forecasting of load increases and system reaction to these increases is essential to the planning process. There are two common time scales of importance to load forecasting; long-range, with time horizons on the order of fifteen or twenty years away and short-range, with time horizons of up to five years distant. Ideally, these forecasts would predict future loads in detail, extending even to the individual customer level, but in practice, much less resolution is sought or required.

Figure 8 indicates some of the factors which influence the load forecast. As one would expect, load growth is very much dependent on the community and its development. Economic indicators, demographic data and official land use plans all serve as raw input to the forecast procedure. Output from the forecast is in the form of load densities (kVA/unit area) for long-range forecasts. Short-range forecasts may require greater detail. Densities are associated with a coordinate grid for the area of interest. The grid data is then available to aid configuration design.

The outputs from a load forecasting program can be detailed as much as desired and feasible. Further, a master grid is developed on a suitable scaled map of the complete service area under study. The master grid presents the load forecasting data and it provides a useful planning tool for checking all geographical locations and taking the necessary actions to accommodate the system expansion patterns. Thus, the master grid facilitates coordination of more detailed information on local area maps. This method provides fast and accurate information and a display of overall future system requirements.

An excellent source of present load data is a Transformer Load Management (TLM) program. Output of a TLM contains a list of all the distribution transformers in the system by their numbers. Once each transformer has been assigned grid coordinates, the TLM file can be an excellent geographical display of the present load and can be used as a starting point for the data base.

By using the energy consumption data, which are readily available from the customer account files, typical demand curves are scaled and the resultant information is used to estimate the peak loading on specific equipment of the system, e.g., distribution transformers, feeders, and ultimately substations.

After establishing the grid system, the next step in a load projection is to fill in the load data using the TLM data and the projected load growth of each square. The grid squares can be classified into three basic types:

1) The ones which are heavily filled with customers. Here the sources of information are the historical TLM data, marketing projections, and demographic data.
2) The ones which are lightly loaded with room for known, or forecasted, expansion.
3) The ones with relatively large vacant areas for which there is no certain load potential or for which forecasting of the timing of future load connections is difficult.

Official land use plans can be used as an additional source of information in forecasting the future load densities in any given area. Also, aerial photography has been used by some utilities to set up the grid system. Furthermore, attempts have been made to develop pattern recognition procedures to aid in reading maps and automatically interpret and process the information.

Three additional code designators covering the remaining factors are assigned to each grid. They are the planning code, saturation code, and class code. The planning code is a district number which is established by metropolitan planning commissions and designates twenty-six different types of districts, each with different growth patterns and rates. The saturation code is assigned by the planning engineer and indicates how "full" the grid is. The class code is also assigned by the planning engineer and reflects the demand expected for each customer in the grid. The initial assignment of these three factors is tedious and unfortunately somewhat subjective. However, once it has been done, the saturation code is updated by the program itself, and the planning code and the class code need to be changed only as a result of some substantial change in the original conditions, e.g., when some new industrial or residential development plans are revealed. The output of the load forecasting program is a forecasted power demand for a given grid area for a specified time period. It becomes a part of the data base management system, where it is available to the other programs.

If, for example, the planner desires to analyze the system performance, the required data is given in the form of grid coordinates, connections, element code numbers, and related parameters. All the physical and electrical characteristics of the system are stored in the data base. For instance, for the primary system the stored parameters are:

1) Conductor: The description, ampacity, positive/zero sequence resistances and reactance.
2) Node: The points on the system for different branches and wire sizes.
3) Section: The parameters for line length, ampacity, positive/zero sequence resistances and reactances, load data (kW, kVAR), growth rates, and special equipment data (capacitors, auto transformers, regulators, etc.)
4) Sequence: The sequence of the interconnections between nodes and sections, the locations of the branches and the terminal points of branches.

B.4.2.2 Substation Expansion

Figure 9 presents some of the factors affecting the expansion of the present system decision to handle the forecasted load. Here, of course, the planner can make the decision himself, without using any of the planning tools, based on information either intangible or difficult to qualify in terms of some suitable form of computerized analysis. Examples are a recent top management decision to delay all new construction projects indefinitely, or an announcement of a substantial number of residential development projects. The forecasted load, load density, load growth, and the distance to the nearest substation may indicate that eventually a substation has to be built to serve the expanding load. Also, under some circumstances it might be advantageous to build the substation right away to prevent some future adversary situations, e.g., complaints by the property holders in the immediate vicinity claiming that the construction of the substation would reduce the value of their property substantially. There might also be some other intangible factors that need to be considered by the planner as he directs the planning process. For instance, the protection limitations are not precisely known until a new system configuration is designed and only at that time can the necessary protection scheme be designed to eliminate any coordination problem.

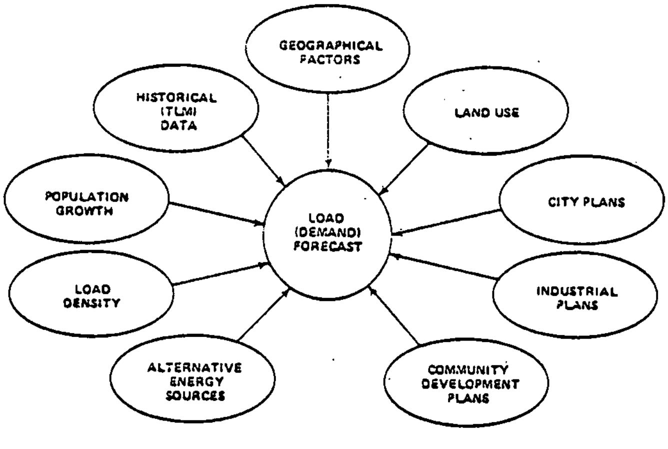


Figure 8. Factors affecting load forecast

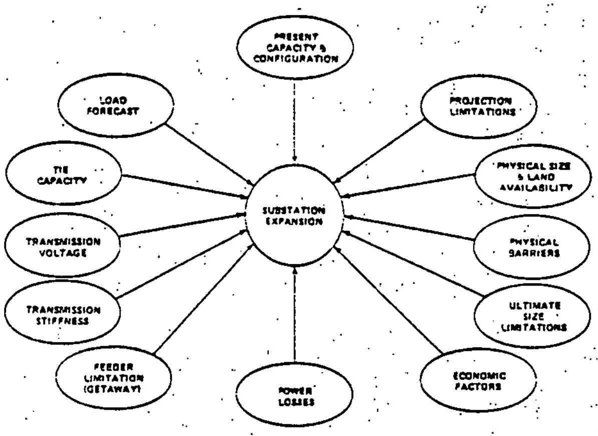


Figure 9. Factors affecting substation expansion

In the system expansion plan, of course, the present system configuration, capacity, and the forecasted loads play major roles. For instance, physical limitations may preclude the installation of any additional feeders. In that case some of the load can be shifted to adjacent substations if the tie line and substation capacities are adequate. Also, the physical size and the availability of adjacent land help to determine whether a particular substation can be expanded. The ultimate substation size limitations, as dictated by company policy and/or other factors, may contribute to this decision. Transmission stiffness, i.e., the conductor capacity of the transmission line, the source capacity, and adequate voltage support for normal and emergency conditions, indicate whether the transmission or subtransmission system can support the additional load in the area.

B.4.2.3. Substation Site Selection

Figure 10 shows the factors that affect substation site selection. The distance from the load centers and from the existing subtransmission lines, as well as other limitations, such as availability of land, its cost, and land use regulations, are important.

The substation siting process can be described as a screening procedure through which all possible locations for a site are passed as indicated in Figure 11. The service region is the area under evaluation. It may be defined as the service territory of the utility. An initial screening is applied using a set of considerations, e.g., safety, engineering, system planning, institutional, economics, aesthetics. This stage of the site selection mainly indicates the areas that are unsuitable for site development. Thus, the service region is screened down to a set of candidate sites for substation construction. Further, the candidate sites are categorized into three basic groups: (1) sites that are unsuitable for development in the foreseeable future; (2) sites that have some promise but are not selected for detailed evaluation during the planning cycle; and (3) candidate sites that are to be studied in more detail.

The emphasis put on each consideration changes from level to level and from utility to utility. Three basic alternative uses of the considerations are: (1) quantitative versus qualitative evaluation, (2) adverse versus beneficial affects evaluation, (3) absolute versus relative scaling of effects. A complete site assessment should use a mix of all alternatives and attempts to treat the evaluation from a variety of prospectives.

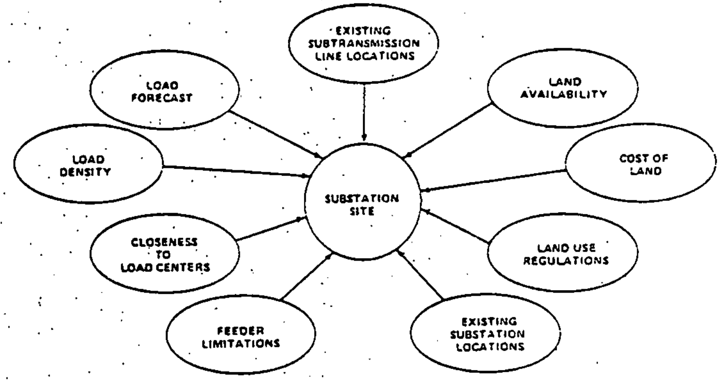


Figure 10. Factors affecting substation siting.

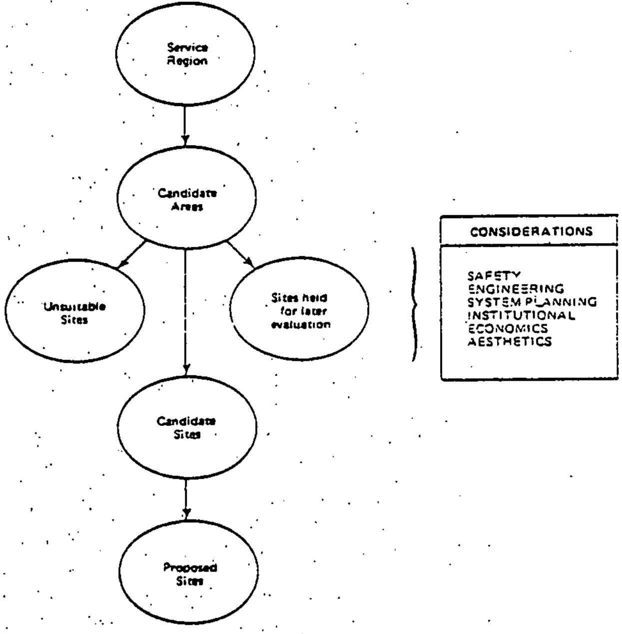


Figure 11. Substation site selection procedure .

B.4.2.4. Other Factors

Once the load assignments to the substations are determined, then the remaining factors, as shown in Figures 12-16 need to be considered. In general, the subtransmission and distribution system voltage levels are determined by company policies and they are unlikely to be subject to change at the whim of the planning engineer unless he can support his argument by running test cases to show substantial benefits that can be achieved by selecting different voltage levels.

Further, because of the standardization and economy that are involved, the planner may not have that much freedom in choosing the necessary sizes and types of capacity equipment. For example, he may have to choose a distribution transformer out of a fixed list of transformers that are presently stocked by his company for the voltage levels that are already established by the company. Any decision regarding addition of a feeder or adding on to an existing feeder will, within limits, depend on the adequacy of the existing system and the size, location, and timing of the additional loads that need to be served.

PRESENT DISTRIBUTION SYSTEM PLANNING TECHNIQUES

In order to gather information on present distribution system planning and design techniques, a questionnaire was distributed to a large number of electric power utility companies. Furthermore, a number of these companies were visited and their distribution planning engineers were interviewed by the authors to identify and obtain descriptions of distribution system planning methods in use.

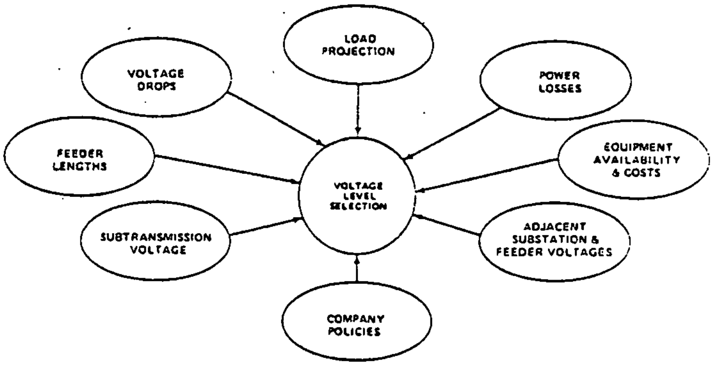


Figure 12. Factors affecting primary voltage selection

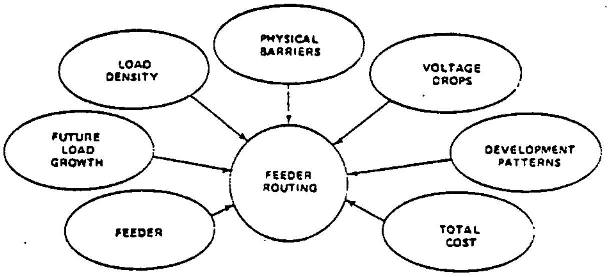


Figure 13. Factors affecting feeder route selection

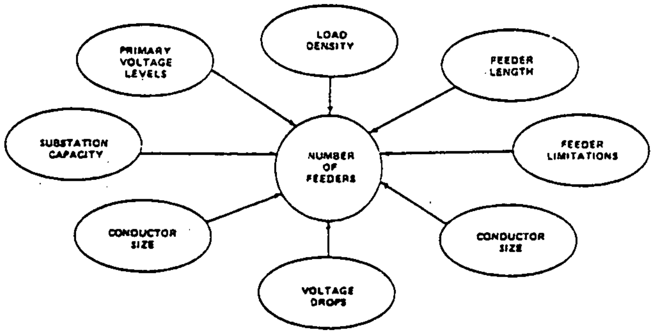


Figure 14. Factors affecting number of feeders

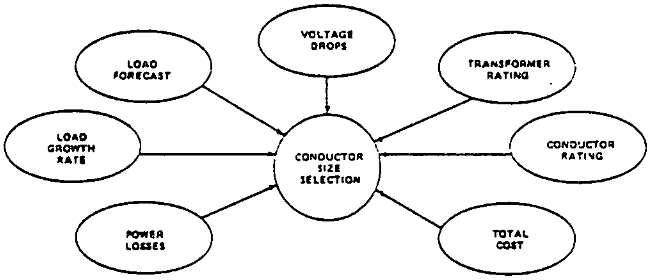


Figure 15. Factors affecting conductor size selection

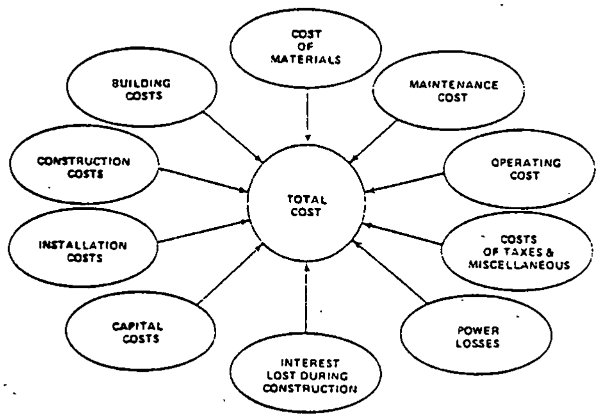


Figure 16. Factors affecting total cost of the distribution system expansion

Today, many electric distribution system planners in the industry utilize computer programs, usually based on ad hoc techniques, such as load flow programs, radial or loop lead flow programs, short circuit and fault current calculation programs, voltage drop calculation programs, and total system impedance calculation programs, as well as other tools such as load forecasting, voltage regulation, regulator setting, capacitor planning, reliability and optimal siting and sizing algorithms, etc. However, in general, the overall concept of using the output of each program as input for the next program is not in use. Of course, the computers do perform calculations more expeditiously than other methods and free the distribution engineer from detailed work. The engineer can then spend his time reviewing results of the calculations, rather than actually making them. Nevertheless, there is no substitute for engineering judgement, based on adequate planning at every stage of the development of power systems, regardless of how calculations are made. In general, the use of these tools and their bearing on the system design is based purely on the discretion of the planner and overall company operating policy.

Figure 17 shows a functional block diagram of the distribution system planning process currently followed by the most of the utilities. This process is repeated for each year of a long-range (15-20 year) planning period. In the development of this diagram, no attempt was made to represent the planning procedure of any specific company but rather to provide an outline of a typical planning process. As the diagram shows, the planning procedure consists of two major activities: load forecasting, distribution system configuration design, substation expansion, and substation site selection.

Configuration Design

Configuration design starts at the customer level. The demand, type, load factor and other customer load characteristics dictate the type of distribution system required. Once customer loads are determined, secondary lines are defined which connect to distribution transformers. The latter provide the reduction from primary voltage to customer-level voltage. The distribution transformer loads are then combined to determine the demands on the primary distribution system. The primary distribution system loads are then assigned to substations that step down from subtransmission voltage. The distribution system loads, in turn, determine the size and location (siting) of the substations as well as the routing and capacity of the associated subtransmission lines. It is clear that each step in this planning process provides input for the steps that follow.

Perhaps what is not clear is that in practice, such a straight-forward procedure may be impossible to follow. A much more common procedure is the following. Upon receiving the relevant load projection data, a system performance analysis is done to determine whether the present system is capable of handling the new load increase with respect to the company's criteria. This analysis, constituting the second stage of the process, requires the use of tools such as a distribution load flow program, a voltage profile and regulation program, etc. The acceptability criteria, representing the company's policies, obligations to the consumers, and additional constraints can include:

1) Service continuity.
2) The maximum allowable peak-load voltage drop to the most remote customer on the secondary.
3) The maximum allowable voltage dip occasioned by the starting of a motor of specified starting current characteristics at the most remote point on the secondary.
4) The maximum allowable peak load.
5) Service reliability.
6) Power losses.

As illustrated in Figure 17, if the results of the performance analysis indicate that the present system is not adequate to meet future demand, then either the present system needs to be expanded by new, relatively minor, system additions or a new substation may need to be built to meet the future demand. If the decision is to expand the present systems with minor additions, then a new additional network configuration is designed and analyzed for adequacy. If the new configuration is found to be inadequate, another is tried, and so on until a satisfactory one is found. The cost of each configuration is calculated. If the cost is found to be too high, or adequate performance cannot be achieved, then the original expand/build decision is reevaluated. If the resulting decision is to build a new substation, a new placement site must be selected. Further, if the purchase price of the selected site is too high, the expand/build decision may need further reevaluation. This process terminates when a satisfactory configuration is attained which provides a solution to existing or future problems at a reasonable cost. Many of the steps in the above procedures can feasibly be done only with the aid of computer programs. A brief catalogue of some of the available tools is presented in Table I.

The program names, in Table I, are accompanied by a general description of the actions performed. No one utility is in possession of all of these programs and generally there has been no attempt to coordinate the input of one program with the output of another. The capabilities summarized in Table I are surely not indicative of the present state of technology with respect to distribution planning tools. There are no doubt other tools available from research institutes, universities and consultants. It does accurately reflect the tools which are presently available to and in use by the nation's utilities and the power industry.

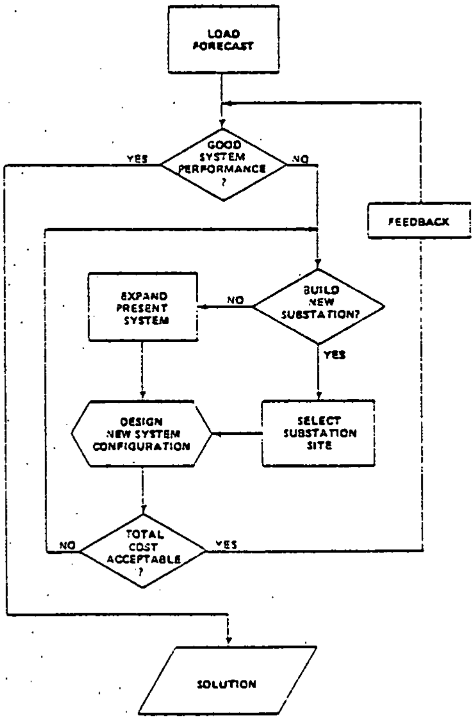


Figure 17. A block diagram of a typical distribution system planning process

TASK B.5: THE IDENTIFICATION AND CATALOGING OF PLANNING OBJECTIVES, CONCEPTS, AND CONSTRAINTS ACCEPTED IN PRACTICE

Introduction

The objective of distribution system planning is to assure that the growing demand for electricity, in terms of increasing growth rates and high load densities, can be satisfied in an optimum way by additional distribution systems, from the secondary conductors through the bulk power substations, which are both technically adequate and reasonably economical. Table 24 presents some additional planning objectives. Distribution system planners must determine the load magnitude and its geographic location. Then the distribution substations must be placed and sized in such a way as to serve the load at maximum cost effectiveness by minimizing feeder losses and construction costs, while considering the constraints of service reliability.

As is well known, distribution system planners have used computers for many years to perform the tedious calculations necessary for system analysis. However, it has only been in the past few years that technology has provided the means for planners to truly take a "systems approach" to the total design and analysis. It is reasonable to assume that the development of such an approach will occupy planners in the 1980's and will significantly contribute to their meeting the challenges previously discussed. In the future, more so than in the past, electric utilities will need a fast and economical planning tool to evaluate the consequences of different proposed alternatives and their impact on the rest of the system to provide the necessary economical, reliable and safe electric energy to consumers. The planner, in the absence of accepted planning techniques, may restate the problem as an attempt to minimize the cost of subtransmission, substations, feeders, laterals, etc., and the cost of losses. In this process, however, he is usually restricted by permissable voltage values, voltage dips, etc., as well as service continuity and reliability. In pursuing these objectives, the planner ultimately has a significant influence on additions to and/or modifications of the subtransmission network, locations and sizes of substations, service areas of substations, locations of breakers and switches, sizes of feeders and laterals, voltage levels and voltage drops in the system, the location of capacitors and voltage regulators, and the loading of transformers and feeders. There are, of course, some factors that need to be considered such as transformer impedances, insulation levels, availability of spare transformers and mobile substations, dispatch of generation, and rates charged to customers.

Table 24. Distribution System Planning Objectives

- Service continuity
- Service reliability
- Quality of service
- Meeting demand
- Cost minimization
- Aesthetics

Impacts of Load Management

In the past, the power utility companies of this nation supplied electrical energy to meet all customer demands at the time the demands are occurred. Recently, however, because of the financial constraints (i.e., high cost of labor, materials, and interest rates), environmental concerns, and the recent shortage (or high cost) of fuels, this basic philosophy has been reexamined and customer load management investigated as an alternative to capacity expansion.

Load management's benefits are system-wide. Alteration of the electrical energy use patterns will affect not only the demands on system generating equipment, but also alter the loading of distribution equipment. The load management may be used to reduce or balance loads on marginal substations and circuits, thus even extending their lives. Therefore, in the future, the implementation of load management policies may drastically affect the distribution of load, in time and in location, on the distribution system, subtransmission system, and on the bulk power system. Since distribution systems have been designed to interface with uncontrolled load patterns, the systems of the future will necessarily be designed somewhat different to benefit from the altered conditions. However, the benefits of load management cannot be fully realized unless the systems planners have the tools required to adequately plan incorporation into the evolving electric energy system. The evolution of the system in response to changing requirements and underchanging constraints is a process involving considerable uncertainty.

Table 25 presents some of the distribution system planning constraints that are encountered in the practice. Further, it would be appropriate to examine what today's trends are likely to portend for the future of the planning process. The central purpose of doing this is to stimulate ideas about how to best meet the increasingly difficult challenges of the near future.

Table 25. Distribution System Planning Constraints

- Economic and financial
- Environmental
- Institutional (company policies)
- Codes, standards, and ordinances
- Geographical and physical boundaries
- Planning deadlines
- Equipment and component standards
- Service continuity
- Service reliability
- Present facilities
- Data availability
- Computational (as a result of the lack of comprehensive planning models)
- Maximum permisable peak-load voltage drop
- Maximum permisable voltage dip
- Maximum permisable peak load
- Power losses
- Feeder getaway, routing, and right-of-ways
- Voltage levels

Economic Factors

There are several economic factors which will have significant effects on distribution planning in the 1980's. The first of these is inflation. Fueled by energy shortages, energy source conversion cost, environmental concerns and government deficits, inflation will continue to be a major factor.

The second important economic factor will be the increasing expense of aquiring capital. As long as inflation continues to decrease the real value of the dollar, attempts will be made by government to reduce the money supply. This in turn will increase the competition for attracting the capital necessary for expansions in distribution systems.

The third factor which must be considered, is increasing difficulty in raising customer rates. This rate increase "inertia" also stems in part from inflation as well as from the results of customers being made more sensitive to rate increases by consumers activist groups.

Demographic Factors

Important demographic developments will affect distribution system planning in the near future. The first of these is a trend which has been dominant over the last fifty years: the movement of the population from the rural areas to the metropolitan areas. The forces which initially drove this migration, economic in nature are still at work. The number of single family farms has continuously declined during this century and there are no visible trends which would reverse this population flow into the larger urban areas. As population leaves the countrysides, population must also leave the smaller towns which depend on the countrysides for economic life. This trend has been a consideration of distribution planners for years and represents no new effect for which account must be taken.

However, the migration from the suburbs to the urban and near urban areas is a new trend attributable to the energy crisis. This trend is just beginning to be visible and it will result in an increase in multifamily dwellings in areas which already have high population densities.

Technological Factors

The final class of factors, which will be important to the distribution system planner, has arisen from technological advances. These advances have been encouraged by the energy crisis. The first of these is the improvement in fuel cell technology. The output power of such devices has risen to the point where in the areas with high population density, large banks of fuel cells could supply significant amount of the total power requirements. Other nonconventional energy sources which might be a part of the total energy grid could appear at the customer level. Among the possible candidates would be solar and wind-driven generators. There is some pressure from consumer groups to force utilities to accept any surplus energy from these sources for use in the total distribution network. If this trend becomes important, it would change drastically the entire nature of the distribution system as is it known today.

FUTURE NATURE OF DISTRIBUTION PLANNING

Predictions about the future methods for distribution planning must necessarily be extrapolations of present methods. Basic algorithms for network analysis have been known for years and are not likely to be improved upon in the near future. However, the superstructure which supports these algorithms and the problem-solving environment used by the system designer is expected to change significantly to take advantage of new methods which technology has made possible. Before giving a detailed discussion of these expected changes, the changing role of distribution planning needs to be examined.

Increasing Importance of Good Planning

For the economic reasons listed above, distribution system will become more expensive to build, expand and modify. Thus it is particularly important that each distribution system design be as cost effective as possible. This means that the system must be optimal from many points of view over the time period from first day of operation to the planning time horizon. In addition to the accurate load growth estimates, components must be phased in and out of the system so as to minimize capital expenditure, meet performance goals and minimize losses.

These requirements need to be met at a time when demographic trends are veering away from what have been their norms for many years in the past and when distribution systems are becoming more complex in design due to the appearance of more active components (e.g., fuel cells) instead of the conventional passive ones.

Cost/Benefit Ratio for Innovation

In the utility industry, the most powerful force shaping the future is that of economics. Therefore, any new innovations are likely to be adopted for their own sake. These innovations will be adopted only if they reduce the cost of some activity or provide something of economic value which previously had been unavailable for comparable costs. In predicting that certain practices or tools will replace current ones, it is necessary that one judge their acceptance on this basis.

The expected innovations which satisfy these criteria are planning tools implemented on a digital computer which deal with distribution systems in network terms. In TASKS A.1 and A.2, a list of currently available such planning tools was given, and one might be tempted to conclude that these tools would be adequate for industry use throughout the 1980's. That this is not likely to be the case may be seen by considering the trends judged to be dominant during this period with those which held sway over the period in which the tools were developed.

New Planning Tools

Tools to be considered fall into two categories: network design tools and network analysis tools. The analysis tools may become more efficient but are not expected to undergo any major changes although the environment in which they are used will change significantly. This environment will be discussed in the next section.

The design tools, however, are expected to show the greatest development since better planning could have a significant impact on the utility industry. The results of this development will show the following characteristics:

(1) Network design will be optimized with respect to many criteria using programming methods of operations research.

(2) Network design will be only one facet of distribution system management directed by human engineers using a computer system designed for such management functions.

(3) So-called network editors will be available for designing trial networks; these designs in digital form will be passed to extensive simulation programs which will determine if the proposed network satisfies performance and load growth criteria.

TASK B.6: THE SELECTION OF PLANNING AND ANALYSIS CONCEPTS AND METHODS WHICH FORM THE CORE OF EFFICIENT DISTRIBUTION PLANNING METHODOLOGIES

Introduction

The assessment established in TASKS A.3, B.1, and B.3 make it clear that a comprehensive mathematical planning model was not available in the literature. The assessment indicated that the major gaps in the literature had to do with the decisions on voltage levels, accomodation of non-linear terms in the objective function, reconductoring decisions, fixed charges on transhipment, location of substations from a given set of prospective sites. and conductor size selection. The literature contained explicit models which address transhipment, incremental substation capacity, variants on site location, fixed charges for substations, and one which purports to track individual transformers. The major gap in the literature was the lack of full exploitation of mixed integer formulations of the distribution networks to analyze these decisions simultaneously.

As a result of this analysis a full scale mixed-integer model for both a single period and dynamic planning horizon has been formulated. Some small scale tests have been run on IBM 370/158 computer to assess the solvability. In TASK B.3 a complete assessment of the solvability problem is presented.

The decision on voltage levels are implicit in these models. By use of the interactive approach, these models can be reformulated by computing the values of parameters at new voltage levels. The voltage level which gives the minimum cost can be selected. Most parameters are unchanged, but of course all power loss curves are affected.

Historical Background

Distribution systems planning requires a complex procedure because: (1) large numbers of variables are involved, (2) the mathematical representation of many requirements and restricting conditions specified by systems configuration is a very difficult task. Some of the approaches used in performing this task in the past include:

1) The alternative policy method, which compares a few alternative policies and selects the best of them.

2) The decomposition approach in which a large problem is divided into several smaller subproblems, and each is solved separately.

3) The linear programming and the integer programming methods which linearize constraint conditions.

4) The dynamic programming approach.

Each method has its own advantages and disadvantages. In long term planning, in particular, a large number of variables is involved and there can exist a number of feasible alternative plans which make the selection of the optimum alternative a very difficult task. The approach used by Lawrence, et al. [1], in their model of "Automated Distribution System Planning", is a good example of the ad hoc models. They included all facilities from the bulkpower transmission lines to the customer's meter. In recent years, there have been a number of advances in the application of mathematical programming to the distribution systems planning models [2-5]. Only recently the state-of-the-art in computer technology has reached the point where the speed and storage capabilities are sufficient to solve a problem of such magnitude as distribution planning, where the interactive nature of the decisions, coupled with the cumbersom amount of data, presents a formidable task even to the most skilled and experienced planner. Juricek, et al. [4], developed a model which employs a load flow analysis to determine future system conditions based on load growth projections and present conditions. A set of possible system modifications composed of combinations of substation expansion or construction and feeder constuction is proposed. This set of modifications is generated by a transportation analysis technique which models the distribution network as a transportation system.

Masud [3] developed a model which included a zero-one integer programming approach to optimize substation transformer capacity, and a linear programming approach to optimize load transfers. The procedure involves first minimizing substation transformer capacities for each year and then optimizing substation transformer capacities for each year and then optimizing load transfers. However, it does not minimize the present value of expansion costs. Recently, Shelton and Mahmoud [6] treated the same task by an interesting approach, combining financial modeling with distribution system expansion decisions, e.g., expanding substations, opening new sites, expanding circuits, and decisions on transformer interchanges.

Each of these models makes approximations, in varying degrees of detail, in the primary feeder network. However, "the greatest level of detail is reached by the Adams and Laughton approach [2] which represents each feeder line segment in terms of capacity and linearized cost, and also considers multiple time periods. They used the model to solve small problems (involving a single substation, 34 small feeder segments and 24 demand locations)" [9]. Later, others, e.g., "Masud [3] and Juricek, et al. [4], developed models that achieved results which involved more realistically sized problems having 10-15 substations. However, in each of these later models, the feeder network was approximated in terms of load transfer capabilities between substation service areas or primary feeder service areas. These approximations, therefore, reduced the capability of such models to include feeder network variable costs directly into the optimization process" [9].

Juricek, et al. [4], and Crawford and Holt [5] recognized the importance of including the load points to represent non-uniform load distribution and feeder cost directly in the optimization process. Crawford and Holt [5] have developed a linear programming approach which utilizes also a transportation algorithm to optimize substation service areas by minimizing the product of demands and distances from substations. The model determines the required loading for each substation one year at a time. Thus, the required substation transformer capacity is determined indirectly. While this technique minimizes distribution feeder losses, it does not necessarily arrive at the optimal expansion plan, since it does not minimize the present value of costs associated with expansion.

A dynamic programming approach to distribution systems planning has been developed by Oldfield and Lang [7] and also by Adams and Laughton [8]. As a compromise between the difficulties due to the large number of variables, plus the complexity of the design process, and the economics to be gained by a search for optimality, Oldfield and Lang have suggested a two-stage planning method; the intention is to provide a method in which to processes of design and optimization are applied consecutively rather than simultaneously. The model, used by Adams and Laughton, determines load transfer schemes and substation installation dates by minimizing the cost of substation transformer losses. Their dynamic programming technique examines all possible combinations of expansion alternatives explicitly for each stage of the study. This approach does not necessarily derive the optimal expansion plan, since minimizing the costs for each year does not necessarily minimize the present value of all costs throughout the study period. Wall, et al. [9], devised a model that contains all the details of the Adams and Laughton model [8] for a single time period, except for the fixed charges on feeder segments. A highly efficient transshipment code is used to solve the model which incorporates several recent signigicant advances, thereby decreasing the time of solution of such problems. They claimed that their model utilizes linear approximations of non-linear cost functions but the explicit equations to achieve that claim were not given.

Distribution Design Techniques Used in Practice

In a given distribution system the existing distribution of demand centers, i.e., demand locations, defines fixed routes along which feeders are located. In order to represent the non-uniform load distribution, a grid system has been devised for a given distribution area, therefore, the centers of each grid represent a demand center. The existing substations, feeders, and feeder routes must be included in the design. Further, any existing or possible interconnecting routes, i.e., tie lines, between existing or potential substation sites must be specified. Thus, the planning engineer, based on the given information, his past experience, and his engineering judgement, selects a radial network configuration by using the available substation sites and feeder routes. Here, the substation may be expanded in some increments where each increment represents either an expansion of an existing substation capacity or the installation of a new substation with an adequate capacity to meet the total system demand. The necessary feeder and transformer sizes are chosen from standardized sets according to thermal rating and voltage regulation restrictions.

Data Preparation

The foundation of the data base of the system for the computer application is the system description in which each significant element of the system is located via grid coordinates, using auxilary files and code numbers as appropriate. Connections between elements are identified and the loads are also assigned by grid coordinates. Therefore, the system model is derived by the standard approach of dividing the geographical distribution area into a system of grids.

The dimension of the grid can be flexible and adjusted according to the need. For rural areas a large grid dimension with a small scale can be used since the loads will be sparsely distributed. For areas where the load density is high, on the other hand, a small grid represents the loads more accurately. A common used grid is a quarter section or 40-acre parcel.

In the system model, each distribution circuit serves a certain number of grids or each grid may be fed by several circuits, depending upon the characteristics of each service area. The present maximum capacity for each circuit is recorded by tabulating the grids served by the circuit, and the total KVA ratings of all transformers in those grids which are fed by that circuit. Table 26 illustrates a sample listing, in which each grid number identifies the grid coordinates of the grid. Thus, the load ratings for each circuit are found by multiplying the total KVA rating by a utilization factor that can be derived from actual measurements of the present loading conditions, i.e., the actual load (in KVA) connected to the circuit divided by the rated KVA capacity of each circuit divided by the rated KVA capacity of each circuit. The future load is then obtained by increasing each of the present loads at the forecasted growth rate. The total present load for each circuit is found by adding the loads connected to that circuit and dividing the total amount of the load by the rated voltage.

TABLE 26. A SAMPLE LISTING

Circuit #	Grid #	Total KVA Rating
172	041067	50
172	041069	175
173	042046	410
173	042048	222
173	043036	65

Feeder Network Data

The feeder network can be represented by means of grid coordinates and the auxiliary data base which provides descriptions of the possible and common line types that can exist in the network. The auxiliary data base also provides information of the demand centers that are connected by the same feeder, the length of each feeder segments connecting the demand centers, and the existing or possible line types that can be used as part of the feeder segment in the operation or design process.

Substation Data

Each existing or potential substation of the system is described by its location, using the grid coordinates, and by its total transformer capacity. In order to simulate and study the system behavior under contingencies, the rated capacity, i.e., its total transformer capacity, of the substation can be reduced to represent the emergency, e.g., loss of feeder segment, loss of number of feeders, the loss of substation transformers, or inadequate substransmission supply, etc.

New Feasible Substation Sites

The substations with their associated feeders make up different substation service areas. The optimal locations of a single substation can be determined using the squared Euclidean distance measure for a given set of service demands by:

$$XS_i = \frac{\sum_{j=1}^{NF} XF_j \cdot DF_j}{\sum_{j=1}^{NF} DF_j} \quad \text{for } j \in S_i$$

$$YS_i = \frac{\sum_{j=1}^{NF} YF_j \cdot DF_j}{\sum_{j=1}^{NF} DF_j} \quad \text{for } j \in S_i$$

where:

$i = 1, 2, \ldots, NS$
$j = 1, 2, \ldots, NF$

XS_i = X - coordinate of substation i

YS_i = Y - coordinate of substation i

XF_j = X - coordinate of feed point, i.e., first customer bus served by the feeder, of feeder service area j.

YF_j = Y - coordinate of feed point of feeder service area j.

DF_j = total demands of feeder service area j

S_i = substation i

NF = number of substations

NS = number of feeders

For a rectilinear distance measure XS_i is a point such that fifty percent or few load values are located on X coordinates which are greater than or equal to XS_i and fifty percent or fewer load values are located on X coordinates which are less than or equal to XS_i. A similar rule applies for the YS_i.

MATHEMATICAL MODEL

The distribution system expansion costs can be categorized into two groups: (1) the substation expansion costs, and (2) the feeder expansion costs. Let the number of existing substations be NEW and NPS be the number of possible sites on which to build substations. Therefore, an electric distribution system can effectively be modeled as a mixed integer programming problem with the substations as sources and the loads on the feeders as demands. Here, the objective function is the minimization of the present worth of the capital, i.e., fixed cost of the distribution system expansion and the present worth of the variable costs associated with the power losses, subject to restrictions which relate substation transformer capacities and feeder ratings to system load projections.

In order to set up the model the following notation is introduced:

For Parameters

SFC_i = present worth of fixed cost of constructing substation i,

FFC_{ij} = present worth of constructing a feeder from substation i to demand center j,

TFC_{ij} = present worth of constructing a tie-feeder from substation i to substation j,

DFC_{ij} = present worth of fixed cost of constructing a feeder from demand center i to demand center j,

FRC_{ij} = present worth of reconductoring feeder from substation i to demand center j,

FBC_i = present worth of fixed cost of adding a bay at substation i (if facility already exists, the fixed charge is zero),

C_{ik} = present worth of fixed cost of adding incremental capacity k to substation i,

(a_s, b_s) = coordinates of points on the envelope curve of the feeder variable cost curves,

(a_{ijs}, b_{ijs}) = coordinates of a point s on the envelope curve of the substation variable cost for route ij,

(a_{ijs}, b_{ijs}) = coordinates of a point s on power loss curve for a specific conductor placed in route ij,

NF = number of feeders emanating from the initial size of a substation,

NF_B = number of feeders per bay added to a substation,

NB_{max} = maximum number of bays that can be added to a given substation,

NES = number of existing substations,

NPS = number of potential substation sites,

NS = number of existing substations and potential substation sites thus,

NS = NES + NPS

m = total number of existing demand centers,

DF_j = total demand of demand center j,

SIC_i = initial capacity of substation i,

Δ_k = incremental capacity size k that can be added to a given substation,

U_{ij} = rated capacity of a feeder connecting origin i to destination j,

G = total number of points required to approximate a given nonlinear curve.

For Power Flow Variables

X_{ij} = quantity transported from substation i to demand center j,

Y_{ij} = quantity transported from substation i to substation j,

Z_{ij} = quantity transported from demand center i to demand center j,

$(RX)_{ij}$ = quantity transported from substation i to demand center j over a reconductored feeder.

For Decision Variables

δ_i = binary integer variable which denotes the decision to select or not to select site i,

δ_i = 0, if a substation does not exist at the site i or will not be built,

δ_i = 1, if a substation is to be built at the site or already exists,

δ_{ij} = 0, if a feeder does not exist between substation i and demand center j,

δ_{ij} = 1, if a feeder does exist or is to be built between substation i and demand center j,

γ_{ij} = 0, if a tie-feeder does not exist between substation i and substation j,

γ_{ij} = 1, if a tie-feeder does exist or is to be built between substation i and substation j,

θ_{ij} = 0, if a feeder does not exist between demand center i and demand center j,

θ_{ij} = 1, if a feeder does exist or is to be built between demand center i and demand center j,

$(R\delta)_{ij}$ = 0, if reconductoring of tie-feeder between substation i and substation j will not be done,

$(R\delta)_{ij}$ = 1, if reconductoring of tie-feeder between substation j will be done,

μ_i = number of bays to be added to the substation at site i,

α_{ik} = 1, if incremental capacity k is to be added to substation i,

α_{ik} = 0, otherwise.

For Linearization Variables

The following variables are used to approximate a nonlinear curve with straight line segments.

t_{ijs} = representation variables for $f(X_{ij})$,

$(ty)_{ijs}$ = representation variables for $f(Y_{ij})$,

$(tz)_{ijs}$ = representation variables for $f(Z_{ij})$,

$(tR)_{ijs}$ = representation variables for $f[(RX)_{ij}]$,

$(\beta X)_{ijs}$ = decision variable to force selection of at most two t_{ijs} to be nonzero,

$(\beta Y)_{ijs}$ = decision variable to force selection of at most two $(ty)_{ijs}$ to be nonzero,

$(\beta Z)_{ijs}$ = decision variable to force selection of at most two $(tz)_{ijs}$ to be nonzero,

$(\beta R)_{ijs}$ = decision variable to force selection of at most two $(tR)_{ijs}$ to be nonzero.

The optimization problem includes choosing the sites to locate substations; determining the optimum amount of incremental capacity to add to existing and/or newly built substation; determining the optimum number of feeders emanating from substations; finding the optimum number of bays required to support the number of feeders chosen; connecting the substations through tie feeders; selecting the connections between demand centers; selecting the connections between substations and demands; the optimum conductor size of each connecting feeder; and the feeder, between substations and demand centers, which should be reconductored in such a way as to minimize the present value of costs and meet the forecasted demands. Two concepts are necessary to introduce which are not fully defined in a mathematical programming format in the previous power system planning literature; (1), decision variables and (2) power loss envelope curves.

A decision variable is a variable whose values are restricted to either zero or one. The one represents a yes decision and the zero represents a no decision or status quo. For example, suppose that the cost of constructing a number of substations at certain sites, out of eight potential substation sites, needs to be represented, then the proposed cost would be equal to

$$\sum_{i=1}^{8} SFC_i \cdot \delta_i$$

If the decision is to build two substations at the potential substation sites three and seven, then

$$\delta_3 = \delta_7 = 1$$

and

$$\delta_1 = \delta_2 = \delta_4 = \delta_5 = \delta_6 = \delta_8 = 0$$

and

$$\begin{aligned}\text{Cost} &= SFC_1(0) + SFC_2(0) + SFC_3(1) + \ldots \\ &\quad + SFC_7(1) + SFC_8(0) \\ &= SFC_3 + SFC_7\end{aligned}$$

Thus, the cost of any combination of decisions can be represented by the general summation. However, this cost function is subject to some logical constraints dictated by the technology of the problem. For example, no feeder emanating can be built from a substation if the substation does not exist. This can be assured by a constraint of the form

$$\sum_{j=1}^{m} \delta_{ij} \le NF \cdot \delta_i \text{ for every substation i.}$$

This works in the following way. If substation i does not exist, then $\delta_i = 0$. Since δ_{ij} can be only zero or one for each j. The equation

$$\sum_{j=1}^{m} \delta_{ij} \le 0$$

implies that each $\delta_{ij} = 0$. Since zero represents the decision not to act, no feeder is to be built. If $\delta_i = 1$, then a substation exists or will be built and

$$\sum_{i=1}^{m} \delta_{ij} \le NF$$

Then up to the number of NF of the δ_{ij} can be one, and up to the number of NF demands can be served from substation i.

The cost of power losses in feeders varies with the conductor size and the power transported. For a given load there is a conductor size which gives the minimum total cost of power losses, lost feeder capacity due to power losses and investment. Figure 18 illustrates the concept.

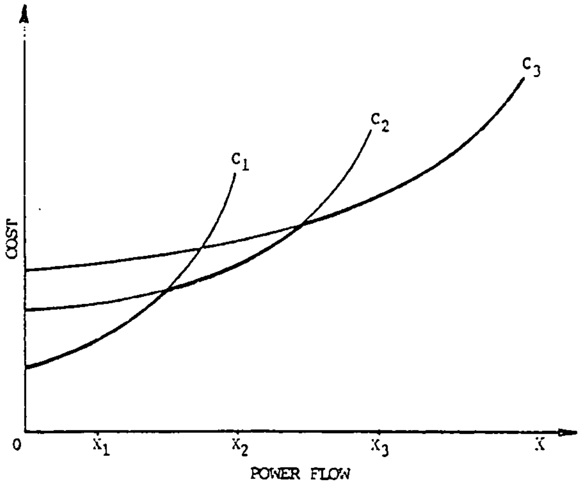


Figure 18. Envelope curve representing minimum cost of transporting power X

In the figure, each of the three curves represent the summation of the cost of investment in an installed feeder, the cost of lost energy due to I^2R losses in the feeder, and the cost of demand lost (i.e., the cost of lost capacity) due to the I^2R losses for a given conductor size [10]. If the power transported is X_1, then the least expensive conductor size is the conductor one, given by C_1 curve. Similarly, the cost of the conductors two and three, given by curves C_2 and C_3, are minimum if the transported power are X_2 and X_3, respectively. It is assumed that the conductor giving optimum cost will be selected in installing a new feeder. Therefore, the variable cost of conductor can be represented by the function.

$$f(X) = \min_i f_i(X)$$

where

$f_i(X)$ = installed feeder investment cost plus lost energy cost due to power losses in feeder plus lost capacity cost of feeder due to the power losses in the feeder for a given conductor size i.

The function f(X) gives the envelope curve which is represented by the heavy line in Figure 19. It is a nonlinear and nonconvex curve. This envelope curve can be approximated by straight line segments and represented by a continuous and zero-one variables. The method by which this can be done is given in the following section.

For a feeder that already exists, the cost of power losses is given by the single curve defined by the given conductor size. This curve is a convex curve and can be represented by straight line segments as shown in Figure 7.

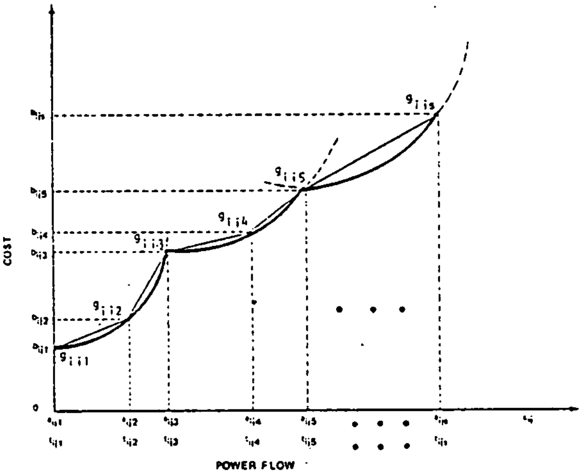


Figure 19. Linear approximation of cost curve $f(X_{ij})$ which represents the present worth of cost of investment, lost energy in the feeder and cost of lost feeder capacity as a function of power flow over route (i,j)

Piecewise Linearization of Nonlinear Cost Functions

In Figure 19, a nonconvex and nonlinear curve has been approximated by line segments between six chosen points on the curve. The curve represents the present worth of installed feeder investment cost, lost energy cost due to power losses in feeder, and lost capacity cost of feeder due to the power losses in the feeder for the optimim conductor size for a given load transported.

The following equations are given to demonstrate the concept of how a piecewise linear approximation of a nonlinear function is done, as shown in Figure 19.

The approximation rests on the idea that if

$$a_{ij1} \le X_{ij} \le a_{ij2}$$

then t_{ij1}, t_{ij2}, the surrogate variables used to represent $f(\cdot)$, can be chosen such that the value of

$$X_{ij} = a_{ij1} \cdot t_{ij1} + a_{ij2} \cdot t_{ij2}$$

where t_{ij1} and t_{ij2} have to satisfy $t_{ij1} + t_{ij2} = 1$ and $t_{ij1}, t_{ij2} \ge 0$. Note the fact that if $t_{ij1} + t_{ij2} > 1$, one cannot find the surrogate variables t_{ij1}, t_{ij2} such that the value of

$$X_{ij} = a_{ij1} \cdot t_{ij1} + a_{ij2} \cdot t_{ij2}$$

lies between a_{ij1} and a_{ij2}. Now consider the points (a_{ij1}, b_{ij1}) and (a_{ij2}, b_{ij2}). If

$$a_{ij1} \le X_{ij} \le a_{ij2} \text{ and } t_{ij1} + t_{ij2} = 1$$

where

$$t_{ij1}, t_{ij2} \ge 0$$

then

$$X_{ij} = a_{ij1} \cdot t_{ij1} + a_{ij2} \cdot t_{ij2} \quad \text{is}$$

such that the point $(X_{ij}, b_{ij1} \cdot t_{ij1} + b_{ij2} \cdot t_{ij2})$ lies on the line segment connecting (a_{ij1}, b_{ij1}) and (a_{ij2}, b_{ij2}).

If a curve is approximated by several points, X_{ij} would lie between a specific pair of coordinates a_{ijs-1} and a_{ijs}. Therefore, the t_{ijs}, corresponding to this pair, must be greater than or equal to zero and all other t_{ijs} must be forced to be zero. For example, consider the following equations where the β_{ijs} carry out the aforementioned function.

$$X_{ij} = \sum_{s=1}^{6} a_{ijs} \cdot t_{ijs} \tag{1}$$

$$\sum_{s=1}^{5} \beta_{ijs} = 1 \tag{2}$$

$$\sum_{s=1}^{6} t_{ijs} = 1 \tag{3}$$

$$\text{Cost} = \sum_{s=1}^{6} b_{ijs} \cdot t_{ijs} \tag{4}$$

$$t_{ij1} \le \beta_{ij1} \tag{5}$$

$$t_{ij2} \le \beta_{ij1} + \beta_{ij2} \tag{6}$$

$$t_{ij3} \le \beta_{ij2} + \beta_{ij3} \tag{7}$$

$$t_{ij4} \le \beta_{ij3} + \beta_{ij4} \tag{8}$$

$$t_{ij5} \le \beta_{ij4} + \beta_{ij5} \tag{9}$$

$$t_{ij6} \le \beta_{ij5} \tag{10}$$

where

$t_{ijs} \ge 0$ and β_{ijs} = either zero or one.

Here, it is desired to find

$$\text{Cost} \cong f(x_{ij}) \tag{11}$$

In order to find the cost, using equation (4), set

$$\beta_{ij2} = 1$$

then equation (2) would force that

$$\beta_{ij1} = \beta_{ij3} = \beta_{ij4} = \beta_{ij5} = 0$$

Thus, equations (5-10) would allow only t_{ij2} and t_{ij3} to be non zero. Hence, equations (1) and (3) can be reduced to

$$x_{ij} = a_{ij2} \cdot t_{ij2} + a_{ij3} \cdot t_{ij3}$$

and $1 = t_{ij2} + t_{ij3}$

these last two equations can be solved for t_{ij2} and t_{ij3} and the approximate cost value can be found from equation (4).

Therefore, in summary

$$a_{ijg-1} \le X_{ij} \le a_{ijg} \qquad (12)$$

then

$$X_{ij} = t_{ijg-1} \cdot a_{ijg-1} + t_{ijg} \cdot a_{ijg} \qquad (13)$$

and

$$f(X_{ij}) \cong t_{ijg-1} \cdot b_{ijg-1} + t_{ijg} \cdot b_{ijg} \qquad (14)$$

Thus, the double summation of $f(X_{ij})$, which would be the present worth of cost of investment, lost energy in the feeder and cost of lost feeder capacity as a function of power flow over route (i,j), can be represented by the following term, that is

$$\sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{G} b_{ijs} \cdot t_{ijs} \qquad (15)$$

and the total of the variable costs of substations can be represented by

$$\sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{G} d_{ijs} \cdot t_{ijs} \qquad (16)$$

a_{ij1} must equal zero in order that X_{ij} equal zero can be represented.

BASE MODEL

The following is the complete mathematical model which, when solved, gives the optimum decision in each of the aforementioned categories. The definition of each term in the objective function and the definition of each constraint are given following the model.

$$\text{Cost} = \underset{(A)}{\sum_{i=1}^{NS} SFC_i \cdot \delta_i} + \underset{(B)}{\sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{g} t_{ijs} \cdot d_{ijs}}$$

$$+ \underset{(C)}{\sum_{\substack{i=1 \\ i,j \in NE}}^{NS} \sum_{j=1}^{m} FFC_{ij} \cdot \delta_{ij}} + \underset{(D)}{\sum_{\substack{i=1 \\ i,j \in NE}}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{g} t_{ijs} \cdot b_{ijs}}$$

$$+ \underset{(E)}{\sum_{i=1}^{NS} \sum_{j=1}^{NS} TFC_{ij} \cdot Y_{ij}} + \underset{(F)}{\sum_{i=1}^{NS} \sum_{j=1}^{NS} \sum_{s=1}^{g} (ty)_{ijs} \cdot b_{ijs}}$$

$$+ \underset{(G)}{\sum_{i=1}^{NS} \sum_{k=1}^{R} C_{ik} \cdot \alpha_{ik}} + \underset{(H)}{\sum_{\substack{i=1 \\ i \ne j}}^{m} \sum_{j=1}^{m} \Theta_{ij} \cdot DFC_{ij}}$$

$$+ \underset{(I)}{\sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{s=1}^{g} (tz)_{ijs} \cdot b_{ijs}} + \underset{(J)}{\sum_{\substack{i=1 \\ i,j \in REC}}^{NS} \sum_{j=1}^{m} (FRC)_{ij} \cdot (R\delta)_{ij}}$$

$$+ \underset{(K)}{\sum_{\substack{i=1 \\ i,j \in REC}}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{g} (tR)_{ijs} \cdot b_{ijs}}$$

$$+ \underset{(L)}{\sum_{\substack{i=1 \\ i,j \in E}}^{NS} \sum_{j=1}^{m} \sum_{s=1}^{g} t_{ijs} \cdot b_{ijs}} + \underset{(M)}{\sum_{i=1}^{NS} \mu_i \cdot FBC_i} \qquad (17)$$

where

E = {i,j | route i,j has a feeder in place}

REC = {i,j | route i,j has a potential cost saving by reconductoring}

NE = {i,j | route i,j has no feeder in place}

E ∪ NE = {i,j | set of all routes i,j}

Any fixed cost term in the above objective function is zero if facility is already in place. The objective function is subject to the following constraints:

$$\sum_{i=1}^{NS} X_{ij} + \sum_{\substack{i=1 \\ i \ne j}}^{m} Z_{ij} - \sum_{\substack{i=1 \\ j \ne i}}^{m} Z_{ji} + \sum_{\substack{i=1 \\ i,j \in REC}}^{NS} (RX)_{ij} \ge DF_j \quad \forall_j \qquad (18)$$

where

$j=1,\dots,m$

$$\delta_{ij} = 1 - (R\delta)_{ij} \qquad \forall_{ij} \qquad (19)$$

where

$i,j \in REC$

$$(RX)_{ij} \le (R\delta)_{ij} \cdot U_{ij} \qquad \forall_{i,j} \in REC \qquad (20)$$

$$\sum_{j=1}^{m} X_{ij} \le SIC_i \cdot \delta_i + \sum_{\substack{j=1 \\ i \ne j}}^{NS} y_{ji} \qquad \forall_i \qquad (21)$$

$$- \sum_{\substack{j=1 \\ i \ne j}}^{NS} Y_{ij} + \sum_{k=1}^{R} \Delta_k \cdot \alpha_{ik}$$

where

$i=1,\dots,NS$

$$X_{ij} \le U_{ij} \cdot \delta_{ij} \qquad \forall_{ij} \qquad (22)$$

where

$i=1,\dots,NS$
$j=1,\dots,m$

$$\sum_{j=1}^{m} \delta_{ij} \le NF\ \delta_i + NF_B \cdot \mu_i \qquad \forall_i \qquad (23)$$

where

i=1,...,NS

$$\sum_{j=1}^{NS} \delta_i \le NS \qquad (24)$$

$$Y_{ij} \le U_{ij} \cdot \gamma_{ij} \qquad \forall_{ij} \qquad (25)$$

where

i=1,...,NS
j=1,...,NS i≠j

$$\gamma_{ij} \le \tfrac{1}{2}(\delta_i + \delta_j) \qquad \forall_{ij} \qquad (26)$$

where

i=1,...,NS
j=1,...,NS i≠j

$$\sum_{k=1}^{R} \alpha_{ik} \le \delta_i \qquad \forall_i \qquad (27)$$

where

i=1,...,NS

$$Z_{ij} \le U_{ij} \cdot \theta_{ij} \qquad \forall_{ij} \qquad (28)$$

where

i=1,...,m
j=1,...,m
i≠j

$$\mu_i \le \delta_i \cdot NB_{max} \qquad \forall_i \qquad (29)$$

where

i=1,...,NS

$$Z_{ij} = \sum_{s=1}^{G} (tz)a_{ijs}$$

$$\sum_{s=1}^{G} (tz)_{ijs} = 1$$

$$\sum_{s=1}^{G=1} (\beta z)_{ijs} = 1$$

$$(tz)_{ij1} \le (\beta z)_{ij1} \qquad \forall_{ih} \qquad (30)$$

$$(tz)_{ij2} \le (\beta z)_{ij1} + (\beta z)_{ij2}$$

$$\vdots$$

$$(tz)_{ijg} \le (\beta z)_{ijg-1} + (\beta z)_{ijg}$$

$$\vdots$$

$$(tz)_{ijG} \le (\beta z)_{ijG-1}$$

where

i=1,...,m
j=1,...m

$$(RX)_{ij} = \sum_{s=1}^{G} (tR)_{ijs} \cdot a_{ijs}$$

$$\sum_{s=1}^{G} (tR)_{ijs} = 1$$

$$\sum_{s=1}^{G-1} (\beta R)_{ijs} = 1 \qquad \forall_{ij} \qquad (31)$$

$$(tR)_{ij1} \le (\beta R)_{ij1}$$

$$\vdots$$

$$(tR)_{ijG} \le (\beta R)_{ij\ -1} + (\beta R)_{ijG}$$

$$(tR)_{ijG} \le (\beta R)_{ijG-1}$$

where

ij ∈ REC

$$X_{ij} = \sum_{s=1}^{G} t_{ijs} \cdot a_{ijs}$$

$$\sum_{s=1}^{G} t_{ijs} = 1$$

$$\sum_{s=1}^{G-1} (\beta X)_{ijs} = 1$$

$$t_{ij1} \le (\beta X)_{ij1} \qquad \forall_{ij} \qquad (32)$$

$$t_{ij2} \le (\beta X)_{ij1} + (\beta X)_{ij2}$$

$$\vdots$$

$$\le (\beta X)_{ijG-1} + (\beta X)_{ijG}$$

$$\vdots$$

$$t_{ijG} \le (\beta X)_{ijG}$$

where

i=1,...,NS
j=1,...,m

$$Y_{ij} = \sum_{s=1}^{G} (ty)_{ijs} \cdot a_{ijs}$$

$$\sum_{s=1}^{G} (ty)_{ijs} \cdot a_{ijs} = 1$$

$$\sum_{s-1}^{G-1} (\beta y)_{ijs} = 1 \qquad \forall_{ij} \qquad (33)$$

$$(ty)_{ij1} \le (\beta y)_{ij1}$$

$$(ty)_{ij2} \le (\beta y)_{ij1} + (\beta y)_{ij2}$$

$$\vdots$$

$$(ty)_{ijG} \le (\beta y)_{ijG-1} + (\beta y)_{ijG}$$

$$(ty)_{ijG} \le (\beta y)_{ijG-1}$$

where

i=1,...,NS
j=1,...,NS
i≠j

$X_{ij} \ge 0$; $t_{ijs} \ge 0$; $(ty)_{ijs} \ge 0$; $(tz)_{ijs} \ge 0$;

$(tR)_{ijs} \geq 0; \quad z_{ij} \geq 0; \quad y_{ij} \geq 0; \quad (RX)_{ij} \geq 0;$

$(\beta X)_{ijs} = 0,1; \quad (\beta y)_{ijs} = 0,1; \quad (\beta z)_{ijs} = 0,1,;$

$(\beta R)_{ijs} = 0,1; \quad \delta_1 = 0,1; \quad \delta_{ij} = 0,1;$

$\Theta_{ij} = 0,1; \quad \alpha_{ik} = 0,1; \quad \gamma_{ij} = 0,1;$

$u_i = 1,2,3,\ldots NB_{max}$

The complete procedure is outlined in flow-chart form in Figure 20. The inequalities can be converted to equalities by addition of slack variables, however, the program used does this automatically.

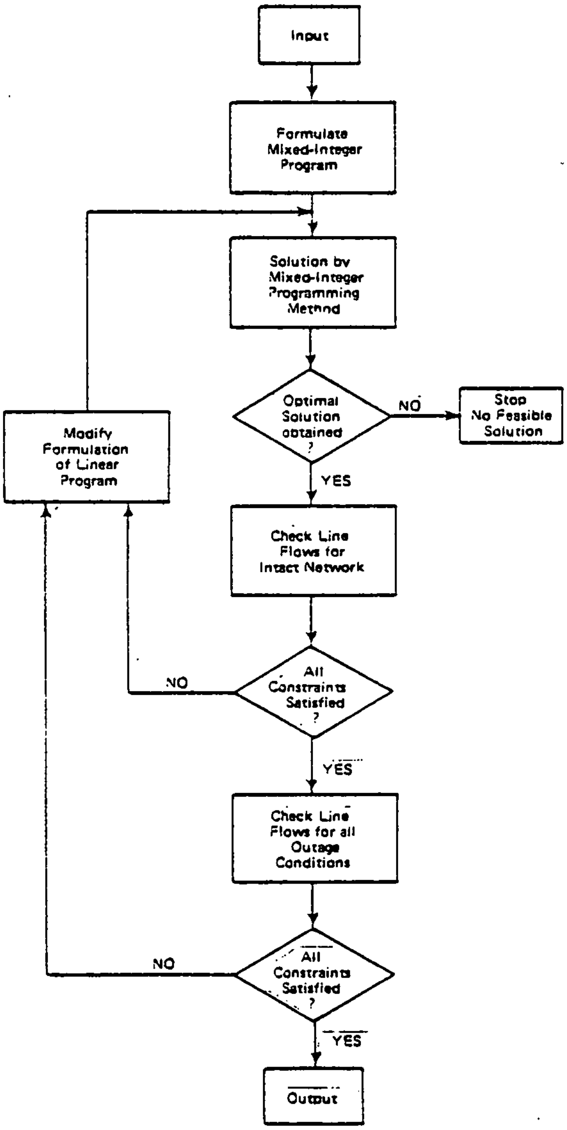


Figure 20. Flow diagram of mixed-integer programming

The following are the explanations of the terms in the objective function.

Term A.: gives the fixed charges for constructing a substation on site i.

Term B: gives the cost of power losses in substation equipment, represented by a piecewise linear approximation of power transported over route (i,j).

Term C: gives the fixed charges of an installed feeder on the route from substation i to demand center j.

Term D: gives the cost of power losses when power quantity X_{ij} is transported, represented by a piecewise linear approximation of the nonlinear cost curve.

Term E: gives the fixed charges of an installed tie-feeder between substation.

Term F: gives the cost of power losses in tie-feeders between substations as a result of transshipment, represented by a piecewise linear approximation of the nonlinear cost curve.

Term G: gives the cost of adding incremental capacity at the various substations.

Term H: gives the fixed charges for installed feeders between demand centers.

Term I: gives the cost of power losses in transshipment between demand centers with piecewise linear representation of the nonlinear cost curve.

Term J: gives the fixed charges of reconductoring a feeder over route (i,j).

Term K: gives the cost of power losses in feeders over reconductored route (i,j), represented by a piecewise linear approximation of the nonlinear cost curves.

Term L: gives the cost of power losses over a route (i,j) with a feeder in place, represented by a piecewise linear approximation of the nonlinear cost curves.

Term M: gives the fixed charges for adding bays to a substation.

The following are the explanations of the constraints used in the model.

Constraint (18): assures that each demand is met.

Constraint (19): is a logic restriction assuring that decision variable for power flow is zero if reconductoring occurs and vice versa.

Constraint (20): assures that there is no power flow over reconductored route if reconductor decision variable is zero.

Constraint (21): assures that the total quantity transported from substations to destinations is less than the substation capacity plus power transfered from other substations.

Constraint (22): stops the power flow if feeder decision variable is zero.

Constraint (23): limits the number of feeders built to the amounts specified by the number of bays available.

Constraint (24): limits the number of sites available.

Constraint (25): assures that the power flow between substations is zero if no feeder exists between substations.

Constraint (26): stops a feeder from being built between substations unless both substations exist (i.e., $\delta_i = 1$ and $\delta_j = 1$).

Constraint (27): assures that incremental capacity cannot be added unless the substation exist (i.e., $\delta_i = 1$).

Constraint (28): assures that there is no power flow between demand centers unless the connecting feeder exists (i.e., $\theta_{ij} = 1$).

Constraint (29): assures that no bay can be added unless the substation exists (i.e., $\delta_i = 1$).

Constraint (30): computes piecewise linearization variables for $f(Z_{ij})$.

Constraint (31): computes piecewise linearization variables for $f[(RX)_{ij}]$.

Constraint (32): computes piecewise linearization variables for $f(Y_{ij})$.

In order to demonstrate how the constraints are formed consider the restriction (18). This constraint simply states the fact that the sum of the power flows from all substations, that is

$$\sum_{i=1}^{NS} X_{ij}$$

plus the power flow received from all other demand centers, that is

$$\sum_{\substack{i=1 \\ i \neq j}}^{m} Z_{ij}$$

minus the power flow sent to other demand centers. that is

$$\sum_{\substack{i=1 \\ j \neq 1}}^{m} Z_{ji}$$

plus the power flow transmitted through all reconductored feeders, that is

$$\sum_{\substack{i=1 \\ i,j \in REC}}^{NS} (RX)_{ij}$$

must be greater than or equal to the demand of the demand center j, i.e., DF_j.

Also, consider the restriction (19) which is a "cut off" constraint. That is, if

$$(RX)_{ij} = 1 \qquad \text{then } \delta_{ij} = 0$$

and by constraint (22) X_{ij} becomes zero. If

$$\delta_{ij} = 1$$

then the same result is obtained using restriction (20).

Further, consider the restriction (27). If the substation does not exist than the substation decision variable is

$$\delta_i = 0.$$

Then

$$\sum_{k=1}^{R} \alpha_{ik} \leq 0$$

that is

$$\alpha_{ik} = 0 \qquad \forall_k$$

and no incremental capacity is added. If, for example,

$$\alpha_{i3} = 1$$

then term (G) in the objective function becomes equal to C_{i3} because all other terms in the summation have a zero factor.

Applications of the Model

To test the solvability of the model a problem was devised and solved. The problem had three different cases that each included power transshipment, feeder routing and substation site selection using linear approximation of nonlinear cost curves, and substation incremental capacity. The data used in determining parameters of the model were provided by the Oklahoma Gas and Electric Company. Each case had three substation sites and eight demand centers. Table 27 gives the description of each individual case. The variable feeder costs were based on the rectilinear distances between points. Using the "sunk cost " concept of engineering economy, fixed costs of existing facilities were assumed to be zero. In estimating future fixed and variable costs a carrying charge rate of 21.88 percent was used and also an inflation rate of ten percent was considered.

Table 27. Case Descriptions

Case	Initial Conditions	No. of Integer Variables	No. of Continuous Variables	No of Constraints
1	no existing feeders or substations	113	197	111
2	one existing substation and two existing feeders	113	197	111
3	one existing substation and three existing feeders	113	197	111

The cases were run on an IBM 370/158 computer using the MPSX mathematical programming system. The solution results are shown in Figures 21,23, and 24. The summary of the computer output data of case number one is shown in Figure 22. The cost of the computer runs, the CPU time, and the value of the optimum solution for each case are given in Table 28.

Table 28. Computational Results

Case	Cost of Computer Run (in $)	CPU Time (in Minutes)	Optimum Solution Value (in $)
1	284	41.5	1,675,000
2	121	17.67	1,471,700
3	115	17.05	1,418,750

In case number one, the computer run time exceeded the allocated CPU time without achieving an optimum solution. Therefore, case number one was restarted with a planner's solution and an optimal solution was achieved. The optimal solution saved ninety-eight thousand dollars in comparison to the planner's solution. The optimal solution was found in twenty-nine minutes, however, twelve additional minutes were required to prove optimality. Previous problems of this size reported in the literature were not solved even after a five hour CPU time despite the fact that a faster computer was utilized.

Sensitivity analysis of case number one were performed. The results of three different sensitivity analyses for case number one are shown in Figures 25-27. Figure 25 shows the results of changing incremental capacity step sizes to two thirds of their original values. Figure 26 shows the results of changing the number of feeders emanating from substation two from four to two. Figure 27 shows the results of changing the demand of demand center two from 2.4 MVA to 6 MVA. Note the difference in feeder connections of Figures 25 and 26 in comparison with Figure 21. However, the change in demand did not force a change in feeder connections, as shown in Figure 27. This study indicates the fact that incorporating all the major distribution planning decisions into a comprehensive planning model is now computationly feasible and some major cost savings can be achieved.

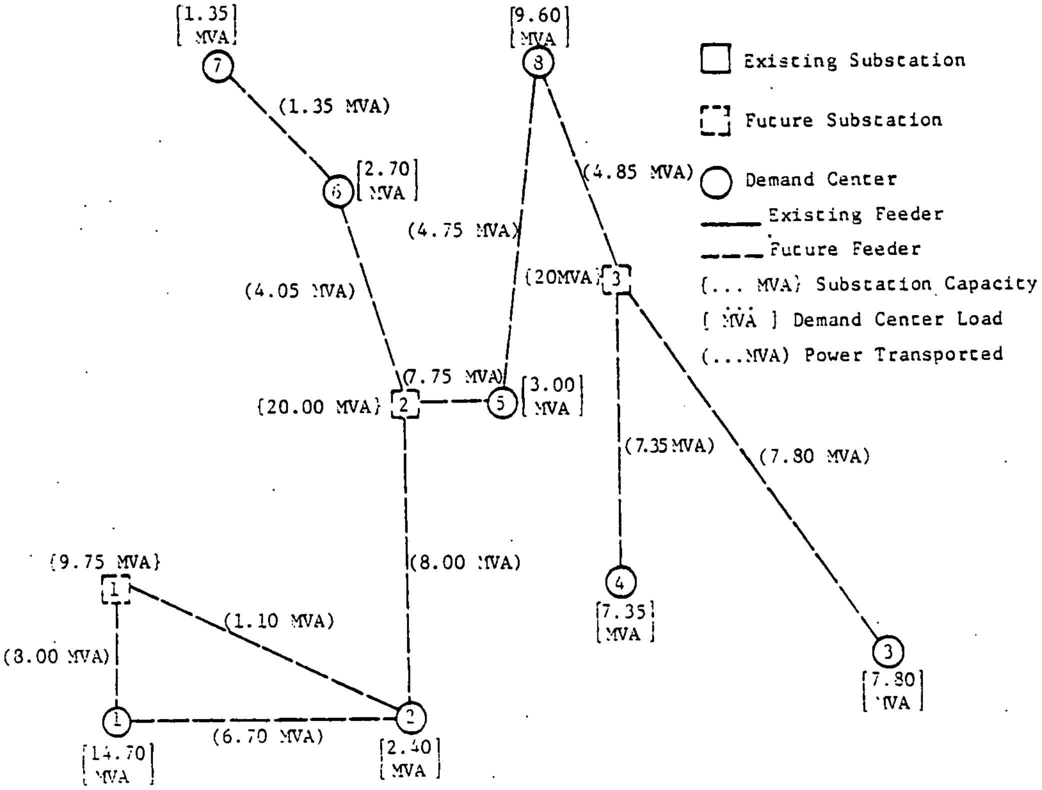
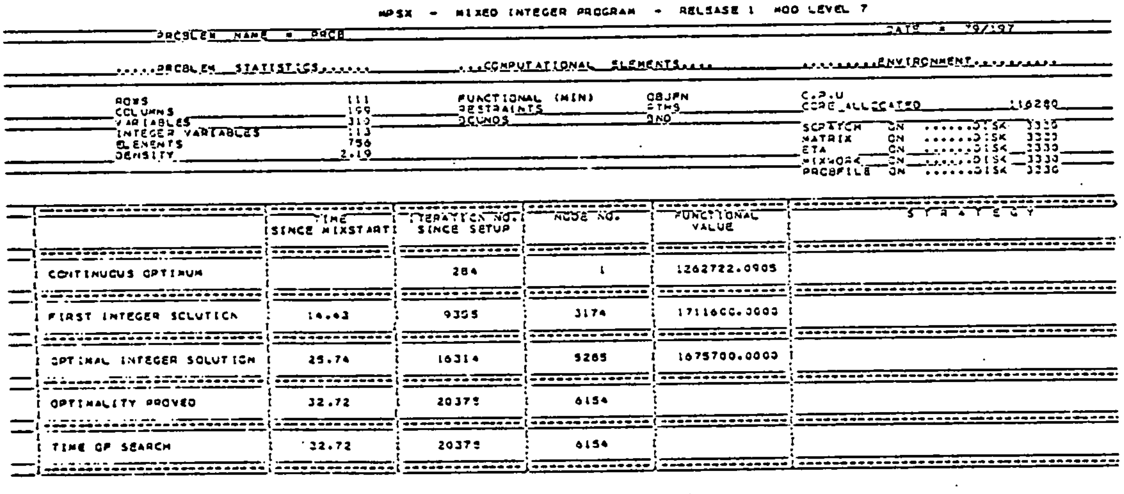


Figure 21. Case #1 solution results



```
                              MPSX  -  MIXED INTEGER PROGRAM  -  RELEASE 1  MOD LEVEL 7
            PROBLEM  NAME  =  PROB                                                          DATE  =  79/197

         .....PROBLEM  STATISTICS.......          ...COMPUTATIONAL  ELEMENTS....          .........ENVIRONMENT.........

            ROWS                        111       FUNCTIONAL (MIN)      OBJFN             C.P.U
            COLUMNS                     199       RESTRAINTS            RHS               CORE_ALLOCATED            116280
            VARIABLES                   310       BOUNDS                BND
            INTEGER VARIABLES           113                                               SCRATCH    ON   ......DISK   3330
            ELEMENTS                    756                                               MATRIX     ON   ......DISK   3330
            DENSITY                    2.19                                               ETA        ON   ......DISK   3330
                                                                                          MIXWORK    ON   ......DISK   3330
                                                                                          PROBFILE   ON   ......DISK   3330

 ---------------------------------------------------------------------------------------------------------------------
 I                           I      TIME     I ITERATION NO.I  NODE NO.   I  FUNCTIONAL   I       S T R A T E G Y
 I                           ISINCE MIXSTARTI SINCE SETUP  I             I     VALUE     I
 ---------------------------------------------------------------------------------------------------------------------
 I CONTINUOUS OPTIMUM        I               I      284     I       1     I 1262722.0905  I
 ---------------------------------------------------------------------------------------------------------------------
 I FIRST INTEGER SOLUTION    I     14.43     I     9355     I    3174     I 1711600.0000  I
 ---------------------------------------------------------------------------------------------------------------------
 I OPTIMAL INTEGER SOLUTION  I     25.74     I    16314     I    5265     I 1675700.0000  I
 ---------------------------------------------------------------------------------------------------------------------
 I OPTIMALITY PROVED         I     32.72     I    20375     I    6154     I               I
 ---------------------------------------------------------------------------------------------------------------------
 I TIME OF SEARCH            I     32.72     I    20375     I    6154     I               I
 ---------------------------------------------------------------------------------------------------------------------

            NUMBER_OF_INTEGER_VARIABLES_NOT_INTEGER_AT_CONTINUOUS_OPTIMUM  =  14
            NUMBER OF INTEGER SOLUTIONS FOUND  =  3
            BRANCHES_ABANDONED_WHILE_COMPUTING  =  107
```

Figure 22. The summary of the input data of the case #1 solution

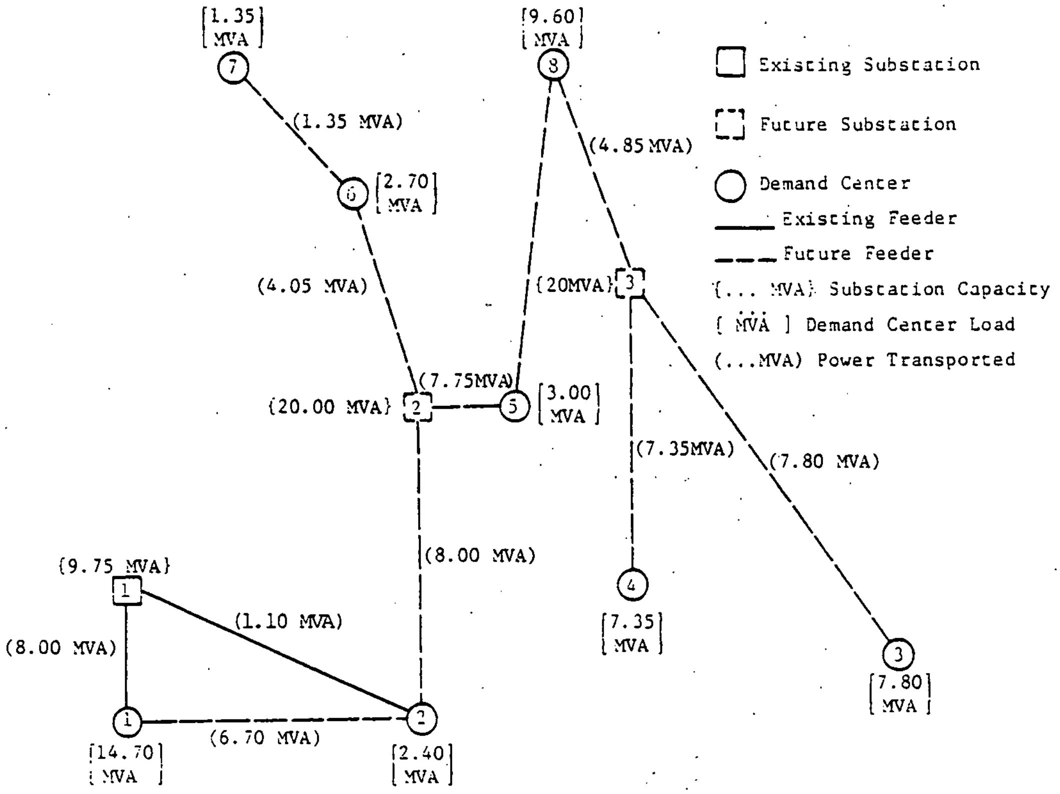


Figure 23. Case #2 solution results

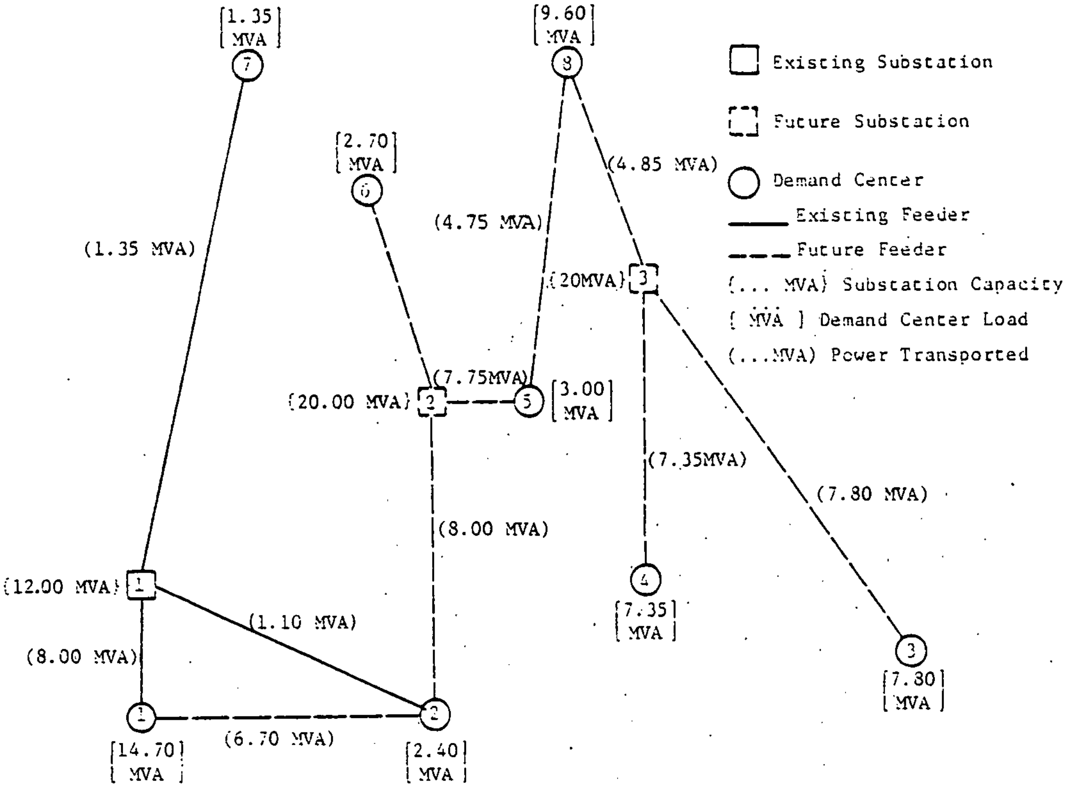


Figure 24. Case #3 solution results

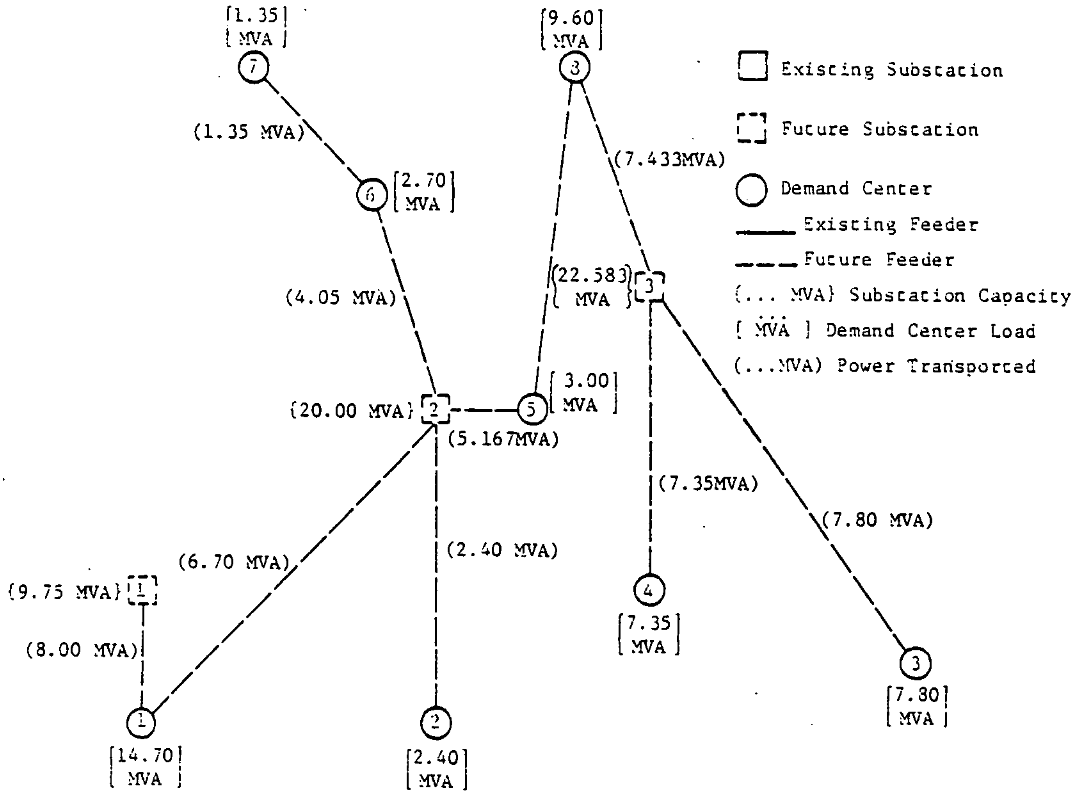


Figure 25. Results of sensitivity analysis #1 for case #1

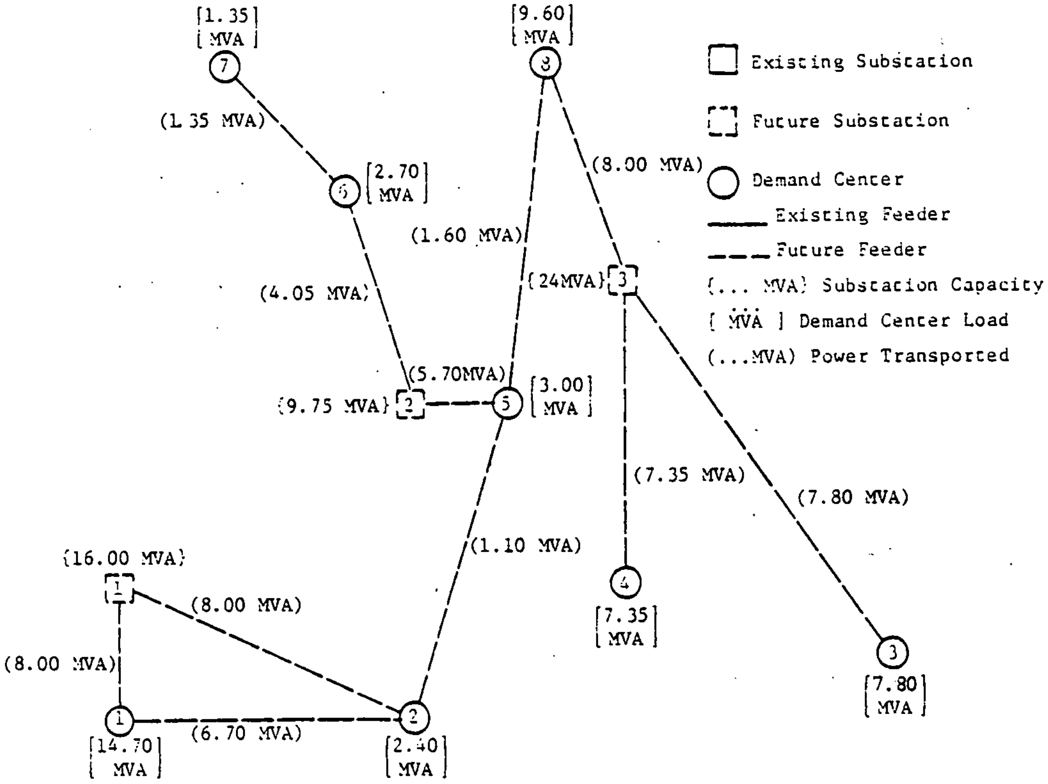


Figure 26. Results of sensitivity analysis #2 for case #1

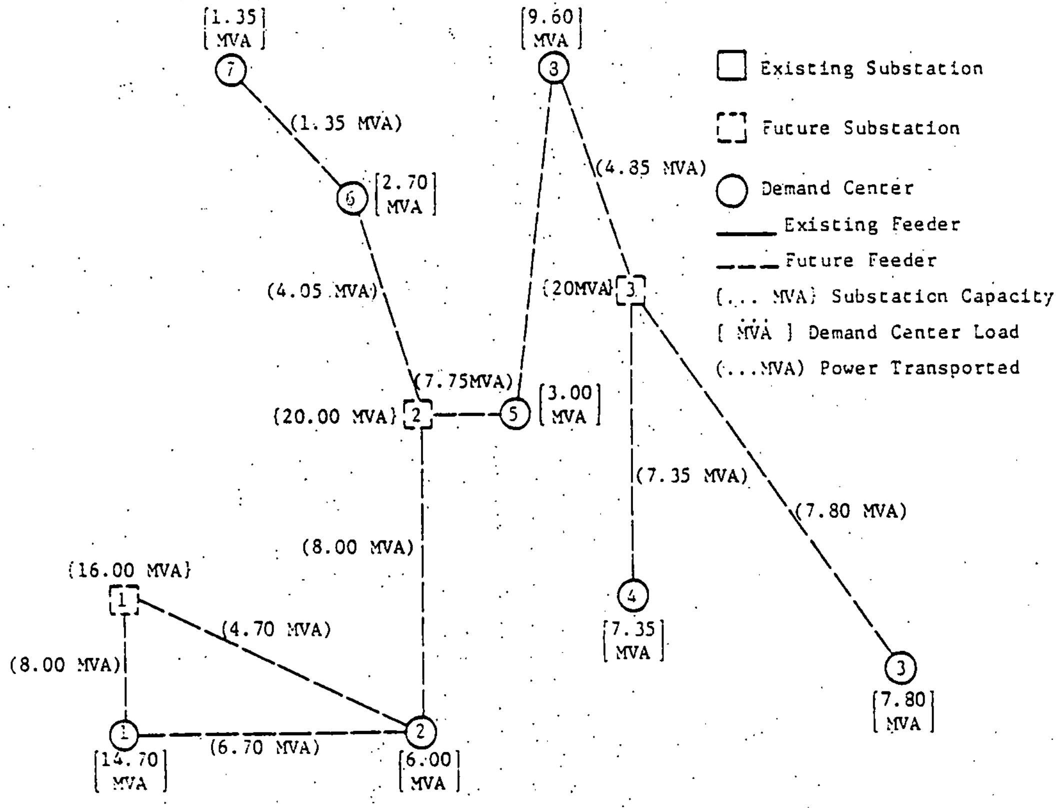


Figure 27. Results of sensitivity analysis #3 for case #1

REFERENCES

[1] R.F. Lawrence, et al., "Automated Distribution System Planning," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-83, pp. 311-316, April 1964.

[2] R.N. Adams, and M.A. Laughton, "Mathematical Programming Systems in Electrical Power Generation, Transmission, and Distribution Planning," in Proc. of 4th PSCC, Grenoble, France, 1972, Paper 1.1/13.

[3] E. Masud, "An Interactive Procedure for Sizing and Timing Distribution Substations Using Optimization Techniques," A Paper, Presented at the IEEE PES Winter Meeting, New York, NY, January 27 - February 1, 1974.

[4] M.J. Juricek, et al., "Transportation Analysis of an Electric Power Distribution System," A Paper, Presented at the IEEE PES Winter Meeting, New York, NY, January 25-30, 1976.

[5] D.M. Crawford and S.B. Holt, Jr., "A Mathematical Optimization Technique for Locating and Sizing Distribution Substations and Driving Their Optimal Service Areas," IEEE TRansactions on Power Apparatus and Systems, Vol. PAS-94, No. 2, pp 230-235, March/April, 1975.

[6] S.E. Shelton and A.A. Mahmoud, "A Direct Optimization Approach to Distribution Substation Expansion," A Paper, Presented at the IEEE PES Summer Meeting, Los Angeles, CA, July 16-21, 1978.

[7] J.V. Oldfield and T. Lang, "The Long Term Design of Electrical Power Distribution Systems," A Paper, Presented at the IEEE Power Industry Computer Applications (PICA) Conference, May 1965.

[8] R.N. Adams and M.A. Laughton, "A Dynamic Programming/Network Flow Procedure for Distribution System Planning," A Paper, Presented at the IEEE PICA Conference, June, 1975.

[9] D.L. Wall and J.E.D. Northcote-Green, "An Optimization Model for Planning Radial Distribution Networks, "A Paper, Presented at the IEEE Summer Meeting, Los Angeles, CA, July 16-21, 1978.

[10] D.C. Yu, "A Study of Computerized Modeling Approach for Distribution Systems Planning", A Master Thesis, University of Oklahoma, Norman, OK, 1979.

[11] T. Gonen, et al., "A Research Progress Report for Electrical Energy Distribution Systems," University of Oklahoma, Norman, OK, Feb. 1979.

[12] T. Gonen, et al., "Toward Automated Distribution System Planning," A Paper, Presented at the IEEE Control of Power Systems Conference, Texas A & M University, College Station, TX, March 19-21, 1979.

[13] B.L. Foote and T. Gonen, "Application of Mixed-Integer Programming to Reduce Sub-Optimization in Distribution Systems Planning, "A Paper, Presented at the Modeling and Simulation Conference, University of Pittsburgh, Pittsburgh, PA, April 25-27, 1979.

GENERAL DYNAMIC MODEL

The parameters of the dynamic model are different than the ones used in the previous base model since the envelope cost curves cannot be used[1]. In the dynamic model, as can be seen in Figure 28, power transmitted over the same feeder route can be different. Therefore, utilization of the same conductor may not give the optimum cost. Thus, all the material costs are included in a fixed cost. As a result of this, all the cost curves for transmitting power are to be based on power losses only.

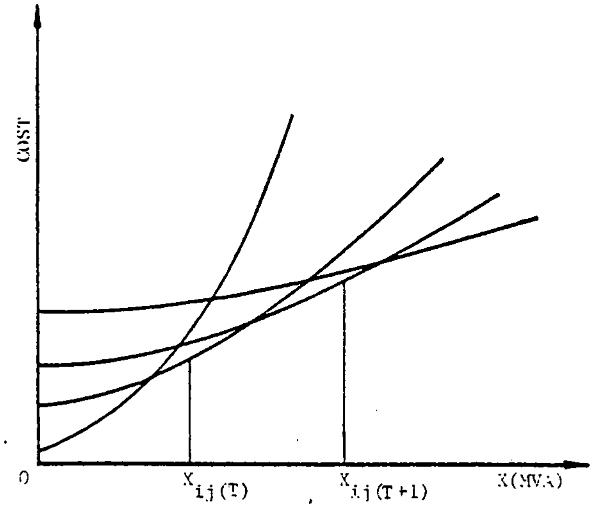


Figure 28

The cost curves required to be calculated as a function of conductor size and time. The cost curves look different with respect to time because the present worth of costs varies with time. The last year of the planning horizon T_h has a curve which includes the present worth of all costs from T_h to L, the last year of the useful life. The cost of power losses must be brought back to T_h and then to time zero. The assumption made here is that for a given power transmission of $X_{ij}(T)$ $T = T_h, T_h + 1, \ldots, L$ is constant.

Figure 29 presents the various power loss curves as a function of conductor size and time. Here, of course, $f_c(T_h)$ is an exception since it includes $L - T_h + 1$ years in its cost accumulation. Whereas, $f_c(T)$ for the value of T.

$$1 \leq T \leq T_h - 1$$

only accumulates costs for one year. In this model costs can be completely accounted for in the feeder fixed charge.

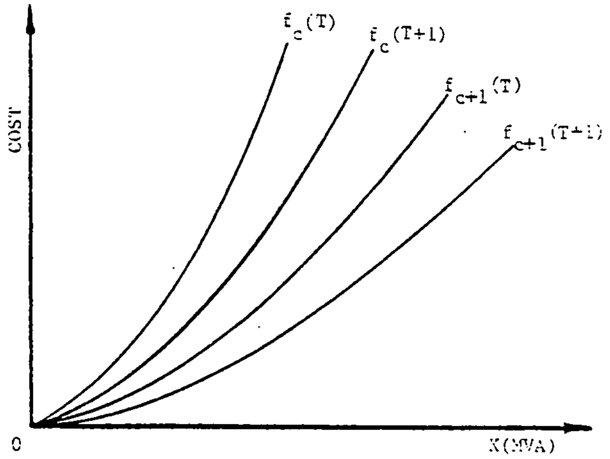


Figure 29

In order to set up the general dynamic, i.e., time-phased, model the following notation is introduced:

For Parameters:

NS = number of existing substations and potential substations,

M = total number of existing demand centers,

T_h = length of the planning horizon over which substations and feeders can be built,

T = a particular time period, $1 \leq T \leq T_h$,

SFC_{iT} = present worth of fixed cost of constructing substation i in year T,

FFC_{ijTc} = present worth of constructing a feeder from substation i to demand center j in year T of conductor size C,

TFC_{ijcT} = present worth of fixed cost constructing a tie-feeder from substation i to substation j in year T of conductor size C,

DFC_{ijcT} = present worth of fixed cost of constructing a feeder from demand center i to demand center J in year T of conductor size C,

FBC_{iT} = present worth of fixed cost of adding a bay at substation i (if facility already exists, the fixed charge is zero) in year T,

FRC_{ijcT} = present worth of reconducting feeder from substation i to demand center j in year T of conductor size c,

---

[1] Bobbie L. Foote and Turan Gonen, "Application of Mixed-Integer Programming to Reduce Sub-Optimization in Distribution Systems Planning", the 1979 Modeling and Simulation Conference, University of Pittsburgh, Pittsburgh, PA, April 25-27, 1979.

(a_{ijscT}, b_{ijscT}) = coordinates of a point s on the power loss curve for a specific conductor placed in route ij of conductor size c in year T,

DF_{jT} = total demand of demand center j in year T,

C = index number of maximum conductor size,

G = total number of points required to approximate a given nonlinear curve,

U_{ijc} = rated capacity of a feeder of conductor size c and connecting origin i to destination j,

NF = number of feeders emanating from the initial size of a substation,

NF_B = number of feeders per bay added to a substation,

NB_{max} = maximum number of bays that can be added to a given substation,

SIC_i = initial capacity of substation i,

C_{ikT} = present worth of fixed cost of adding incremental capacity k to substation i in year T.

For Power Flow Variables

X_{ijcT} = quantity transported from substation i to demand center j with conductor size c in year T,

Y_{ijcT} = quantity transported from substation i to substation j with conductor size c in year T,

Z_{ijcT} = quantity transported from demand center i to demand center j with conductor size c in year T,

$(RX)_{ijcT}$ = quantity transported from substation i to demand center j over a reconductored feeder with new conductor size c in year T.

For Decision Variables

δ_{iT} = binary integer variable which denotes the decision to select or not to select size i in year T,

δ_{iT} = 0, if a substation does not exist at the site i or will not be built in year T,

δ_{iT} = 1, if a substation is to be built at the site or already exists in year T,

δ_{ijcT} = 0, if a feeder with conductor size c does not exist between substation i and demand center j in year T,

δ_{ijcT} = 1, if a feeder with conductor size c does not exist or is to be built between substation i and demand center j in year T,

α_{ijcT} = 0, if a tie-feeder with conductor size c does not exist between substation i and substation j in year T,

α_{ijcT} = 1, if a tie-feeder with conductor size c does exist or is to be built between substation i and substation j in year T,

θ_{ijcT} = 0, if a feeder with conductor size c does not exist between demand center i and demand center j in year T,

θ_{ijcT} = 1, if a feeder with conductor size c does exist or is to be built between demand center i and demand center j in year T,

$(R\delta)_{ijcT}$ = 0, if reconducting, with conductor size c, of a tie-feeder between substation i and substation j will not be done in year T,

$(R\delta)_{ijcT}$ = 1, if reconductoring, with conductor size c, of a tie-feeder between substation i and substation j will be done in year T,

α_{ikT} = 1, if incremental capacity to be added to substation i in year T,

α_{ikT} = 0, otherwise.

μ_{iT} = number of bays to be added to the substation at site i in year T.

For Linearization Variables

The following variables are used to approximate a nonlinear curve with straight line segments.

t_{ijs} = representation variables for $f(X_{ij})$,

$(ty)_{ijs}$ = representation variables for $f(Y_{ij})$,

$(tz)_{ijs}$ = representation variables for $f(Z_{ij})$,

$(tR)_{ijs}$ = representation variables for $f[(RX)_{ij}]$,

$(\beta X)_{ijs}$ = decision variable to force selection of at most two t_{ijs} to be nonzero,

$(\beta Y)_{ijs}$ = decision variable to force selection of at most two $(tY)_{ijs}$ to be nonzero,

$(\beta Z)_{ijs}$ = decision variable to force selection of at most two $(tZ)_{ijs}$ to be nonzero,

$(\beta R)_{ijs}$ = decision variable to force selection of at most two $(tR)_{ijs}$ to be nonzero.

In the above definitions $t \geq 0$ and

$$(\beta X) \Leftrightarrow t$$
$$(\beta y) \Leftrightarrow (ty)$$
$$(\beta z) \Leftrightarrow (tz)$$
$$(\beta R) \Leftrightarrow (tR)$$
$$(ty) \Leftrightarrow Y$$
$$(tz) \Leftrightarrow Z$$
$$(tR) \Leftrightarrow (RX)$$

The optimization problem includes the timing of choosing the sites to locate substations; determining the optimum amount of incremental capacity to add to existing and/or newly built substations; determining the optimum number of feeders emanating from substations; finding the optimum number of bays required to support the number of feeders chosen; connecting the substations through tie-feeders; selecting the connections between demand centers; selecting the connections between substations and demands; the optimum conductor size of each connecting feeders; and the feeders between substations and demand centers which should be reconductored in such a way as to minimize the present value of costs and meet the forecasted demands.

DYNAMIC PLANNING MODEL

The following model is the complete mathematical model which, when solved, gives the optimum decision in each of the aforementioned categories. The definition of each term in the objective function and the definition of each constraint are given following the model.

$$\text{Cost} = \underbrace{\sum_{i=1}^{NS} \sum_{T=1}^{T_h} SFC_{iT} \cdot \delta_{iT}}_{(A)} + \underbrace{\sum_{T=1}^{T_h} \sum_{s=1}^{G} \sum_{i=1}^{NS} \sum_{j=1}^{m} t_{ijsT} \cdot d_{ijst}}_{(B)}$$

$$+ \sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{T=1}^{T_h} \sum_{c=1}^{C} FFC_{ijTc} \cdot \delta_{ijTc} \qquad i,j \in NE$$

(C)

$$+ \sum_{s=1}^{G} \sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{T=1}^{T_h} \sum_{c=1}^{C} t_{ijsTc} \cdot b_{ijscT}$$

(D)

$$+ \sum_{T=1}^{T_h} \sum_{i=1}^{NS} \sum_{j=1}^{NS} \sum_{c=1}^{C} TFC_{ijcT} \cdot \gamma_{ijcT}$$

(E)

$$+ \sum_{s=1}^{G} \sum_{i=1}^{NS} \sum_{j=1}^{NS} \sum_{c=1}^{C} \sum_{T=1}^{T_h} (ty)_{ijsTc} \cdot b_{ijscT}$$

(F)

$$+ \sum_{i=1}^{NS} \sum_{k=1}^{R} \sum_{T=1}^{T_h} C_{ikT} \cdot \alpha_{ikT}$$

(G)

$$+ \sum_{j=1}^{m} \sum_{i=1}^{m} \sum_{c=1}^{C} \sum_{T=1}^{T_h} \theta_{ijcT} \cdot DFC_{ijcT}$$

(H)

$$+ \sum_{s=1}^{G} \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{T=1}^{T_h} \sum_{c=1}^{C} (tz)_{ijsT} \cdot b_{ijscT}$$

(I)

$$+ \sum_{i=1}^{NS} \sum_{T=1}^{T_h} \mu_{iT} \cdot FBC_{iT}$$

(J)

$$+ \sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{c=1}^{C} \sum_{T=1}^{T_h} (FRC)_{ijcT} \cdot (R\delta)_{ijcT} \qquad i,j \in REC$$

(K)

$$+ \sum_{s=1}^{G} \sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{T=1}^{T_h} \sum_{c=1}^{C} (tR)_{ijscT} \cdot b_{ijscT} \qquad i,j \in REC$$

(L)

$$+ \sum_{s=1}^{G} \sum_{i=1}^{NS} \sum_{j=1}^{m} \sum_{T=1}^{T_h} \qquad i,j \in E$$

(M)

$$t_{ijscT} \cdot b_{ijscT}$$

where

$E = \{i,j \mid \text{route } i,j \text{ has a conductor in place}\}$

$REC = \{i,j \mid \text{route } i,j \text{ has a potential cost saving by reconductoring}\}$

$NE = \{i,j \mid \text{route } i,j \text{ has no conductor in place}\}$

$\dot{N}E = \{i,j \mid \text{set of all routes } i,\dot{T}\}$

Any fixed cost term in the above objective function is zero if facility is already in place. The objective function is subject to the following constraints:

$$\sum_{\substack{i=1\\ i\neq j}}^{NS} \sum_{c=1}^{C} X_{ijcT} + \sum_{\substack{i=1\\ i\neq j}}^{m} \sum_{c=1}^{C} Z_{ijcT} + \sum_{c=1}^{C} \sum_{\substack{i=1\\ i,j \in REC}}^{NS} (RX)_{ijcT} \geq DF_{jT} \qquad \forall_{j,T} \quad (34)$$

where

$j = 1, \ldots, m$

$T = 1, 2, \ldots, T_h$

$$\delta_{ijcT} \leq \sum_{T=1}^{T} \delta_{iT} \qquad \forall_{i,j,c,T} \qquad (35)$$

where

$c=1, \ldots, C$
$i=1, \ldots, NS$
$j=1, \ldots, m$
$T=1, \ldots, T_h$

$$X_{ijcT} \leq \sum_{T=1}^{T_h} U(c) \cdot \delta_{ijcT} \qquad \forall_{i,j,c,T} \qquad (36)$$

where

$c=1, \ldots, C$
$T=1, \ldots, T_h$
$j=1, \ldots, m$
$i=1, \ldots, NS$

$$Z_{ijcT} \leq \sum_{T=1}^{T_h} U(c) \cdot \theta_{ijcT} \qquad \forall_{i,j,c,T} \qquad (37)$$

where

$c=1, \ldots, C$
$T=1, \ldots, T_h$
$i=1, \ldots, m$
$j=1, \ldots, m$
$i \neq j$

$$Y_{ijcT} \leq \sum_{T=1}^{T} U(c) \cdot \gamma_{ijcT} \qquad \forall_{i,j,c,T} \qquad (38)$$

where

$i=1, \ldots, NS$
$j=1, \ldots, NS$
$c=1, \ldots, C$
$T=1, \ldots, T_h$

$$\sum_{c=1}^{C} \sum_{j=1}^{m} X_{ijcT} \leq SIC_i . \sum_{T=1}^{T} \delta_{iT} + \sum_{c=1}^{C} \sum_{j=1}^{NS} Y_{jicT} -$$

$$\sum_{c=1}^{C} \sum_{j=1}^{NS} Y_{ijcT} + \sum_{k=1}^{R} \sum_{T=1}^{T} C_{ikT} \alpha_{ikT} \qquad \forall_{i,T} \qquad (39)$$

where

$i=1,2 \ldots NS$
$T=1,2 \ldots T_h$

$$\sum_{k} \alpha_{ikT} \leq \sum_{T=1}^{T} \delta_{iT} \qquad \forall_{i,T} \qquad (40)$$

where

$i=1,2 \ldots NS$
$T=1,2 \ldots T_h$

$$\sum_{T=1}^{T} \delta_{iT} \leq 1 \qquad \forall_{i,T} \qquad (41)$$

where

$i=1,2 \ldots NS$
$T=1,2 \ldots T_h$

$$\sum_{j=1}^{m} \sum_{c=1}^{C} \sum_{T=1}^{T} \delta_{ijcT} \leq NF \cdot \sum_{T=1}^{T} \delta_{iT} + NF_B \cdot \left(\sum_{T=1}^{T} u_{iT} \right)$$

$$\forall_{i,T} \qquad (42)$$

where

$i=1,2 \ldots NS$
$T=1,2 \ldots T_h$

$$\delta_{ijcT} = 1 - \sum_{T=1}^{T} (R\delta)_{ijcT} \qquad \forall_{i,j,c,T} \qquad (43)$$

where

$i,j \in REC$

$T=1, \ldots, T_h$

$$\sum_{c=1}^{C} \sum_{T=1}^{T_h} (R\delta)_{ijcT} \leq 1 \qquad \forall_{i,j} \qquad (44)$$

where

$i,j \in REC$

$$(RX)_{ijcT} \leq U(c) \sum_{t=1}^{T} (R\delta)_{ijct} \qquad \forall_{i,j,c,T} \qquad (45)$$

where

$i,j \in REC$

$T=1, \ldots, T_h$

$$\sum_{c=1}^{C} \sum_{T=1}^{T_h} \delta_{ijcT} \leq 1 \qquad \forall_{i,j} \qquad (46)$$

where

$i=1, \ldots, NS$
$j=1, \ldots, m$

$$\sum_{c=1}^{C} \sum_{T=1}^{T_h} \theta_{ijcT} \leq 1 \qquad \forall_{i,j} \qquad (47)$$

where

$i=1, \ldots, m$
$j=1, \ldots, m$
$i \neq j$

$$\sum_{c=1}^{C} \sum_{T=1}^{T_h} \gamma_{ijcT} \leq 1 \qquad \forall_{i,j} \qquad (48)$$

where

$i=1, \ldots, NS$
$j=1, \ldots, NS$
$i \neq j$

$$\sum_{c=1}^{C} \gamma_{ijcT} \leq \tfrac{1}{2} \left(\sum_{T=1}^{T} \delta_{iT} + \sum_{i=1}^{T} \delta_{jT} \right) \qquad \forall_{i,j,T} \qquad (49)$$

where

$i=1, \ldots, NS$
$j=1, \ldots, NS$
$i \neq j$
$T=1, \ldots, T_h$

$$\sum_{T=1}^{T} \mu_{iT} \leq \left(\sum_{T=1}^{T} \delta_{iT} \right) . \mu_{max} \qquad \forall_{i,T} \qquad (50)$$

where

$i=1,2 \ldots NS$
$T=1,2 \ldots T_h$

$$Z_{ijcT} = \sum_{s=1}^{G} (tz)_{ijcTs} \cdot a_{ijcst}$$

$$\sum_{s=1}^{G} (tz)_{ijcTs} \leq 1$$

where

$i=1,2 \ldots m$
$j=1,2 \ldots m$
$c=1,2 \ldots C$
$T=1,2 \ldots T_h$
$i \neq j$

$$\sum_{s=1}^{G-1} (\beta z)_{sijcT} \leq 1 \qquad \forall_{i,j,c,T} \qquad (51)$$

$$(tz)_{1ijcT} \leq (\beta z)_{1ijcT}$$

$$\vdots$$

$$(tz)_{gijcT} \leq (\beta z)_{(g-1)ijcT} + (\beta z)_{gijcT}$$

$$\vdots$$

$$(tz)_{GijcT} \leq (\beta z)_{(G-1)ijcT}$$

where

$i=1,2 \dots NS$
$j=1,2 \dots NS$
$T=1,2 \dots T_h$
$c=1,2 \dots C$

$$(RX)_{ijcT} = \sum_{s=1}^{G} (tR)_{sijcT} \cdot a_{sijscT}$$

$$\sum_{s=1}^{G} (tR)_{sijcT} = 1$$

$$\sum_{s=1}^{G-1} (\beta R)_{sijcT} = 1 \qquad \forall i,j,c,T \qquad (52)$$

$$(tR)_{1ijcT} \leq (\beta R)_{1ijcT}$$
$$\vdots$$
$$(tR)_{gijcT} \leq (\beta R)_{(g-1)ijcT} + (\beta z)_{gijcT}$$
$$\vdots$$
$$(tR)_{GijcT} \leq (\beta R)_{(G-1)ijcT}$$

where

$i,j \in REC$
$c=1,2 \dots C$
$T=1,2 \dots T_h$

$$Y_{ijcT} = \sum_{s=1}^{G} (ty)_{sijcT} \cdot a_{ijscT}$$

$$\sum_{s=1}^{G} (ty)_{sijcT} = 1$$

$$\sum_{s=1}^{G-1} (\beta y)_{sijcT} = 1 \qquad \forall i,j,c,T \qquad (54)$$

$$(ty)_{1ijcT} \leq (\beta y)_{1ijcT}$$
$$\vdots$$
$$(ty)_{gijcT} \leq (\beta y)_{(g-1)ijcT} + (\beta y)_{gijcT}$$
$$\vdots$$
$$(ty)_{GijcT} \leq (\beta y)_{(G-1)ijcT}$$

$i=1, \dots, NS$
$j=1, \dots, NS$
$c=1, \dots, C$
$T=1, \dots, T_h$
$i \neq j$

$$X_{ijcT} \leq U(c) \cdot \delta_{ijcT} \qquad \forall i,j,T \qquad (55)$$

where

$i,j \in E$

$$X_{ijcT} = \sum_{s=1}^{g} t_{sijcT} \cdot a_{ijscT}$$

$$\sum_{s=1}^{g} t_{sijcT} = 1$$

$$\sum_{s=1}^{g-1} t_{sijcT} = 1$$

$$t_{1ijcT} \leq (\beta X)_{1ijcT} \qquad \forall i,j,c,T \qquad (56)$$
$$\vdots$$
$$t_{gijcT} \leq (\beta X)_{(g-1)ijcT} + (\beta X)_{gijcT}$$
$$\vdots$$
$$t_{gijcT} \leq (\beta X)_{(g-1)ijcT}$$

where

$i,j \in E$, c is known

and

$T = 1, \dots, T_h$.

Also

$$X_{ijcT} \geq 0; \; Y_{ijcT} \geq 0; \; Z_{ijcT} \geq 0$$

$$(RX)_{ijcT} \geq 0 \quad (ty)_{sijcT} \geq 0; \quad (tz)_{sijcT} \geq 0$$

$$t_{sijcT} \geq 0; \; (tR)_{sijcT} \geq 0$$

$$(\beta X)_{sijcT} = 0,1; \; (\beta Y)_{sijcT} = 0,1; \; (\beta Z)_{sijcT} = 0,1$$

$$(\beta R)_{sijcT} = 0,1; \; \delta_{iT} = 0,1; \; \delta_{ijcT} = 0,1$$

$$(R\delta)_{ijcT} = 0,1; \; \theta_{ijcT} = 0,1; \; \alpha_{ikT} = 0,1$$

$$\gamma_{ijcT} = 0,1; \; u_{iT} = 1, \dots, NB_{max}$$

The following are the explanations of the terms in the objective function:

Term A: gives the fixed charges for constructing a substation on site i over the time horizon T_h.

Term B: gives the cost of power losses in substation equipment, represented by a piecewise linear approximation of power transported over route (i,j) over the time horizon T_h.

Term C: gives the fixed charges of an installed feeder of size C on the route from substation i to demand center j over the time horizon T_h.

Term D: gives the cost of power losses when power quantity X_{ij} is transported over conductor C represented by a piecewise linear approximation of the nonlinear cost curve over the time horizon T_h.

Term E: gives the fixed charges of an installed tie-feeder of size C between substations over the time horizon T_h.

Term F: gives the cost of power losses in tie-feeders between substations as a result of transshipment over conductor C, represented by a piecewise linear approximation of the nonlinear cost curve over the time horizon T_h.

Term G: gives the cost of adding incremental capacity at the various substations over the time horizon T_h.

Term H: gives the fixed charges for installed feeders of size C between demand centers over the time horizon T_h.

Term I: gives the cost of power losses in transshipment between demand centers over conductor C with piecewise linear representation of the nonlinear cost curve over the time horizon T_h.

Term J: gives the fixed charges for adding bays to a substation over the time horizon T_h.

Term K: gives the fixed charges of reconductoring a feeder over route (i,j) over the time horizon T_h.

Term L: gives the cost of power losses in feeders over reconductored route (i,j), represented by a piecewise linear approximation of the nonlinear cost curves during the time period T_h.

Term M: gives the cost of power losses of conductor C over a route (i,j) with a feeder in place, represented by a piecewise linear approximation of the nonlinear cost curves during the time horizon T_h.

The following are the explanations of the constraints used in the model:

Constraint (34): assures that each demand is met during the time horizon T_h.

Constraint (35): stops a feeder, connecting substation i to demand center j, from being built unless the substation has been built by time T.

Constraint (36): assures that no power flow over a feeder route connecting i to demand center j can occur unless the feeder has been built by time T.

Constraint (37): assures that there is no power flow between demand centers unless the connecting feeder exist (i.e., $\theta_{ijcT} = 1$).

Constraint (38): assures that the power flow between substations is zero if no feeder exists between substations.

Constraint (39): assures that the total quantity transported from substations to destinations is less than the substation capacity plus power transferred from other substations for any given time during the time horizon T_h.

Constraint (40): assures that incremental capacity cannot be added unless the substation exist (i.e., $\delta_{iT}=1$) by time T.

Constraint (40): assures that the number of feeders be built, if at all, by time T.

Constraint (41): assures that the number of feeders allowed at time T is less than or equal to the number of initial feeders plus the number of feeders per bay times the number of bays added by the time T.

Constraint (42): assures that there is no power flow over old feeder if a new reconductored feeder has been placed in time period T.

Constraint (43): assures that only one conductor size is used to reconductor the feeder and that the reconductoring occurs only once.

Constraint (44): assures that power flow over reconductored feeder route occurs only if a reconductoring decision has been made.

Constraint (45): assures that only one conductor size can be chosen for a given feeder between substation i and demand center j during the time horizon T_h.

Constraint (46): assures that only one conductor size can be chosen for a given feeder between demand center i and demand center j during the time horizon T_h.

Constraint (47): assures that only one conductor size can be chosen for a given tie-feeder between substation i and substation j during the time horizon T_h.

Constraint (48): assures that no feeder can be built between substations at time T unless both substations exist (i.e., $\delta_{iT}=1$ and $\delta_{jT}=1$) by time T.

Constraint (49): assures that the number of bays are limited and that bays cannot be added at time T unless a substation exists by the time T.

Constraint (50): computes the piecewise linearization variables of the power flow between demand centers for $f(Z_{ij})$ i.e., the term (I) in the objective function.

Constraint (51): assures that the number of bays are limited and that bays cannot be added at time T unless a substation exists by the time T.

Constraint (52): computes the piecewise linearization variables of the power flow between demand centers for $f(Z_{ij})$ i.e., the term (I) in the objective function.

Constraint (53): computes the piecewise linear approximation of the power loss cost function of the power transmitted over reconductored feeder routes in conjunction with the term (L) in the objective function.

Constraint (54): computes the piecewise linear approximation of the power loss cost function of the power transmitted over feeder routes between substation i and demand center j at time T over conductor size C in conjunction with the term (D) in the objective function.

Constraint (55): computes the piecewise linear approximation of the power loss cost function of the power transmitted over tie-feeder routes between substation i and substation j at time T over conductor size C in conjunction with the term (F) in the objective function.

Constraint (56): stops the power flow if feeder decision variable is zero.

Constraint (57): computes the piecewise linear approximation of the power loss cost function of the power loss curves for conductors in place with a given conductor size in conjunction with term (M) in the objective function.

OPTIMIZATION MODEL FOR DISTRIBUTION PLANNING USING REAL AND REACTIVE POWER FLOWS

Let the number of existing substations be NES and the number of possible sites on which to build substations be NPS, as were previously. Assume that the planner has to make a number of decisions. For example, building a new substation, or substations, capacity additions to existing substations, addition of feeders from substations to demand centers, installation of new capacitor banks, and determining optimum conductor sizes.

Assume that real and reactive power requirements of each demand center are known. Then the model can be developed in such a manner to include the aforementioned properties at the minimum cost. Therefore, let S_{ij}, P_{ij} and Q_{ij} be the total appearant power, real power, and reactive power transmitted from substation i to demand center j, respectively. Thus, the following equation relates the three variables as

$$S_{ij}^2 = P_{ij}^2 + Q_{ij}^2 \qquad \forall i,j \qquad (57)$$

Assuming that the utility company's policy dictates to achieve a power factor of 0.9, as is the usual case, the following restriction can be written as

$$P_{ij} \geq 0.9 S_{ij} \qquad \forall i,j \qquad (58)$$

Since the total real power transmitted from all of the substations must satisfy the real power demand of the demand centers

$$\sum_{i=1}^{n} P_{ij} \geq DP_j \qquad \forall j \qquad (59)$$

where

$n = NES + NPS$,

DP_j = total real power demand of demand center j.

Since the reactive power demand of the load centers is met by transmitting reactive power from the substations and by installing capacitor banks, the following restriction can be written as

$$Q_j + \sum_{i=1}^{n} Q_{ij} \geq DQ_j \qquad \forall j \qquad (60)$$

where

Q_j = magnitude of the reactive power supplied by the capacitor bank located at demand center j,

DQ_j = total reactive power demand of the demand center j.

The power transmitted from a substation must be less or equal to the capacity of the substation plus any additions to the capacity. Therefore

$$\sum_{j=1}^{m} S_{ij} \leq CAP_i \cdot \delta_i + \sum \Delta_k \cdot \alpha_{ik} \qquad \forall i \qquad (61)$$

where

CAP_i = initial capacity of the substation,

δ_i = 0, if a substation does not exist at the site i or will not be built,

δ_i = 1, if a substation is to be built at the site or already exists,

Δ_k = k'th incremental capacity size available to increase the capacity of a given substation,

α_{ik} = a zero-one decision variable.

Since only one capacity addition is allowed to be made at any given time

$$\sum_{k=1}^{k} \alpha_{ik} \leq 1 \qquad \forall i \qquad (62)$$

Assuming that the permissable voltage drop is limited to 5 percent of base voltage (or to 0.05 pu V)

$$\frac{\sqrt{S_{ij}|Z_{ij}|}}{V_B} \leq 0.05 \text{ pu V} \qquad (63)$$

or

$$\sqrt{S_{ij}|Z_{ij}|} \leq 0.05\, V_B \qquad (64)$$

or

$$S_{ij}|Z_{ij}| \leq 25 \times 10^{-4} V_B^2 \qquad (65)$$

where

S_{ij} = apparent power, in kVA,

$|Z_{ij}|$ = magnitude of the Thevenin equivalent impedance, in ohms,

V_B = base voltage, in kV.

where

$$|Z_{ij}| = R_{ij}^2 + X_{ij}^2$$

R_{ij} = resistance of the Thevenin equivalent impedance, in ohms,

X_{ij} = reactance of the Thevenin equivalent impedance, in ohms,

In a system of overhead lines the reactive component of the Thevenin equivalent is much larger than its resistive component, contrary to a system of underground cables. However, by installing fixed and variable capacitor banks at the substation and/or load centers to comply with the restriction (58), the mangnitude of the reactive component of the Thevenin equivalent impedance may be negligible, with a small error involvement. Therefore,

$$|Z_{ij}| \cong R_{ij}$$

Hence

$$S_{ij} \cdot R_{ij} \leq 25 \times 10^{-4} V_B^2$$

or

$$S_{ij} \leq \frac{1}{R_{ij}} 25 \times 10^{-4} V_B^2$$

where

R_{ij} = resistive component of the Thevenin equivalent, in ohms per mile,

S_{ij} = apparent power transported from origin i to demand center j in kVA,

V_B = base voltage, in kV.

Of course, as a result of the installment of the capacitor banks, the copper losses or I^2R losses of the system decreases considerably.

Also,

$$S_{ij} \leq \delta_{ij} \cdot M \qquad (66)$$

where

S_{ij} = apparent power transported from origin i to destination j, in kVA.

δ_{ij} = 0, if a feeder does not exist between origin i and destination j,

δ_{ij} = 1, if a feeder does exist or is to be built between origin i and destination j,

M = A very large number.

Since feeders cannot be built unless a substation exits,

$$\sum_{j=1}^{m} \delta_{ij} \leq NF(\delta_i + \Theta_i) \qquad \forall_i \qquad (68)$$

where

NF = number of feeders that can emanate from a bay, at a substation,

δ_i = 0, if a substation does not exist at the site i or will not be built,

δ_i = 1, if a substation is to be built at the site or already exists,

Θ_i = the number of bays allowed per bay,

$\Theta_i = 1,2,3 = \Theta_{max}$

The following restriction keeps a bay from being added unless a substation exists or is to be built,

$$\Theta_i \leq \delta_i (\Theta_{max}) \qquad \forall_i \qquad (69)$$

Since additional capacity cannot be added unless bays are added, thus

$$\alpha_{ik} \leq \Theta_i \qquad \forall_i \qquad (70)$$

In general, the costs involved are

SFC_i = present worth of fixed costs of substation (including land, right-of-ways, and some construction costs),

SVC_i = present worth of variable costs of substation i,

BAC_i = present worth of fixed costs of adding a bay at substation i,

FFC_{ij} = present worth of fixed costs of a feeder connecting origin i to destination j,

C_{ik} = present worth of cost of the k'th capacity increment to the i'th substation (without including the cost of bay),

CAP_j = cost of adding capacitor bank at j.

Assuming the cost of power losses has a nonlinear function, the total cost or objective function which needs to be minimized is

$$Z = \sum_{i=1}^{n} SFC_i \cdot \delta_i + \sum_{i=1}^{\Theta max} BAC_i \cdot \Theta_i + \sum_{i=1}^{n} \sum_{j=1}^{m} FFC_{ij}$$

$$\cdot \delta_{ij} + \sum_{i=1}^{n} \sum_{k=1}^{k} C_{ik} \cdot \alpha_{ik}$$

$$+ \sum_{i=1}^{n} SVC_i \cdot S_{ij} + \sum_{i=1}^{n} \sum_{j=1}^{m} FVC_{ij} \cdot S_{ij} + \sum_{j=1}^{n} CAP_j \cdot Q_j \qquad (71)$$

where

$$Q_j \geq 0; \ S_{ij} \geq 0; \ P_{ij} \geq 0; \ Q_{ij} \geq 0$$

and

$$\delta_i = 0,1; \quad \delta_{ij} = 0,1$$

$$\alpha_{ik} = 0,1$$

$$\Theta_i = 1,2 \ldots, \Theta_{max}$$

The model developed is a non-linear mixed-integer programming model. There may be several refinements. For example, a fixed charge can be used for the installment of capacitor banks by defining $(FQ)_j = 0,1$ and requiring

$$(FQ)_j \geq \frac{Q_j}{Q_{max}} \qquad (72)$$

Thus, the following term needs to be added to the objective function,

$$\Sigma(FQ)_j \cdot CFC \qquad (73)$$

where

CFC = fixed charge for capacitor bank installation.

THE CONCEPTUAL DISTRIBUTION SYSTEM PLANNING MODEL

Introduction

In general, the ultimate criteria for a good distribution system planning decision are: (1) the one which provides the best result from a set of alternatives, and (2) the one which is reached most quickly and economically. Therefore, the planning engineer has to use some computerized analytical techniques to evaluate the technical and economic merits of alternative plans. There is a need for large scale analysis as well as information retrieval and display, in choosing the best alternative required an interactive problem solving environment based on a data base management system and a library of application programs, with an analytical capability that is far beyond the conventional algorithms. The systems approach would enable the planning engineer to simulate the distribution system in order to develop and evaluate, using cost sensitive evaluation models, plans for a given year or to treat the planning activity as a continuous process. Currently, only a few electric utility companies organize their planning activities in a way similar to this conceptual system. Most of them still treat the planning process not on a integrated system basis, but rather, on the basis of a partitioning of the total system into subproblems which, in turn, results a nonoptimal system planning.

Future Nature of Distribution Planning

"Predictions about the future methods for distribution planning must necessarily be extrapolations of present methods. Basic algorithms for network analysis have been known for years and are not likely to be improved upon in the near future. However, the superstructure which supports these algorithms and the problem-solving environment used by the system designer is expected to change significantly to take advantage of new methods which technology has made possible. Before giving a detailed discussion of these expected changes, the changing role of distribution planning needs to be examined".[2]

Increasing Importance of Good Planning

Because of the economic reasons, distribution system will become more expensive to build, expand and modify. Thus it is particularly important that each distribution system design be as cost effective as possible. This means that the system must be optimal from many points of view over the time period from first day of operation to the planning time horizon. In addition to the accurate load growth estimates, components must be phased in and out of the system so as to minimize capital expenditure, meet performance goals and minimize losses.

These requirements need to be met at a time when demographic trends are veering away from what have been their norms for many years in the past and when distribution systems are becoming more complex in design due to the appearance of more active components (e.g., fuel cells) instead of the conventional passive ones.

Cost/Benefit Ratio for Innovation

In the utility industry, the most powerful force shaping the future is that of economics. Therefore, any new innovations are not likely to be adopted for their own sake. These innovations will be adopted only if they reduce the cost of some activity or provide something of economic value which previously had been unavailable for comparable costs. In predicting that certain practices or tools will replace current ones, it is necessary that one judge their acceptance on this basis.

The expected innovations which satisfy these criteria are planning tools implemented on a digital computer which deal with distribution systems in network terms. In TASK A.1 and TASK A.2, a list of currently available such planning tools are given, and one might be tempted to conclude that these tools would be adequate for industry use throughout the 1980's. That this is not likely to be the case may be seen by considering the trends judged to be dominant during this period with those which held sway over the period in which the tools in were developed.[3]

New Planning Tools

Tools to be considered fall into two categories: network design tools and network analysis tools. The analysis tools may become more efficient but are not expected to undergo any major changes although the environment in which they are used will change significantly. This environment is discussed in the next section.

The design tools, however, are expected to show the greatest development since better planning could have a significant impact on the utility industry. The results of this development will show the following characteristics:

(1) Network design will be optimized with respect to many criteria using programming methods of operations research.

(2) Network design will be only one facet of distribution system management directed by human engineers using a computer system designed for such management functions.

(3) So-called network editors[4] will be available for designing trial networks; these designs in digital form will be passed to extensive

---

[2] John C. Thompson and Turan Gonen, "Distribution System Planning: Past, Present, and the Future-Part II: The Future", IEEE MEXICON-79 Conference, Mexico City, Mexico, Sept. 10-12, 1979.

[3] Turan Gonen and John C. Thompson, "Distribution System Planning: Past, Present, and The Future-Part I: Past and Present", IEEE MEXICON-79 Conference, Mexico City, Mexico, Sept. 10-12, 1979.

[4] John C. Thompson and Turan Gonen, "An Interactive Distribution System Planning Model", the 1979 Modeling and Simulation Conference, University of Pittsburgh, PA, April 25-27, 1979.

simulation programs which will determine if the proposed network satisfies performance and load growth criteria.

The Central Role of the Computer in Distribution Planning

As is well known, distribution system planners have used computers for many years to perform the tedious calculations necessary for system analysis. However, it has only been in the past few years that technology has provided the means for planners to truly take a "systems approach" to the total design and analysis. It is the central theses of this section that the development of such an approach will occupy planners in the 1980's and will significantly contribute to their meeting the challenges previously discussed.

THE SYSTEMS APPROACH

A collection of computer programs to solve the analysis problems of a designer does not necessarily constitute an efficient problem solving system nor even does such a collection when the output of one can be used as the input of another. The systems approach to the design of a useful tool for the designer begins by examining the types of information required and its sources. The view taken is that this information generates decisions and additional information which pass from one stage of the design process to another. At certain points, it is noted that the human engineer must evaluate the information generated and add his inputs. Finally, the results must be displayed for use and stored for later reference. With this conception of the planning process, the systems approach seeks to automate as much of the process as possible, insuring in the process that the various transformations of information are made as efficiently as possible. One representation of this information flow is shown in Figure 30. Here, the outer circle represents the interface between the engineer and the system. Analysis programs forming part of the system are supported by a database management system which stores, retrieves, and modifies various data on distribution systems.

The Database Concept

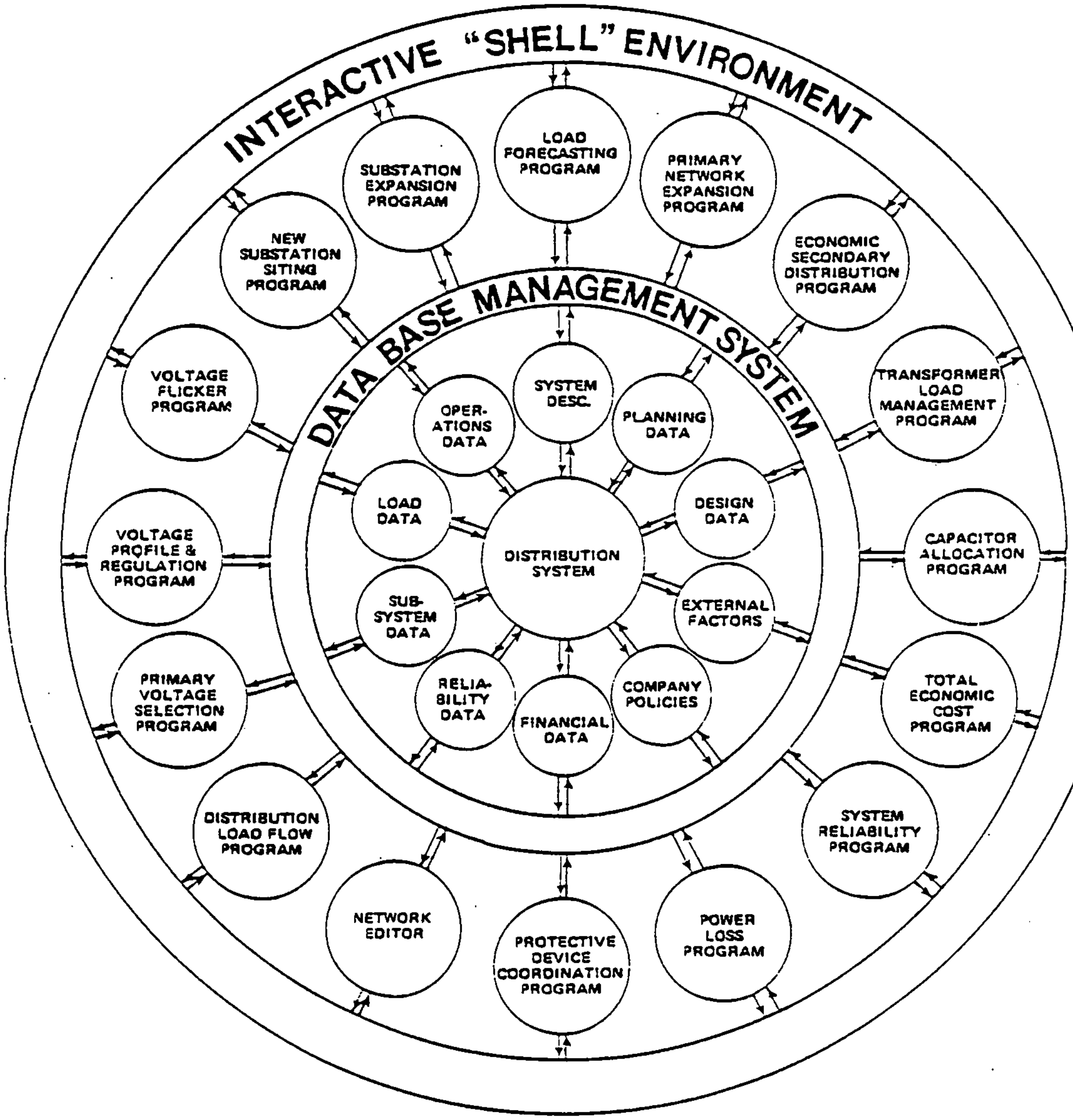


Figure 30. A schematic view of a distribution system planning system

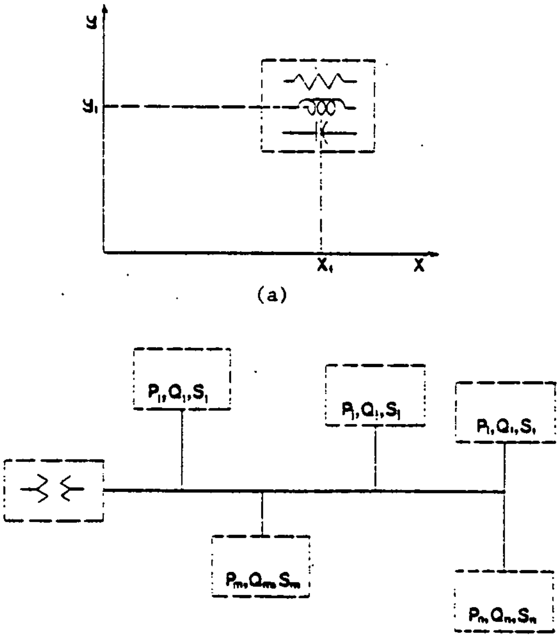
As suggested in Figure 30, the database plays a central role in the operation of such a system. It is in this area that technology has made some significant strides in the past five years so that not only is it possible to store vast quantities of data economically, but it is also possible to retrieve desired data with access times on the order of seconds. The database management system provides the interface between the process which requires access to the data and the data itself. The particular organization which is likely to emerge as the dominant one in the near future is based on the idea of a relation. A more detailed discussion of such a scheme is presented in TASK B.10 and Appendix F.

Networks

The second key element of the conceptual model is the generalized notion of a network. For purposes of logical economy, it is required that all important quantities be associated with components of the distribution network. At first thought, it may appear that this is not desirable since conventionally, much of the system description has been associated with a coordinate grid system. In the model, however, a component of the distribution system is represented as a vertex in a directed graph. Connections between system components are represented as arcs in the graph.

Consider, for example, the total load in some area of the service grid. It may consist of resistive, inductive and capacitive elements, which are distributed over the area. This load may be modeled as a vertex which has parameters such as coordinate position, area, and load coefficients associated with it. This idea is illustrated in Figure 31. Figure 31(a) represents the load for a given area modelled by a graph vertex and Figure 31(b) shows that vertex and other similar vertices combined to form a subnetwork. Other network components, such as substations, transformers, distribution lines, etc. are modelled in a similar way.

The relational data base scheme is flexible enough that a network can be stored conveniently as a collection of relations. Thus, the entire model of a distribution system is represented as a network and all data pertaining to the model is stored in a collection of relations, including the network representation itself.



(a)

(b)

Figure 31. Network representation: (a) "Lumped" network components, (b) a sample network

Planning Programs

The third important ingredient of the model consists of the planning programs in TASK A.1 and TASK A.2. The system has been designed to incorporate the functions of all the programs presented in Table 1 plus additional programs which extend the system's capabilities beyond those in use by the nation's utilities. Some of the required programs are:

1. Load forecasting program,
2. Primary network expansion program,
3. Substation expansion program,
4. New substation siting program,
5. Economic secondary distribution program,
6. Transformer load management program,
7. Distribution load flow program,
8. Primary voltage selection program,
9. Voltage profile and regulation program,
10. Voltage flicker program,
11. Capacitor allocation program,
12. Power loss program,
13. Protective device coordination program
14. System reliability program
15. Total economic cost program,
16. Network editor.

The load forecasting program is a set of programs to predict demands by various customer classes from available data. The primary network expansion program, the substation expansion program, and the new substation siting program are the subset programs of the aforementioned distribution system planning models utilizing mixed-integer programming. The economic secondary distribution program of the distribution transformer and secondary combination subsystem checks all designs against user-furnished criteria which may include: (1)allowable voltage drops and voltage flicker, (2)allowable maximum transformer overloading, (3)power losses in the distribution transformer and secondary system, (4)phase-load balance for the primary system, (5) investment cost of the secondary system, and (6)other engineering and economic considerations.

The transformer load management (TLM) program provides a computerized data base for distribution transformer loading patterns from which area load growth projections can be developed. The distribution load flow program calculates the real and reactive power flows and flows and bus voltage levels in a given distribution network. The program can be improved to perform fault and phase-balance studies as well as the computation of current flows, voltage profiles,and power factors. Further, the program can be designed to provide ratings for sectionalizer or recloser and regulators.

The primary voltage selection program performs an economic study by generating rapidly the various cost elements of alternative system designs suggested by the planner and based on load densities, voltage control requirements, power losses in conductors and transformers, the effects on overall system reliability, protection schemes, and possibly a number of external and internal factors to choose the most economic distribution voltage levels.

The voltage profile and regulation program calculates the voltage profiles along primary feeders to determine voltage drop deviations and the requirements for voltage regulation. The voltage flicker program computes the voltage fluctuations and temporary voltage dips below the minimum allowable voltage level due to customer utilization apparatus, e.g., motor starting, welding, electric furnaces, etc.

The capacitor allocation program calculates the optimum size of capacitor banks and their locations along a primary feeder circuit. The power loss programs computes the power losses of alternative system designs.

The protective device coordination program computes the optimum ratings of fuses, the ratings and settings of reclosers and sectionalizers, and the pickup current and time delay settings of overcurrent relays.

The system reliabilty program computes the system reliability indices of alternative system designs based on geographical location, weather conditions, customer mix, primary voltage level, overall system configuration, and a number of other factors. The total economic cost program calculates the cost of alternative system designs.

System Organization

A schematic diagram of the system is shown in Figure 30. The major functional components are the shell program, i.e. and executor or overload program; the planning programs which run under its control; a network editor which allows the distribution system planner to construct and modify networks; and finally, the data base management system which both supports the network editor functions affecting the data base as well as supplying additional capabilities for data organization, retrieval and input-output.

The Shell

One of the design goals of the implemented system was to make its operation as independent as possible of the host hardware/software. In order to achieve this portability, given the present technology, it was necessary to simulate many of the functions of a computer's operating system with the distribution system planning model (DSPM) itself. These functions are located in the shell program which provides an interactive interface between the distribution system planner and the functions and programs of the DSPM. Some of the important functions of the shell are summarized in Table 29.

Table 29. Shell Commands

COMMAND	FUNCTION
INVOKE	Connect the shell to a particular data base which supplies all data required by programs running under control of shell.
EXECUTE	Begin the execution of a program under control of the shell.
PIPELINE	Execute a series of programs; the output of a previous program becomes the input of a succeeding program.
EXECUTE CONTROL FILE	A file containing shell commands is executed.
CREATE	Create a file for use by programs under the shell's control.
OPEN	Prepare a file for reading or writing.
CLOSE	Terminate processing for a file.
UNLINK	Remove a file from the system.
SAVE	Permanently retain a file.
LIST FILES	List all files assigned to a particular user name.

The Network Editor

Among the processes subordinate to the shell is the network editor. In planning power distribution systems, it is necessary to have a model of the distribution network. For the purposes of this discussion, such a model consists of a graph, whose vertices are network components such as transformers, loads, etc., and edges which represent connections among the components.

An engineer involved in planning will need to have the capability of constructing network models and will also need to modify such models. Additionally, the investigator may wish to construct several related models to explore alternatives in power system design. The network editor serves to facilitate and support the construction and modification of power system network models.

The features of the network editor is discussed in terms of network objects, control mechanisms and command functions. A primitive network object is comprised of a name, an object class description and a connection list. The name, uniquely identifies the object class description contains a list of parameters which specify the object class and a list of input and output ports. The connection list defines the components to which the object of interest is connected.

Several control mechanism features provide the planner with natural tools for correct construction and modification. The first of these is that of state variables. These are used to identify a particular setting of a control. State variables are included in the object class description of an object and may have either numerical or boolean values. In the case of a variable capacitor, the capacitance would be considered a numerical state variable. In the case of a switch, the setting (off, on) is a boolean state variable. The values of state variables may be changed under the engineer's control without modifying the network by replacing one component by another. This is not the case for object class parameters which are not state variables. A numerical state variable may be assigned any arithmetic expression whose operands are numerical state variables or boolean constants.

A second important control mechanism feature is that of the design switch. A design switch may be inserted into any network object. The resulting object is then controlled, made to "appear" and "disappear" by a boolean state variable corresponding to the design switch. If a design switch is included in an object class description, then all objects which are instantiations of the object class change state whenever the corresponding boolean state variable is changed. Several design switches may be controlled by a single state variable; accordingly, several alternative network configurations may be obtainable by altering one or two state variables.

In order to correctly and automatically connect several input ports of one device to several "corresponding" output ports of a second device connection descriptors are used. These may be thought of as polarized connectors which permit connections to be made only one way. During object definition, each port gets a "pin number", and this number determines the corresponding port on all other devices.

In order to manipulate connections, the network editor provides a facility for assigning cursors to network objects. A cursor is a pointer which specifies a particular network object. Some cursors point to objects within the network of interest while others point to "parts" i.e., objects which are not currently connected to the network. These parts may be primitive or composite objects.

Composite objects are objects formed from network objects by including them in a subnetwork. Such composite objects have object class descriptions which consists of the union of object class descriptions of their components and connection lists representing the unspecified portions of the connection lists of their components. A composite object is to a network as a FORTRAN subroutine is to a program. The major network editor commands which the user has at his disposal are shown in Table 30. Also see Appendix F for a complete functional description of the network editor.

The final major subsystem running under the aegis of the shell is the data base management system (DBMS). All transactions against the data base are processed by the DBMS. This insures that the integrity of the data base is maintained, regardless of which process initiates a transaction, be it the designer himself via the shell, the planning programs or the network editor. The capabilities of the DBMS are primarily those found in any current relational data base. The most important are summarized in Table 31.

Transactions are specified in terms of a dialog based on the relational algebra introduced by Codd[5]. This form was preferred over others such as the relational calculus[6] because of its procedural nature and the relative ease of implementation. All of the major system components have been introduced. Their use in actual planning procedures is illustrated in the next section.

Table 30. Network Editor Commands

COMMAND	FUNCTION
INVOKE NETWORK EDITOR	Begin edit session.
CREATE NETWROK OBJECT	A netwrok object is created.
DISCARD NETWORK OBJECT	Eliminate an object from system.
ADD OBJECT	An object not initially contained in the netwrok is connected to the network.
REMOVE OBJECT	An object contained in the network is disconnected from the network.
CONNECT NETWORK OBJECTS	Two objects contained within the netwrok are connected together.
DISCONNECT NETWORK OBJECTS	The connection between two network objects is broken; command illegal if this is teh only connection.
SET STATE VARIABLE	State variables are assigned values.
FIND OBJECT	A cursor is positioned to point at the object. A simple <u>FIND</u> can be done, knowing the object's unique name. More complicated <u>FINDS</u> can be done by specifying the surrounding context for the object.
DEFINE NETWORK CLASS	Create a new object class.

Distribution System Planning with the Model

One of the most important design goals of the system is to create a constructive problem solving environment. This consideration dictated that the system must interface with the user interactively; that the abstractions underlying the interactions must be familiar to the planner, i.e., the objects with which he deals in the system must be common electrical components; and that the dialects in which the planner communicates with the system must be simple and natural.

To demonstrate the extent to which these goals have been met, an example will be given of a network modification as it might be done using the DSPM[6].

Assume that it is required to determine if a given network can meet the projected short-range load growth. The load growth program is invoked by the execution of a shell command. One of the parameters of the command is the name of the network to be expanded. Another parameter is the version designator of the data base which contains the data required by the load growth program. Stored as part of the shell command is the query which, in addition to invoking the DBMS to retrieve the load growth program's input data, causes this data to be reformatted as a sequential file suitable for input to the load growth program. None of this underlying mechanism is normally visible to the user. When the shell command terminates, a new network has been created with the additional load that was forecast by the load growth program.

Table 31. DBMS Functions

COMMAND	FUNCTION
SELECT	Create a new relation from a subrelation of an existing one.
PROJECT	Attributes of one relation used to a new relation as a subrelation of another.
JOIN	A new relation is formed from two old ones with common attributes.
MINUS	Logical minus of two relations considered as sets.
UNION	Logical union of two relations considered as sets.
INTERSECT	Logical intersection of two relations considered as sets.
CREATE RELATION	Create a template with which to define a relation.
DESTROY RELATION	Remove a relation from data base.
INPUT RELATION	Read data from an input file into a pre-defined relation.
DELEGATE	Specify security status for a relation.
UPDATE	Modify the information in a relation.
LIST	Output the contents of a relation.

---

[5] E.F. Codd, "A Relational Model of Data for Large Shared Data Banks", <u>Comm</u>. ACM 13, 6 June 1970, pp. 377-397.

[6] E.F. Codd, "Relational Algebra", <u>Courant Computer Science Symposia</u> 6, "Data Base Systems", New York, May 1971, Prentice-Hall, New York, 1971.

[6] See Appendix G for more complete examples.

At this point, the planner may, using the network editor, inspect various network components to obtain a feeling for the probable effects of the increased load. Alternatively, he may simply elect to execute a shell command to obtain a voltage profile for the network. This command terminates with a list of network components which exhibit overvoltage or undervoltage conditions. If no such conditions are detected, the planner may then proceed to perform a reliability assessment. In the simplest case, this consists of executing a fuse coordination program.

If an abnormal voltage condition was detected by the voltage profile program, the planner may proceed to modify the network in order to correct the problem. Suppose for simplicity, that low voltage is indicated on one particular bus. The planner would first examine the neighborhood within the NETWORK, containing the problem bus. This is done by invoking the network editor, supplying the bus designator and specifying that a display is wanted of the network surrounding that bus. The display not only outputs the network topology but, in addition, the voltage and load levels of the components. From these data, the planner may determine the best way to correct the problem.

TASK B.7. THE DETERMINATION OF THE EXTENT TO WHICH EXISTING DIGITAL COMPUTER PROGRAMS, IMPLEMENTING THE CONCEPTS AND METHODS IDENTIFIED IN TASK B.6 ARE SOFTWARE COMPATIBLE

The system view which resulted in the planning distribution model described in the section on Task B.6 has eliminated the concern with which the present task was charged. This is because the Data Base Management System (DBMS) acts as an intermediary between all analysis programs. All input data is formatted by the DBMS so that it conforms to the requirements of the individual program and similarly, all output data from a given program is transformed into a canonical form before being stored by the DBMS. Thus, all that is required to satisfy compatibility conditions for any program is a description of its input and output formats. By construction then, all analysis software is compatible.

TASK B.8: THE SELECTION OF A SET OF EXISTING PROGRAMS WHICH BEST MEET THE NEED OF THE DISTRIBUTION PLANNING METHODOLOGIES AND ARE SUFFICIENTLY SOFTWARE COMPATIBLE TO PERMIT USE ON THE PROJECT

In TASK B.1 criteria were defined to aid in selecting models, both conceptual and implemented computer programs, to carry out the task of generating a production version of an interactive distribution planning tool. Programs that pass this criteria are eligible for the assessment.

Once a program or a model has passed the criteria defined in TASK B.1, selections among eligible programs must be made on criteria which are design oriented and/or ease of implementation considerations. The design considerations are the size of network that can be analyzed and nonrestrictive assumptions which allow more general use of the program. The implementation considerations exclude company-peculiar data entry and good documentation that facilitates transferability.

The functional system defined in the executive summary plus the models defined in TASK B.6 set up requirements which effectively eliminate current implemented systems planning models as suboptimal. This system forms a total economic cost package.

None of the above considerations which eliminated implemented programs that were assessed imply that the rejected programs were either incorrect or poorly conceived. They simply did not meet the design requirements posed by this research in TASK B.6.

The computer programs described in TASK A.1 and TASK A.2 did not meet all the implementation-related criteria and all the design and transferability-related criteria. However, some programs meet most of the criteria and therefore should be pointed out.

The computer program P1 exceeded the core storage requirements, as described in TASK B.1, even after extensive reductions in core storage have been made by overlay techniques. However, this program is in an interactive mode and performs a number of important analysis functions. It can be used to study distribution load flow, voltage profile and regulation, voltage flicker, capacitor allocation and power loss based on the system growth. Table 30 presents the results of the analysis and selection process.

Table 30. Computer Program Selection and Analysis

COMPUTER PROGRAM FUNCTION	PROGRAM	COMMENTS
Load forecasting		***
Primary network expansion		**
Substation expansion		**
Substation siting		**
Economic secondary distribution		***
Transformer load management	P8	*
Distribution load flow	P1, P14	P1=R, P14=*
Primary voltage selection		**
Voltage profile and regulation	P1, P6, P25	P1=R, P6=*, P25=*
Voltage flicker	P1, P6, P25	P1=R, P6=*, P25=*
Capacitor allocation	P1, P25	P1=R, P25=*
Power Loss	P1, P14	P1=R, P14=R=*
Protective device coordination	P20, P23	***
System reliability	P17	***
Total economic cost		**
Network editor		**

* = Needs interactive development.
** = Designed in this research.
*** = Needs to be developed.
R = Recommended for the function.

TASK B.9: THE ANALYSIS OF THE DATA REQUIREMENTS OF ALL PHASES OF THE DISTRIBUTION PLANNING PROCESS

INTRODUCTION

The purpose of this task was to identify all of the relavent data necessary for distribution system planning. In this section of the report, an overview of the required data will be provided so that when the detailed data base is presented in the section describing TASK B:11, the major outlines will still be visible.

Sources of Data

The three primary sources of data considered by the present analysis were the following:

(1) Information obtained by direct contact with utility companies.*
(2) Data requirements of computer programs available to this project.
(3) Data requirements of the conceptual models considered by this project.

Types of Data

The data considered can be grouped under two major classifications. Data on the distribution system itself, and planning data about forces and effects which may change the distribution system in the future. Distribution system data consist of inventories of hardware with accompanying status and topological or network data which defines the energy distribution system itself. Planning data consist of descriptions of future demographic trends and anticipated political action (e.g., zoning changes, bond issues, etc.) which will change distribution requirements from their present values.

Data Organization

Although there are many possible ways to analyze and organize distribution system data, the approach taken here was to concentrate on a hardware/network description. This decision was suggested by the fact that most distribution engineers tend to view the system in this way. The natural hierarchy within this organization consists of substation descriptions, primary distribution system descriptions, secondary distribution system descriptions, and network terminal descriptions. The description of hardware components does not vary too much from one component to another regardless of which element of the hierarchy, contains the components. Thus while the total amount of data is large, it can be organized in a logically parsimonius way by regarding it in this manner. In the following sections, each element of the hierarchy will be examined.

GENERIC COMPONENT DESCRIPTIONS

All of the electrical components of the networks can be described using the descriptors listed in Table 31. In this section, the significance of each entry in the table will be examined.

Organizational Description. Each utility has certain designations which it assigns to its equipment. This information would be contained here. e.g., substation number, division code, etc.

Network Data. In the section which discusses TASK B.10, the explicit form which network data takes for each component will be presented. Suffice it to say here that all the information necessary to place the substation in the network would be included in this descriptor.

Electrical Description. A complete electrical description of the component is provided by this descriptor.

Automation Code. Many network components are subject to having their parameters modified by automatic planning programs. This descriptor indicates what limitations in this regard are in force for the given equipment.

Operating Status. This descriptor indicates whether the component is presently in service or is undergoing repair, modification etc.

Service Code. This descriptor destinguishes between network components which are actually in the network, planned for the network or being experimented with, using simulation tools, by the designer.

Installation Data. Used for both engineering and accounting purposes, this descriptor tells the job number, the number of man-hours required, and last change required for installation.

Coordinates. Geographic coordinates; either map numbers or street address are in common use.

Manufacturing Data. Typical information provided by this descriptor would include manufacturer, model type, serial number and a brief functional description.

Physical Description. Physical characteristics, such as dimensions, weight, etc. are described.

Reliability Data. Data useful for estimating mean time between failures.

Cost Data. Installation costs, replacement costs, depreciation, interest, maintenance expense, taxes, insurance, and in some cases, operation costs are provided by this descriptor.

Record Status. This descriptor relates to the data base rather than the component. It reflects the currancy of the information, data of last update and who among the user community has access to the information.

Time History Data. Energy use as a function of time is required for some equipment; e.g., customer meter records. This descriptor captures that data.

Table 31. Generic Component Descriptors

Organizational Description
Network Data
Electrical Description
Automation Code
Operating Status
Service Code
Installation Data
Coordinates
Manufacturing Data
Physical Description
Reliability Data
Cost Data
Record Status
Time History Data

---

*Oklahoma Gas and Electric Company was particularly helpful in releasing detailed information on their data base.

With the generic description defined, attention will be focused next on the different elements in the distribution system hierarchy.

SUBSTATION DATA

Physically and electrically speaking, the typical substation consists of the components shown in Table 32.

Table 32. Substation Components

- Transformers
- Breakers/Relayed Reclosers
- Bus Configuration
- Vacuum Switches
- Manual Switches
- Capacitors
- Reactors

In our view, there is a description of the substation based on the generic descriptors of Table 31, just as there is such a description for each component shown in Table 32. The detailed description is presented under TASK B.10.

As a brief illustration of how the generic description applies to the substation, consider the descriptor Electrical Description. For the substation, it is Table 32 suitably expanded to include the detailed properties of each such component, as well as total capacity, existing load, etc. The descriptor, Network Data, describes the bus configuration within the substation. The descriptor, Physical Description, supplies service area size, getaways and so on.

The applicability of the generic description is just as comprehensive for the other elements in the hierarchy as is discussed below.

LINE DATA

Line data can be separated into primary and secondary distribution catagories. In this overview, however, that will not be done. Instead, the listing of device types shown in Table 33 will be presented.

Each element shown in this table can be described using the generic description. Since detailed expansions for each of these devices in given in the sequel, no further discussion will be provided here.

Table 33. Line Section Elements

- Line Section
- Distribution Transformer
- Step-up/Step-down Transformer
- Boost and Buck Transformer
- Shunt Capacitor
- Series Capacitor
- Step Regulator
- Induction Regulator
- Line Sectionalizing Device
- Line Recloser
- Static Recloser
- Line Reactor
- Tie Sectionalizing Device

METER DATA

Meters represent the termini of the distribution network. The meter is the component closest to the customer and provides a measure of consumption on a per customer basis. As are the other major classes of network elements, a meter associated with an active customer is described by the generic descriptors of Table 31.

The time history data contained in the meter gives a consumption for any major section of the distribution system. Since these sums directly correlate with geographic areas, a measure of consumption as a function of area may be calculated. This consumption function is important for use by the various planning programs which attempt to predict load growth in given areas of the network

PLANNING DATA

The planning data, as it is presently collected and used in practice, is the only signigicant portion of the data that is not explicitly network-oriented. Some preliminary research concerned with relating even load growth forecasts directly to a network model has been attempted by this research group but these efforts are at such an early stage that any consideration of them in this report would be inappropriate.

Planning, as it is currently practiced, requires energy consumption data, which, as discussed above, is carried on individual equipment in the network. In addition, it requires data on the geographical area in which the distribution system is located. This data, compiled for every element of a grid system covering the system is shown in Table 34.

Table 34. Planning Data

- Land Use Plans
- Planning Code
- Saturation Code
- Class Code

These data are explained below.

Land Use Data. This category includes community environment factors such as developer's plans, zoning ordinances, and the attractiveness of the area for further development.

Planning Code. This datum is a measure which is established by metropolitan planning commissions and designates twenty-six different growth patterns and rates.

Saturation Code. Assigned by the planning engineer, this number indicates how much "room" exists in a grid element for further growth.

Class Code. Also assigned by the planning engineer, the class code is computed based on a projection of future demand by each present customer in the grid.

SUMMARY

The object of this task was to examine the data classes pertinent to the distribution system planning process. This examination yielded a generic description which applies to all of the data related to a distribution network. The only data which is not subsumed under this classification scheme is the planning data which must be treated separately. The generic description will be expanded in the section describing TASK B.11 to fully describe the complete data base required by the distribution planner.

TASK B.10: THE CONCEPTUAL DESIGN OF A DATA BASE SUFFICIENT TO MEET THE DATA REQUIREMENTS OF A DISTRIBUTION PLANNING METHODOLOGIES DEVELOPED

INTRODUCTION

The purpose of this section is to explain the conceptual basis for the detailed database that will be presented in the section of TASK B.11. Actually, this section will discuss considerably more than just the ideas underlying the database, since it is important to explain the concepts underlying the other major components of the methodology identified by this research.

For this reason, the present section is divided into subsections, the first three of which discuss issues directly related to the database. The last one discusses a program called the Network Editor. This discussion has some thoughts in common with that found under TASK B.6; however the point of view rests firmly on the foregoing discussion of the database and thus is able to clarify a number of points which could not be adequately covered in the treatment of TASK B.6.

The introduction to databases begins with the inspection of the state-of-the-art.

Present State of Database Technology

Today's Database technology can be described in terms of three distinct data organizations: the relational model, the network model and the hierarchical model[1,2].

The relational model, which is the scheme chosen as the basis of the conceptual model for this study was introduced by Codd[3]. It may be developed by considering attributes to be identifiers taken from a finite set $A_1, A_2, \ldots, A_n$. Each A_i has associated with it a set of values called a domain, written as $dom(A_i)$. A relation on the set of attributes, $R(A_1, A_2, \ldots, A_n)$, is a subset of the Cartesian product

$$dom(A_1) \times dom(A_2) \times \ldots \times dom(A_n).$$

An element of this subset $(a_1, a_2, \ldots, a_n)$ is called a tuple. A relation may be simply visualized as a table of rows and columns. The rows represent tuples, while columns represent the values of a particular attribute contained in the relation. An example is shown in Table I, where partial data on transformers of a particular class have been used to construct a relation. A database consists of one or more relations. Keys are attribute values which uniquely specify tuples. For example, in Table 34 the serial number is a key because no two transformers have the same serial number. In some relations, more than one attribute must be specified to obtain a unique tuple.

Table 34. The Relation Transformer

Transformer:

SERIAL #	X-Y COORD	VOLTAGE RATING (LINE/LOAD)	POWER LOSS (NO LOAD/FULL LOAD)	PORTS (LINE SECTION/SOURCE)
17324816	296-015	12.5/2.4	140/210	448/12715
59371140	112-431	12.5/2.4	140/210	321/5521
74266312	030-030	2.4/3.6	250/400	25/4374
46164229	075-340	12.5/0.22	20/30	3147/2745
24279531	215-009	12.5/214	140/210	440/4493
17764697	317-445	12.5/0.22	20/30	3147/12772
33473748	075-110	2.4/3.6	250/400	312/5541
00277410	075-200	2.4/3.6	250/400	312/5595

The oldest form of database organization or data model is the hierarchical model. The concept is illustrated in Figure 32 where an employee database ordered by organization level is shown. The major levels in the hierarchy are university, colleges, departments, employee classification and employee. The key or pathname uniquely identifying Professor Jones in the Math Department is:

university.arts&science.math.faculty.jones.

stored under this key might be the professor's salary, number of dependants, social security number and so on.

The shape or topology of the storage structure is that of a tree. Each of the interior nodes (e.g., those on the levels university, colleges, departments, and employee classification) can have any number of descendants. Note however, that in the hierarchical scheme, if a professor holds an appointment in more than one department, the data model has serious problems since the structure to represent that is no longer a tree.

The network model was introduced to generalize the hierarchical model and eliminate the above difficulty*. Figure 33 illustrates a network data model for a collection of authors, (Fezziwig, Wilkins, Crachet, and Dickens), scholarly papers, (denoted as PF1, PF2, PW1, etc.) and Journals in which these papers appeared (CACM, BIT, EBL). Consider the succession of arrows shown in Figure 33 which begin and end with Fezziwig. They serve to associate the papers PF1 and PF2 with Fezziwig. If the circuit which includes the bottom portion of PFL is examined, it will be seen that it also contains PW1 and the journal designator CACM. Thus it relates the papers PF1 and PW1 with the journal CACM. In like manner, each of the other papers is related to both an author and a journal. At the same time, a single author is related to many papers and a single journal is related to many papers. It is the ability to represent the many-to-many relationship which makes the network data model more flexible than hierarchical model.

Before justifying the choice of one of these models over the others, one of the key concepts of the proposed methodology must be examined. Only after this discussion, will it be possible to evaluate all of the factors to be considered.

The Network as the Fundamental Basis for the Methodology

One of the conventional ways of describing distribution systems is via a system of geographical grids. The methodology and technology to be discussed here is based on a different representational approach. The most prominant feature of a distribution system is that its components can be represented as elements of a graph. Specifically, each element may be considered to be a node in a graph. This includes line sections and other components which might otherwise be thought of as constituting edges in the network graph. If all the physical components of the network are considered to be nodes, then the edges serve merely to associate network elements with each other as the topology of the physical network requires.

Mathematically speaking, a graph consists of a 4-tuple $G(N,E,s,t)$, where N is a finite set of nodes, E

---

[1] Date, C.J., An Introduction to Database Systems, 2nd Edition, Addison-Wesley, Reading Mass., 1977.

[2] Fry, J.P., and Sibley, E.H., "Evolution of Data Base Management Systems", ACM Computing Surveys 8, No 1, (1976), pp. .

[3] Codd, E.F., "A Relational Model of Data for Large Shared Data Banks", Comm. ACM 13, No. 6 (1970), pp. 377 - 397.

*The use of the term "network model" for a data model is unfortunate since in later discussion, the term "network model" will refer to a conceptual scheme central to our preferred methodology. The nomenclature is quite well-established in the database literature however, and is the only one appropriate.

is a finite set of edges, s is the source map

s: E→N

which maps each edge e ∈ E onto an element n ∈ N. Saying it in an alternative way, s(e) is the node in N from which the directed edge e originates.

t is the target map

t: E→N

which defines on which node a given edge terminates. Consonent with the mathematical description, each physical component of the network must have source ports and target ports. In general, target ports will be at a higher electric potential than source ports.

Figure 34(a) shows the conventional description for part of a distribution system and Figure 34(b) shows the equivalent graph representation.

For every node in the network, there is a set of generic component descriptors* which can only be specialized to describe that particular component type. Thus, this network model and the accompanying component descriptors serve to describe the entire distribution system.

CHOICE OF THE RELATIONAL MODEL

As discussed above, there are three well-known data models in use. Upon what basis should one make a decision favoring the adoption of one over the others? The answer lies in the form of the data to be processed. Other considerations being equal, the model which "naturally fits" the data is the one which should be chosen. In the present case, each model will be superposed on the data to determine which should be the most satisfactory.

Non-hierarchical Nature of the Data

If the totality of data required to describe the distribution system is considered, it is difficult to discern a hierarchy, particularly if one adheres to the directed graph conception. This itself is a many-to-many relation and as such is not amendable to being cast in a hierarchical mold.

Even if one elects to represent the distribution system with two different data models--one for the graph description and the other for the component descriptors--it is not clear that the hierarchical model has much to offer. Network components are not usually thought of as being subordinate to each other and since it is the subordinate relationship which must be present for a hierarchical view to succeed, this model does not appear feasible.

Difficulties with the Network Data Model

One might be inclined to think that because the conceptual model of the distribution system is that of a graph (network), the correct data model would be the network model. In some database schemes for distribution system planning, this is indeed the case[4] but the choice is difficult to justify based on the most current technology.

The basic problem is the complexity of the specification description. To retrieve a given record of information, one must essentially traverse the network, until arriving at the data required. In the simple scheme portrayed in Figure 33 this would present no real problem but for retrieval and update in a system with realistic complexity, it is felt that the system imposes an unnecessary burden on the engineer or other user who considers the database and its database management system to be just a tool to support his work.

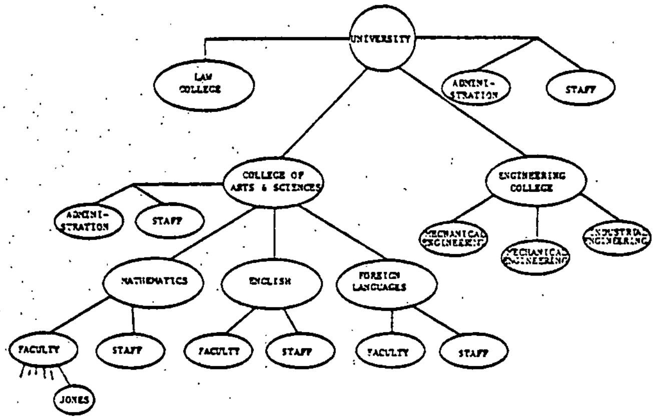


Figure 32

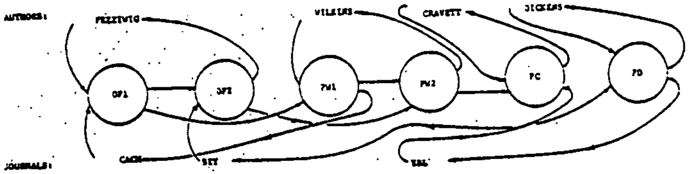
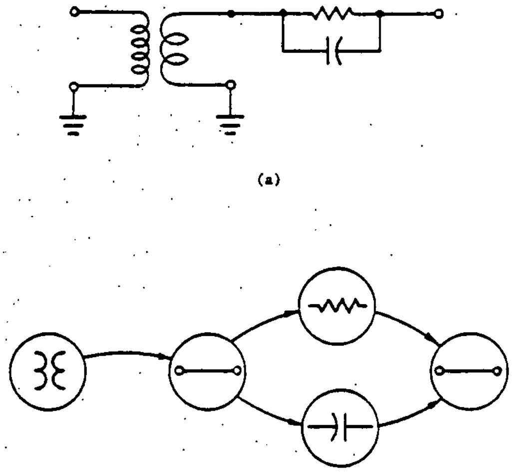


Figure 33


(a)

(b)

Figure 34

---

*See the previous section describing TASK B.9 for the definition of component descriptors.

[4] Fagan, J.E. and O'Dell, M.D., private communication on the Oklahoma Gas and Electric Company Database Management System, Spring, 1979.

THE RELATIONAL ALGEBRA

Accompanying the relational model are two language vehicles for encoding queries and other operations against the database. Both were introduced by Codd[5] and have since been widely accepted. One language is called the relational algebra and the other, the relational calculus. Because there is considerable difference in the level of difficulty encountered in implementing one of these relative to the other and since this affects the implementation efforts discussed in the section on TASK B.12, the distinction between the two will be drawn below.

Relational Algebra vs. Relational Calculus

In the most fundamental sense, relations are sets. To specify subrelations (subsets) of relations (sets) there are, therefore, two approaches: either specify a set of operations which can selectively be used to extract the subset of interest or define the desired subset by stating all of the constraints which distinguish the subset from the set in which it is contained. The first approach is the basis for the relational algebra; the second approach is the basis for the relational calculus.

A complete set of operations for the relational algebra is to be found in Appendix F. The discussion presented here will consider only three of the most important.

Selection. Selection specifies a subset of tuples within a relation for which a particular predecate is true. For example, performing a selection over the relation TRANSFORMER of Table 34 for the predicate

VOLTAGE RATING = '12.5/2.4'

yields the relation shown in Table 35.

Table 35

SELECT(TRANSFORMER.VOLTAGE RATING= '12.5/2.4'):

SERIAL #	X-Y COORD	VOLTAGE RATING (LINE/LOAD)	POWER LOSS (NO LOAD/FULL LOAD)	PORTS (LINE SECTION/SOURCE)
17324016	296-015	12.5/2.4	140/210	442/12715
59371148	112-431	12.5/2.4	140/210	321/5521
24279531	215-009	12.5/2.4	140/210	448/4493

Projection. The projection of a relation $R(A_1, A_2, \ldots, A_n)$ over the set of attributes $\{A_i, A_j, \ldots, A_k\}$ is the relation

$Rp(A_i, A_j, \ldots, A_k) \subseteq R(A_1, A_2, \ldots, A_n)$.

TRANSFORMER(SERIAL#, X-Y COORD, VOLTAGE RATING, WATTS LOSS, PARTS)

shown in Table 34 over the set of attributes

{SERIAL#, VOLTAGE RATING, WATTS LOSS}

gives the relation shown in Table 36.

Table 36

P = PROJECT(TRANSFORMER,SERIAL#,VOLTAGE RATING,POWER LOSS):

SERIAL #	VOLTAGE RATING (LINE/LOAD)	POWER LOSS (NO LOAD/FULL LOAD)
17324816	12.5/214	140/210
59371148	12.5/214	140/210
74266312	2.4/3.6	250/400
46164229	12.5/0.22	20/30
24279531	12.5/214	140/210
17764697	12.4/0.22	20/30
33473748	2.4/3.6	250/400
00277410	2.4/3.6	250/400

Join. The operation join is used to make a connection between attributes that appear in different relations. Let $R(A_1, A_2, \ldots, A_n)$ and $S(A_1, B_2, \ldots, B_m)$ be two relations. Then

JOIN(R,S) FOR R.A = S.A. =

$\{(a_1, a_2, \ldots, a_n, b_2, \ldots, b_n) \mid (a_1, a_2, \ldots, a_n) \in R(A_1, A_2, \ldots, A_n)$ and $(a_1, b_2, \ldots, b_m) \in S(A_1, B_2, \ldots, B_m)\}$

For example, JOIN(P,COST) FOR P.SERIAL# = COST.SERIAL# for the relations in Tables 36 and 37 respectively, yields the relation of Table 38.

In contrast to the relational algebra, the relational calculus does not consist of a set of operations but instead is comprised of statements which define the qualifications which the required relation must satisfy.

The general form for the query is

GET ⟨new relation name⟩ (⟨tuple specification⟩): ⟨attribute predicate⟩

where*

⟨tuple specification⟩ ::= $R_1.A_1$, $R_2.A_2$,111

⟨attribute predicate⟩ ::= expression involving the logical operators >, >=, =, ≠, <=, <, and, or, not and operands, $R_i.A_i$

It is understood that A_i is an attribute name of relation R_i.

As specific examples of relational calculus statements, the statements equivalent to the operations illustrated for the relational algebra will be given.

Table 37

COST

SERIAL #	$-COST
74266312	5000
24279531	3500
33473748	5000
17764697	1500

Consider first

GET S (TRANSFORMER.SERIAL#, TRANSFORMER.X-Y COORD, TRANSFORMER.VOLTAGE RATING, TRANSFORMER.POWER LOSS, TRANSFORMER.PORTS):TRANSFORMER.VOLTAGE RATING '12.5/2.4'

---

[5] Codd, E.F., "Relational Algebra", Courant Computer Science symposia 6, "Data Base Systems", New York, May 1971, Prentice Hall, New York, 1971.

*The symbol "::=" may be read "is defined to be".

This statement generates the relation shown in Table 35. Notice that the attribute names i.e., SERIAL#,X-Y COORD, etc., specified in the parentheses are the attributes which appear in the new relation s (the ⟨tuple specification⟩ is a paradigm for tuples of the new relation). The ⟨attribute predicate⟩

```
   TRANSFORMER.VOLTAGE RATING = '12.5/2.4',
```

acts just as the predicate in the selection operation does.

A calculus statement which produces the relation shown in Table 35 is the following:

```
   GET P (TRANSFORMER.SERIAL#, TRANSFORMER.VOLTAGE
         RATING, TRANSFORMER.POWER LOSS)
```

This statement has no ⟨attribute predicate⟩. This means there is no qualification which the attributes in the paradigm must meet. Thus P contains all such tuples.

The join example which produced Table 38 can be effected by means of the statement

```
   GET Table 38 (P.SERIAL#, P.VOLTAGE RATING, P.POWER
                LOSS, COST.S-COST):  P.SERIAL#
                COST.SERIAL#
```

Table 38

JOIN(P, COST):

SERIAL #	VOLTAGE RATING (LINE/LOAD)	POWER LOSS (NO LOAD/FULL LOAD)	S-COST
74266312	2.4/3.6	250/400	5000
24279531	12.5/214	140/210	3500
17764697	12.5/0.22	20/30	1500
33473748	2.4/3.6	250/400	5000

As a final example of the power of the relational calculus, suppose it is required to determine which of the transformers listed in Table 34 have been in service less than six months. Also, the coordinates of such transformers are required. Assume that in addition to Table 34, Table 39 is available, showing transformer serial number, status (in or out of service) and date when transformer first acquired this status. Using an existential quantifier and a RANGE statement, the query may be formulated as

```
   RANGE  SERVICE X
   GET NEW (TRANSFORMER.SERIAL#,TRANSFORMER.X-Y COORD):
           ∃(X.SERIAL# = TRANSFORMER.SERIAL# and X.STA-
             TUS DATE - CURRENT DATE < 180 and X.STATUS
             = 'In')
```

This query may be understood as follows. The RANGE statement specifies that the variable X is to range over all tuples in the relation SERVICE (Table 39). A new relation named NEW (Table 40) is to be created with attribute names SERIAL# and X-Y COORD. The conditions which the tuples in TRANSFORMER, from which these values are to be drawn are:

(1) There must be a tuple in SERVICE which has the same SERIAL#.

(2) The arithmetic difference of STATUS DATE less CURRENT DATE must be less than six months (180 days).

(3) The transformer must be in service.

Assuming that CURRENT DATE is 79300, the resulting relation is shown in Table 40.

To obtain the same relation NEW, using the relational algebra would require the following sequence of operations.

```
   JTS = JOIN(TRANSFORMER,SERVICE) FOR TRANSFORMER.
         SERIAL# = SERVICE.SERIAL#

   SLT = SELECT JTS WHERE SERVICE.STATUS = 'In' and
         SERVICE.STATUS DATE - CURRENT DATE < 180

   NEW = PROJECT SLT OVER SERIAL#,X-Y COORD
```

This series of operations can be combined into one statement as

```
PROJECT(
        SELECT(
               JOIN(TRANSFORMER, SERVICE)
               FOR TRANSFORMER.SERIAL#=SERVICE.SERIAL#)
        WHERE SERVICE.STATUS='In' and
        SERVICE.STATUS DATE-CURRENT DATE<180
OVER SERIAL #, X-Y COORD.
```

Table 39

SERVICE

SERIAL#	STATUS (In or Out)	STATUS DATE (Julian)
17324816	In	75001
59371148	In	79250
74266312	Out	79297
46164229	In	71155
24279531	In	69042
17764697	In	76091
33473748	Out	79297
00277410	In	65300

Table 40

NEW

SERIAL#	X-Y COORD.
59371148	112-431

It is clear, however, that the Relational Algebra statements even when combined as above, are procedural in nature rather than specification statements.

Implementation Difficulties

Though the features of the relational calculus are highly desirable because of their power and conceptual simplicity, the question of their implementation is quite another matter. In fact, as far as is known, no database management systems have fully implemented the relational calculus (such, work is preceding but the technical problems are severe[6,7]).

For this reason, all of the implementation effort associated with this project centered around a relational model based on the relational algebra. The details of this effort will be found in the section on TASK B.12. More involved examples using the relational algebra for distribution planning will be found in Appendix G.

The last topic which rightfully belongs with the conceptual underpinnings of the research effort is the

---

[6] Stonebraker, M., Wong, E., Kreps, P., "The Design and Implementation of INGRES", ACM Transactions on Database Systems 1, No. 3 (1976).

[7] Palermo, F.P., "A database Search Problem", Information Systems: Coins IV (ed., J.T. Tou), New York, Plenum Press (1974).

work to develop a program model for manipulating network objects. This model is considered next.

THE NETWORK EDITOR

The Network Editor is one of the important conceptual tools in the methodology espoused by this report. In fact, the system as it is seen here consists of these major components: The applications/analysis programs, the Data Base Management System, the Network Editor and the Shell program. The system view which places each of these in the proper perspective has already been presented in the section on TASK B.6. In this part of the Report, a fuller description of the Network Editor will be given with an emphasis not so much on its operational capabilities but rather on its conceptual bases.

The Need For Such A System

As was explained above in this section, one of the chief concepts of the methodology is that of a network. The intention is that since the engineer/planner is used to thinking in terms of network concepts, the dialogues with the planning system should be in network terms. For this reason, two ideas were amalgamated: a vehicle for manipulating network objects combined with concepts from the single most successful interactive computer tool available, the text editor. Therefore, the intent behind the Network Editor is to create a tool which figuratively speaking, feels natural in the planners hand, which allows him to create or alter networks as easily as text editors allow their users to manipulate text.

Another important reason for wanting to include the Network Editor in the methodology is just the sheer diversity of network objects and the complexity of their interconnections. The designer's representational abilities must span the range from simple components such as individual transformers or customer meters to entire substations which consist of subnets. For this task, one would like to have a conceptually elegant and simple to use tool.

To understand the Network Editor as it is proposed here, it is first necessary to examine the concepts upon which it rests.

Guiding Conceptions

There is an important distinction between primitive network objects and compound network objects. Descriptions of the former require no subnet information. Thus the individual transformer or capacitor or even line element is not comprised of a network but is to be found in a network i.e., is atomic from the viewpoint of the network being described. On the other hand, a substation, consisting as it does, of a collection of transformers, busses, meters, etc., can only be described by the inclusion of network information. For this reason, it must be considered a compound object.

As has already been described briefly, a network is modeled by means of a directed graph whose nodes or vertices are physical components and whose edges are descriptors which define how the physical components are interconnected. Edges connect one node to another via ports. Ports may or may not have an internal structure of their own. If a feeder carries three phases, the source and load ports will be described by triples, each element of the triple being referred to as a "pin." In some cases, a single multiport may join two or more ports with a smaller number of pins. For example, in a urban residential neighborhood, a three-phase feeder may divide into three single phase feeders, each of which serves say, twenty-five families. Thus an edge is defined not only by the source and target functions described above, but also by the type of port found at each end of the edge. In this way some integrity constraints may be built into the system, preventing the user of the Network Editor from joining network objects with incorrectly specified ports.

In a graph theory sense, each edge has a color as do most nodes. Some colors represent single phases. Let us say that red, green, and blue each represent a distinct phase while white arbitrarily represents a three-phase component. In addition, define the "sum" of the colors red, green and blue to be the color white and the sum of any two of the colors to be a color which is not one of any of the four. Junction rules for the way in which ports can connect may be expressed very simply (for most components): an edge with a given color must connect two network objects with that same color. In cases where there is a separation of phase (as in some residential areas) then the "color sum" must be constant on either side of the junction. Thus a white feeder can meet three subfeeders so long as one is red, one is green and one is blue. In this way the Network Editor prevents any phase confusion.

Some network components have their non-topological parameters completely defined by constants. For example, a capacitor may have its rating fixed at 20 kVA. Typically however, many network components have one or more variable parameters; the tap settings on a transformer may be variable; the level for a voltage regulator may be adjustable. This means that specifying the class to which such a network object belongs does not specify the values for its variable parameters. One way these parameters may acquire their values is to have them defined when the object, an instantiation of an element from an object class, is created. The other way these values may be set is for them to be associated with state variables. Such state variables may assume integer, floating point or Boolean values. More than one object may have its parameters specified by the same variable. Thus the designer may alter the setting of several devices by changing the value of a single state variable.

State variables are clearly useful for defining the settings of switches. A generalization of this notion however, leads to an even more powerful idea, that of design switches. Suppose it is required to consider the effects on performance of making some relatively minor modifications to a network. For example, suppose there are several alternative locations in the network for a combination of voltage regulators and capacitor banks. The question to be answered is which combination at which locations, gives the best performance. The straightforward way to determine the answer is to make a list of all the possibilities to be tested and then implement the network corresponding to each in turn. A more elegant solution makes use of the notion of a design switch. At the time an object is created (becomes the instantiation of an element from an object class), a state variable can be assigned as a design switch. When the switch is "on," the object appears in the network just as any other normal object. When the design switch is "off." the object is virtual and for all application purposes, disappears. Thus by turning design switches off and on, the configuration of the network can be greatly varied. In the above example of capacitors and voltage regulators, these objects may be placed in the network all at one time to be controlled by design switches. By successively turning switches off and on the various alternatives may be created and analyzed by the applications programs.

Another important concept incorporated into the Network Editor is that of the cursor. A cursor or pointer is used as a variable in a program to allow the planner to refer to various elements in the network. When, for example, an object is created, a cursor variable is associated with the object. The operations to be described below all use cursors to refer to the objects upon which they operate.

Underlying the notion of an editor for a network is the fact that networks can be described in terms of text strings. These strings, which are known as K-formulas, permit one to visualize some network operations as operations on a text string.[8] In particular, one may specify some particular collection of network components and invoke the Network Editor to find all places in the network where this configuration is present. Figure 35 illustrates the general situation. The subnet shown in the insert is sought in the larger network. When it is found, it may be replaced by another network configuration, just as one text string may be substituted for another by a text editor. For example, suppose one has identified a particular subnetwork as something to be represented by a compound object. Using context searching, the Network Editor can locate all locations where the subnet appears then replace the collection of network objects comprising it with compound objects. This might be done to study the variation of several parameters in the compound object and could be done more easily this way than by direct variation of all occurrances of the subnet in the original network.

The last important concept associated with the Network Editor is the relation of network objects to entires in the database. This problem is solved by treating a network as a relation. The relation has attributes corresponding to name, object class, and a list of connections to other objects in the network. A primitive object is just a tuple in such a relation, while a compound object is represented by an entire relation.

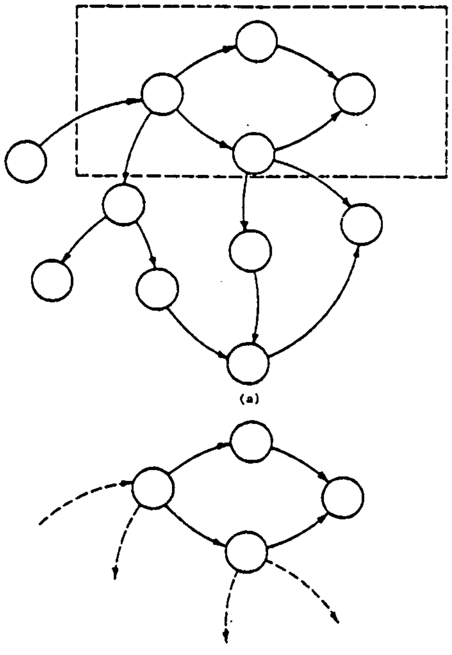
The commands which the Network Editor accepts are listed in detail in Appendix F, Section 3.3.3. An abridged description of them also appears in the section discussing TASK B.6, Table 30). No further discussion of the operations will be given here. Instead, consideration will be given to the last major component of the system representing the planning methodology, the Shell.

THE SHELL PROGRAM

Every system has its executor or monitor. For the Distribution Planning System, that overseer is called the Shell. Its general purpose is to serve as a command line interpreter to which the user directs his commands and inquiries. In effect, however, it is a miniature operating system, implementing some of the more basic functions traditionally associated with operating systems. Some of its features have been borrowed from the Unix operating system developed at Bell Laboratories.[9] In what follows, the basic notions incorporated into the Shell will be discussed.

Purpose of the Shell

In concept, the Planning System is to be highly portable. However, standing in an almost dichotomous position to this requirement are the necessities for create files, execute programs in fixed sequences, monitoring exceptional conditions and supporting several interactive users simultaneously. These latter functions conventionally have been assumed by operating systems, which logically speaking exist at a lower level than so-called user or applications programs.



(a)

(b)

Figure 35

When one considers raising what have generally been operating system functions to the applications program level, as the methodology suggested here proposes, the first question to be examined is whether that is even possible. The answer which one can make to this question is weakly dependent on time--our views concerning operating systems ten years ago were much different than they are today--but to be objective, any answer given must be in the present tense.[10] In consequence, the answer today is that it depends on the native operating system on top of which, the Planning System is to run. All of the prototype system implemented to date has been implemented on a Unix-based system (TASK 3.12 of this report describes the implementation effort in detail). As a result, few actual difficulties have been encountered. Implementing the system on an operating system native to IBM 360 or 370 machines would be considerably more difficult might not even be possible unless performance goals were greatly compromised. Thus while portability remains a laudable ideal, it appears at present to be an unreachagle one. Conversely, it is expected that in the next five years, as general understanding of what constitutes a hospitable user environment becomes more widespread, that implementation difficulties which today might exist for an arbitrarily chosen operating system will be greatly overcome and this time scale is sufficiently short

---

[8] Thompson, J.C. and Atkins, G., "The Use of Generalized K-Formulas for the Syntactic Description of Data Structures," submitted to a technical journal.

[9] Ritchie, D.M., and Thompson, K., "The UNIX Time-Sharing System," Comm. ACM, 17, (1974), pp 365-375.

[10] The greatest change in thinking about operating systems has been due to the impact that Unix and Multics have had on the computer science community. See Feiertag, R.J., and Organick, E.I., "The Multics input-output System", Proc. Third Symposium on Operating Systems Principles, Oct. 18-20, 1971. ACM. New York, pp. 35-41.

that considerations expressed here might have a timely effect on emerging technology. Seen from this view, the Shell represents a collection of functions which the distribution planner needs to have available and which form an umbrella covering the other system components discussed in this report.

Shell Functions

Each user is known to the Shell and as a result to the Data Base Management System via a userid. This is an alphanumeric string which serves to associate all files, relations, networks, and database with the user. The Shell authenticates a userid by receiving a password from the user. Presenting the proper password allows a user to receive all of the privileges associated with his userid. If his userid is that of the Database Administrator for a Database then all of the commands reserved for the DBA are available for his use.

The Shell allows the user to create files and remove files which are associated with his userid. This file processing is done in such a way that the user is isolated from the normal file handling features of the native operating system. Thus, the user is not burdened with learning the syntax which accompanies file processing on the native system. He has only to learn the simple syntax accepted by the Shell.

One of the most significant problems solved through use of the Shell mechanism is that of software compatibility with respect to the applications/analysis programs. Since these programs were written by different individuals at different companies and institutions, there is no uniformity of format for input and output data. In some cases, the same data is required by different programs but in different units! Two approaches suggested themselves. In one, the plan would be to modify the internals of each program so that data requests were programmed as calls to the DBMS. This clearly would be a difficult task, requiring the intimate understanding of the internals of each program. The other approach required the use of a filter which would be placed between the program and the DBMS. The filter would function as a very sophisticated FORMAT statement specifying queries to the database and conversion of data from the form stored in the database to a form compatible with the requirements of the program. The process is represented in Figure 36. A request to the Shell by the user for the execution of a program results in the identification by the Shell of the proper filters to be used. The filter program queries the database and creates a file which matches the demands for the input data to the applications program. The Shell then causes this program to execute using the input file created by the user. The output from the applications program is collected on a file which is run through a second filter program which reformats the data into a form acceptable to the database. The only data necessary for this filtering to take place are the specifications on the formats for input (output) to (from) the applications program. These specifications are of course part of the database. They are stored in a relation owned by the Shell itself.

In general, the situation will be more complicated, since the actual sequence which the planner will often want to execute will involve using the output of one program as the input or partial input to succeeding programs. It is clear, however, that the same principle is involved, only the details are more complicated. For such a case, a number of filters are required. The situation is reflected in Figure 36. The concept is an adaptation of the notion of pipes, found in Unix.[11]

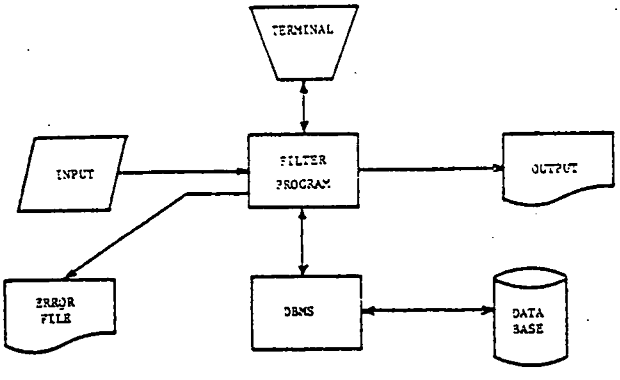


Figure 36

The other important feature borrowed from Unix is that of asynchronous command processing. Processes which consume considerable processor time may be allowed to run "in the background," while the planner submits other commands to the Shell.

The functional descriptions of all the Shell commands are to be found in Appendix F. The interested reader will find there a number of commands which have not been discussed here.

SUMMARY

This section has discussed each of the major systems upon which the planning methodology rests. It differs from the information in the section on TASK B.6, since there an effort was made to outline a systems view of the entire methodology. In this section, the emphasis has been placed on explaining the concepts underlying each of the systems.

The Data Base Management System has as its data model, a relational scheme. Its query language is based on a variant of Codd's relational algebra. The relational algebra was preferred over the relational calculus because of implementation difficulties.

The Network Editor, perhaps the most innovative of the systems envisioned by this project, is an interactive program which allows the planner to deal with networks in a manner similar to the way in which one deals with word processing using a text editor. One may define, modify or delete networks using simple commands. Of particular power are the features which permit context searching in the network, modification of object status using state variables and the forming of compound objects from simpler network components.

---

[11] Ritchie, D.M., and Thompson, K., ibid., p. 370

The last major system is the Shell which makes available to the planner many of the features of a real-time operating system. In principle, these features would be independent of the machine on which the Planning System were running and would provide a fully portable basis for executing planning programs.

TASK 3.11: A DETAILED DESIGN OF THE DATABASE CHARACTERIZED IN TASK 3.9

INTRODUCTION

The database presented in this section is divided into two major parts: one consisting of network data and the other consisting of planning data.

The network data is comprised of fourteen tables. The first thirteen describe a particular component of the distribution system network and the last describes the substation. Each such table is to be considered as a relation. All of the tables together form the network database. The planning data is presented in Table 55. All of the data is correlated with planning activities in Figure 41.

Basic Assumptions

In agreement with the discussion given in the section on TASK 3.10, it is assumed that each network component (Tables 41-54) forms a vertex in the system network graph. To realize this, each component has associated with it Network Data which establishes its place in the network. As a result, there is no separate table (relation) which serves as a network topology description. The topology description has been "distributed" among the network components. If the planner requires a consolidated topological description, then it may, of course, easily be obtained but it is not retained in the database as a separate entity.

In the same way, other data such as some cost data have been distributed throughout the network and associated with individual pieces of equipment. This results in a smaller but no less complete database configuration.

Two other assumptions deserve mention, both concerned with the boundaries of the network. The first is that the Distribution System begins with the breakers, reclosers, capacitors and reactors on the low voltage side of the substation and does not include the low voltage bus bar network of the substation. This view which is not universally held within the industry was adopted in part because the bus network did not conform to the generic descriptions proposed in TASK 3.9.

The second assumption made regarding the extent of the distribution system was that the network description should include the meter at the customer site. This convention, which is also not standard within the industry, was made so that consumption data would also fit into the network description of the system.

The complete database is presented in the next two sections. To preserve the outline sketched in the material on TASK 3.9. generic descriptors are included in all the tables as subtitles. Explanations of terms not thought to be part of the industry's standard vocabulary are explained in footnotes placed at the bottom of the tables in which the terms are found.

NETWORK DATA

Table 41. Data Description of a Line Section

- Device Identifier
    - Line Section Identifier
- Organizational Description
    - Division Code
    - District Code
- Network Data
    - Source Identifier
    - Load Identifier
    - Circuit Number
    - Substation Identifier
- Electrical Description
    - Phase Code
    - KVA Load ABC (Winter)
    - KVA Load ABC (Summer)
    - KVA Demand ABC (Winter)
    - KVA Demand ABC (Summer)
    - Voltage Limit Code
- Automation Code
- Operating Status
- Service Code
- Installation Data
    - Wire Installation Data
    - Job No.
        - Orig. Install.
        - Last Change
        - Most Significant Job
- Coordinates
    - X-Y Coords (Source)
    - X-Y Coords (Load)
- Manufacturing Data
- Physical Description
    - Length (ft)
    - Express/Tie Code
    - Branch or End Code
    - No. of Branches
    - Pri. Wire Code
    - Neut. Wire Code
    - Pri. CMD
    - Duplicate End Pt.
    - No. Branches
- Reliability Date
- Cost Data
    - No. Customers
- Record Status
- Time History Data

Table 42. Distribution Transformer

- Device Identifier
  - Distribution Transformer Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Source Identifier
  - Load Identifier
  - Circuit No.
  - Substation Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating (Line kV)
  - Voltage Rating (Load kV)
  - Voltage Rating (Name Place)
  - Z%
  - R
  - X
  - Watts Loss Full Load $3\phi + 1\phi$
  - Watts Loss No Load $3\phi + 1\phi$
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Data
  - Job Number
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
  - Wire Installation Data
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer $ Year
  - Model/Type
  - Serial No.
  - Drawing Ref.
- Physical Description
  - Connection Code
  - Mounting Type
  - Width
  - Depth
  - Height O.A.
  - Weight
  - Insulation
  - Amount of Oil
- Reliability Data
- Cost Data
  - Cost
- Record Status
- Time History

Table 43. Boost and Buck Transformer

- Device Identifier
  - Boost and Buck Transformer
    - Serial No.
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Source Identifier
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
- Voltage Rating (Line kV)
- Voltage Rating (Line kV)
- KVA Rating
- %Boost/%Buck
- Z
- R
- X
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job Number
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
  - Wire Installation Data
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year
  - Model or Type
  - Serial No.
  - Drawing Ref.
- Physical Description
  - Connection Code
  - Mounting Type
  - Width
  - Depth
  - Height
  - Weight
  - Insulation
- Reliability Data
- Cost Data
  - Cost
- Record Status
- Time History

Table 44. Shunt Capacitor

- Device Identifier
  - Shunt Capacitor Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating
  - No. of Unit 1's (A)
  - KVAR Rating of Unit 1 (A)
  - No. of Unit 2's (B)
  - KVAR Rating of Unit 2 (B)
  - No. of Unit 3's (C)
  - KVAR Rating of Unit 3 (C)
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job Number
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
  - Distance from Load Point
- Manufacturing Data
  - Equipment Manufacturer & Year (Unit 1)
  - Equipment Manufacturer & Year (Unit 2)
  - Equipment Manufacturer & Year (Unit 3)
  - Control Manufacturer & Year
- Physical Description
  - Cap Control Code
  - Cap Override Code
  - Controls On
  - Controls Off
  - Blocked 1
  - Blocked 2
  - Controls - Max
  - Controls - Min
  - Control Style Code
  - Insulation
- Reliability Data
- Cost Data
- Record Status
- Time History

Table 45. Series Capacitor

- Device Identifier
  - Series Capacitor Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating of Equipment
  - No. of Units in Series
  - No. of Units in Parallel
  - Voltage Rating - Parallel
  - KVAR Rating - Series
  - KVAR Rating - Parallel
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job Number
    - Orig. Installation
    - Last Change
    - Most Significant Change
  - Purchase Order
  - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Distance from Load Pt.
  - Location Address
- Manufacturing Data
  - Manufacturer & Year
  - Model or Type
  - Serial No.
  - Drawing Reference
- Physical Description
  - Insulation
  - Mounting Type
  - Duty Code
- Reliability Data
- Cost Data
  - Cost
- Record Status
- Time History

Table 46. Line Regulator

- Device Identifier
  - Line Regulator Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Station Identifier
  - Substation Identifier
  - Circuit No.
- Electrical Description
  - Phase Code
  - Voltage Rating
  - KVA Rating - From Nameplate
  - CT Ratio
  - PT Ratio
  - % Regulation - From Nameplate
  - R
  - X
  - % Regulation - From Nameplate
  - R
  - X
  - % Regulation - From Nameplate
  - R
  - X
  - % Regulation - As Set
  - Regulator Setting - KVA - As Set
  - Compensation Settings per φ
  - R (A,B,C)
  - X (A,B,C)
  - Band Width (A,B,C)
  - Time Delay (A,B,C)
  - No Load Voltage (A,B,C)
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job Number
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year (A,B,C)
  - Model or Type (A,B,C)
  - Serial No. (A,B,C)
  - Drawing Ref.
- Physical Description
  - Mounting Type
  - Width (A,B,C)
  - Depth (A,B,C)
  - Height O.A. (A,B,C)
  - Weight (A,B,C)

Table 46. (cont'd)

  - Insulation (A,B,C)
  - Control Type
- Reliability Data
- Cost Data
  - Cost (A)
  - Cost (B)
  - Cost (C)
  - Total Station Cost
- Record Status
- Time History

Table 47. Sectionalizing Device

- Device Identifier
  - Sectionalizing Device Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Station Identifier
  - Substation Identifier
  - Circuit No.
- Electrical Description
  - Phase Code
  - Voltage Rating
  - Ampere Rating (if fuse)
  - Fault Current Rating
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
  - Coordinates
    - X-Y Coordinates
    - Location Address
  - Manufacturing Data
    - Manufacturer & Year
    - Model/Type
    - Serial No.
  - Physical Description
    - Connection Code
    - Fuse/Blade Type Code
    - TCC Curve Reference Code
    - Mounting Type
  - Reliability Data
  - Cost Data
  - Record Status
  - Time History

Table 48. Line Recloser

- Device Identifier
  - Line Recloser Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Data
  - Phase Code
  - Voltage Rating (from Name Plate)
  - Ampere Rating (from Name Plate)
  - F.C. Interrupting Rating
  - Series Fuse Rating
  - Series Fuse Type Code
  - Bypass Rating
  - Bypass Fuse Type Code
  - No. of Cycles
  - Recloser Code
  - Quick Trip Current
  - No. of Quick Trips
  - Retarded Trip Current
  - Quick, Retarded, Ext Retarded Curve Ref. Code
  - No. of Ret, Ext Retarded Trips
- Automation Code
- Operating Status
  - Data & Reading of Counters
- Installation Data
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year (A,B,C)
  - Model/Type (A,B,C)
  - Serial No. (A,B,C)
- Physical Description
  - No. of Units in Station
  - Mounting Type
  - Width (A,B,C)
  - Depth (A,B,C)
  - Height (A,B,C)
  - Weight (A,B,C)
  - Gal. of Insulation (A,B,C)
- Reliability Data
- Cost Data
  - Cost (A,B,C)
  - Total Station Cost
- Record Status
- Time History

Table 49. Static Recloser

- Device Identifier
  - Static Recloser Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating (Name Plate)
  - Ampere Rating (Name Plate)
  - F.C. Interrupting Rating (Name Plate)
  - No. of Cycles
  - Bypass Fuse Rating
  - Bypass Fuse Type Code
  - Ampere Rating as Set
  - Current Transf. Ratio
  - Amperage Tap
  - Amperage Plug
  - Ground Relay Plug
  - Ground Relay Curve Ref.
  - Ground Relay Amps as Set
  - Static Recloser Code
  - Reclosing Times
  - Resetting Time
  - Operations to Lockout
  - Quick Trip Current
  - Retarded Trip Current
  - Instantaneous Curve Ref.
  - Retarded Curve Ref.
  - Time Delay Type
  - No. of Quick Trips
  - No. of Retarded Trips
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
    - Most Significant
  - Purchase Order
  - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year
  - Model/Type
  - Serial No.
- Physical Description
  - Voltage Source Code
  - Width
  - Depth

Table 49. (cont'd)

  - Height
  - Weight
  - Insulation
  - Cal. of Insulation
- Reliability Data
- Cost Data
  - Cost
  - Total Station Cost
- Record Status
- Time History

Table 50. Reactor

- Device Identifier
  - Reactor Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating
  - Ampere Rating
  - R
  - X
  - No. Units in Station
  - KVAR Rating Per Phase (A,B,C)
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Manhours
  - Job No.
  - Original Installation
  - Last Change
  - Most Significant
- Purchase Order
- Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year (A), (B), (C)
  - Model/Type (A), (B), (C)
  - Serial No. (A), (B), (C)
  - Drawing Reference
- Physical Description
  - Mounting Type
  - Width (A), (B), (C)
  - Depth (A), (B), (C)
  - Height (A), (B), (C)
  - Weight (A), (B), (C)
  - Insulation

Table 50. (cont'd)

- Reliability Data
- Cost Data
  - Cost (A), (B), (C)
  - Total Station Cost
- Record Status
- Time History

Table 51. Tie Sectionalizing Device

- Device Identifier
  - Sectionalizing Device Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
  - TIE Substation Identifier
  - TIE Circuit No.
  - Tie Load Identifier
  - Connection Code
- Electrical Description
  - Phase Code
  - Voltage Rating (Name Plate)
  - Ampere Rating (Fuse)
  - Fault Current Rating
  - Switch Group Code
  - Fuse/Blade Type Code
  - TCC Curve Reference Code
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Date of Installation
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
    - Most Significant
    - Purchase Order
    - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year
  - Model/Type
- Physical Description
  - Mounting Type
- Reliability Data
- Cost Data
- Record Status
- Time History

Table 52. Customer Meter

- Device Identifier
  - Customater Meter Identifier
- Organizational Description
  - Division Code
  - District Code
  - Customer No.
- Network Data
  - Source Identifier
  - Circuit No.
  - Substation Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating
  - Ampere Rating
  - Demand/KWh
- Automation Code
- Operating Status
  - Calibration Date
- Service Code
- Installation Data
  - Installation Date
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year
  - Model/Type
  - Serial No.
- Physical Description
  - Mounting Type
  - Width
  - Depth
  - Height
  - Weight
- Reliability Data
- Cost Data
  - Cost
- Record Status
- Time History
  - Customer Records for Last Two Years by Month

Table 53. Substation Breaker/Relayed Recloser

- Device Identifier
  - Recloser Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Load Identifier
  - Circuit No.
  - Substation Identifier
  - Station Identifier
- Electrical Description
  - Phase Code
  - Voltage Rating (Name Plate)
  - Ampere Rating (Name Plate)
  - F.C. Interrupting Rate (Name Plate)
  - No. of Cycles
  - Bypass Fuse Rating
  - Bypass Fuse Type Code
  - TCC Curve
  - Ampere Rating (As Set)
  - Ampere Rating (Maximum Possible)
  - Current Transf: Ratio
  - Low Instantaneous Tap
  - High Instantaneous Tap
  - Time Delay Type
  - Time Delay Tap
  - Time Lever Curve
  - Reclosing Relay Type
  - Ground Relay Type
  - Ground Relay Tap
  - Ground Relay Time Lever Curve
  - Ground Relay Ampere (As Set)
  - Reclosing Times (3)
  - Resetting Time
  - Operations to Lockout
  - Low Instantaneous Current
  - High Instantaneous Current
  - Time Delay Current
  - No. of Instantaneous Trips
  - No. of Time Delay Trips
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Installation Date
  - Installation Manhours
  - Job No.
    - Original Installation
    - Last Change
    - Most Significant
  - Purchase Order
  - Maintenance Record
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Manufacturer & Year

Table 53. (cont'd)

- Model/Type
- Serial No.
- Physical Description
  - Voltage Source Code
  - Width
  - Depth
  - Height
  - Weight
  - Insulation
  - Gal. of Insulation
- Reliability Data
- Cost Data
  - Cost
  - Total Station Cost
- Record Status
- Time History
- Coordinates
  - X-Y Coordinates
  - Location Address
- Manufacturing Data
  - Contractor Identification List
- Physical Description
  - Service Area Size in Square Miles
  - Yard Dimensions
  - Getaway Type Code
  - No. of Getaways
  - Drawing Ref.
- Reliability Data
- Cost
  - Original Construction Cost
  - Replacement Cost
  - Depreciation Rate
  - Annual Costs
    - Interest
    - Depreciation
    - Insurance
    - Taxes
    - Operation & Maintenance
  - Total Cost of Subnetwork
- Record Status

Table 54. Substation

- Device Identifier
  - Substation Identifier
- Organizational Description
  - Division Code
  - District Code
- Network Data
  - Breaker/Relayed Recloser Identifier List
  - Capacitor Identifier List
  - Reactor Identifier List
- Electrical Description
  - Capacity (In KVA or MVA)
  - Primary & Secondary Voltage Levels
  - Load in KW, KVAR
- Automation Code
- Operating Status
- Service Code
- Installation Data
  - Year of Installation
  - Installation Manhours
  - Maintenance Record
- Time History
  - Cost of Purchased Power (By Month)
  - Cost of Generated Power (By Month)
    - Demand Cost in $/kW

PLANNING DATA

There is some data associated with the distribution system which are not directly related to the network. These data are generally classified as planning data. For the purposes of this report, they have been separated into two tables, one listing general planning data (Table 55) and the second shows parameter construction modules and their parameters (Table 56). Table 57 shows the applications appropriate for each general data classification.

Table 55. Planning Data

- Land Use Plans
- Planning Code
- Saturation Code
- Class Code
- Company Policies
- Transformeter Load Management Data
- Right of Way Acquisition Costs
- Demographic Influences
- Innovation Trends
- Economic Conditions

Table 56. Data Requirements of the Subprograms

SUBPROGRAM		INPUT DATA
#	Name	
1	SFC_i	a) cost of land b) substation construction cost
2	FFC_{ij}	a) grid values of endpoints b) cost of material and labor per feeder mile
3	TFC_{ij}	a) grid values of SS_i and SS_j b) cost of material and labor per tie-feeder mile
4	DFC_{ij}	a) grid values of demand centers i and j b) cost of material and labor per feeder mile
5	FRC_{ij}	a) grid values of demand center j and SS_i b) cost of material and labor per feeder mile
6	FBC_i	Value of FBC_i
7	C_{ik}	a) Transformer sizes b) Transformer costs c) Transformer inventory
8	Power Loss Curves	a) energy cost b) demand cost due to fast capacity as a result of energy losses in feeders c) grid values of endpoints.

Table 57. Application and Summary of Data Requirements

	SUBSTATION Service Area Size	RATINGS: Transformer kVA	Impedance (%)	Voltage	Winding Connections	Location	CONSTRUCTION COSTS: Original	Replacement	Date of Installation	Structure Type	Loading Data (kW, kWh, p.f.)	PRIMARY DISTRIBUTION LINE SECTION DATA: Location Information	Conductor Information Wire Size & Type	Number of Phases	Structure Information	Reliability Data: Customer Data	Loading Data	Hours in Service	Equipment Data Switchgear and Protective Equipment	Voltage Regulation Equip.	Metering, Telemetry	Reliability Infor.
Load Projection											●						●					
Substation Expansion	●	●		●	●	●					●											
Substation Loading						●					●						●					
Substation Siting	●					●																
Number of Feeders											●	●	●	●		●	●					
Feeder Routing												●				●						
Fault Current Study		●	●	●	●							●	●	●	●				●			
Voltage Profile		●		●	●						●	●	●	●		●	●			●		
Capacitor Location		●		●							●	●	●	●			●			●		
Voltage Regulator Location											●	●	●	●			●			●		
Transformer Load Management											●	●	●	●		●	●					
Reliability																●	●	●				●
Load Management											●					●	●					
Economic Conductor Sizing				●							●		●	●	●		●					
Equipment Inventory		●																	●	●	●	
Plant Investment		●					●	●	●	●		●	●	●	●				●	●	●	
Tax Computation												●				●						

Table 57. Cont'D

	SECONDARY DISTRIBUTION									MISCELLANEOUS INFORMATION								
	Line Section Data Location Information	Conductor Information	Structure Information	Equipment Data Distribution Transformer kVA	Voltage	Impedance (%)	Number of Customers Served	Loading Data	Switchgear & Fusing Information	Operating Cost Data Cost of Generation/Purch. Power	Annual Operating Costs	Right of Way Acquisition Costs	Weather Data	Load Projection Data, Growth Rates & Types	Land Use	Geography	Aerial Photography	Population
Load Projection								●					●	●	●	●	●	●
Substation Expansion																		
Substation Loading													●	●				
Substation Siting															●	●	●	●
Number of Feeders																		
Feeder Routing															●	●	●	●
Fault Current Study	●	●		●	●	●			●									
Voltage Profile	●	●		●	●		●	●										
Capacitor Location				●				●										
Voltage Regulator Location																		
Transformer Load Management				●	●		●	●										
Reliability									●									
Load Management							●	●										
Economic Conductor Sizing										●	●							
Equipment Inventory				●					●									
Plant Investment	●	●	●	●														
Tax Computation	●						●											

TASK 3.12: THE IMPLEMENTATION OF THE DATABASE ON A MODERN DIGITAL COMPUTER IN A MANNER WHICH MAXIMIZES THE TRANSPORTABILITY OF DATA AND RELATED SOFTWARE AND IS COMPATIBLE WITH OTHER SOFTWARE USED IN THE PLANNING METHODOLOGY.

> Da habt ihr's nun! mit Narren sich beladen
> Das kommt zuletzt dem Teufel selbst zu Schaden.
>
> *Faust*, Part II, Act 1.

INTRODUCTION

Portability is a requirement having two major aspects. The first is the portability of the data itself; this is the easiest condition to satisfy. Data in character form may be ported easily from one computing environment to another. The second is the portability of the software which processes the data and this is complicated by the nature of the application to which the software is to be devoted. A database management system generally requires the use of the full repertoire of machine instructions for a given computer. Current software technology suggests that only high-level languages be used to develop a database management system but the most widespread high-level languages were not seen as satisfactory for such an effort. As a solution to the implementation language problem, the language C was chosen.[1] Once the software was fully implemented, it was then translated (by hand) into Ratfor[2] which in turn, is translated into Fortran IV by the Ratfor translator. Finally, the resulting Fortran source is optimized, using an approach suggested by Knuth.[3]

This effort has resulted in the implementation of three software products: a canonifier program which allows the database to be software-compatible with all of the application programs; a database management system supporting the suggested database; an optimizer program which transforms the Fortran output of the Ratfor translator into efficient code; and lastly, a breadboard version of the Shell program (described under Tasks 3.6 and 3.10).

THE CANONIFIER PROGRAM

A systems view of the Canonifier is shown in Figure 37. The program runs under interactive control via the terminal but the usual source of control data is from the input specification file. This file contains statements which may be thought of as very general format statements. The program functions in two modes. In one mode, it reformats data which has been extracted from the database using queries. After being reformatted, the data may serve as input to an analysis program. Thus this mode is called input mode. The other mode is the converse of input mode. It accepts output from an analysis program and reformats it so that it may be incorporated into the database.

The error file (Figure 37) is used to note any specification errors encountered either in the terminal input or in the data found in the input specification file. If any errors are detected, no output file is generated. Instead, the user uses the text editor to correct the input specification file or re-enters data from the terminal (or both) until no further errors are found.

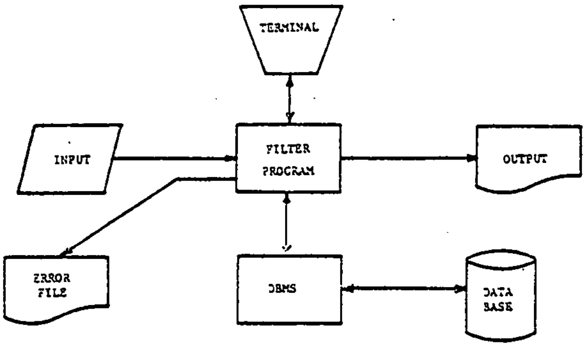


Figure 37.

Detailed Description

Input specification records can be classified into five different types. These are listed and explained in Table 58. The general form of the format specification is shown in Table 59. Conversion characters and their meanings are shown in Table 60.

Table 58. Input Specification Record Types

Statement Type	First Character In Statement	Statement Body	Meaning
Message	m	<message>	Character String message terminated by newline character displayed on terminal
Data Record	d	<string>	string terminated by newline character written unaltered to output file
Query Request	q	<query list>	One or more queries passed to DBMS
Query Format	b	<format specification>	character string terminated by newline character considered as <format specification> for database buffer file
Terminal Format	a	<format specification>	character string terminated by newline character considered as <format specification> for terminal input data

---

[1] Kernighan, B.W., and Ritchie, D.M., The C Programming Language, Prentice-Hall, Englewood Cliffs, N.J., 1978.

[2] Kernighan, B.W., and Plauger, Software Tools, Prentice-Hall, Englewood Cliffs, N.Y. 1977.

[3] Knuth, D.E., Structured Programming with goto Statements," Computer Surveys, Vol. 6 (1974) pp 261-301.

Table 59. General Form of <format specification>

$$\left\{\begin{matrix} c \\ \langle\text{rfac 1}\rangle \end{matrix}\right\} . \left\{\begin{matrix} d \\ \langle\text{rfac 2}\rangle \end{matrix}\right\} , \left\{\begin{matrix} \langle\text{char}\rangle \\ \langle\text{f\_spec}\rangle \end{matrix}\right\} \dots$$

<rfac 1>	- integer specifying the number of lines to be expected by the Canonifier Program; if "C" is used, value read from terminal via <f_spec> of type C.
<rfac 2>	- integer specifying number of times the format specification following the comma is to be repeated; if "d" is used, value read from terminal via <f_spec> of type d.
<char>	- Any ASCII character except "%" or newline
<f_spec>	- if "%%," copied to output buffer as "%"; if %[-][<f_width>][.<precision]<con_char> then a data item from the terminal is expected;
[-]	- optional minus sign specifies left adjustment of data in output field
<f_width>	- field width specifies total number of characters data item is to occupy on output file.
<precision>	- if the number is to be interpreted as a floating point number, <precision> specifies number of digits to right of decimal point. Otherwise, it specifies the number of non blank characters comprising the data object. See succeeding table.
<con_char>	- see succeeding table.

Table 60. Conversion Characters

<can_char> may be any of the following:

Character	Type	Meaning
i	integer	Maximum number of digits which may be accepted; if input exceeds this number, an error will result.
c	<rfact 1>	Same as above.
d	<rfact 2>	Same as above.
f	floating point	Number of digits to right of decimal point; if input exceeds this number, least significant digits truncated.
e	floating point with scaling factor	Same as above
s	string	Maximum string length; if input too long, truncate on right; if too short, pad with blanks on right.

As an example, suppose it is required to output an n by m matrix of real numbers. Elements on the same row should be separated by commas. Each integer should be no longer than four digits, and should occupy five spaces on the output line. All integers are to be read from the user terminal. A possible input specification record is:

```
m enter 2 numbers to be used as # of rows & columns.
a 1.1,n=%c, m=%d
m enter the entire matrix 1 row for 1 input line
m each element is 4 or less digits long, separated by
  commas.
a c.d, %5.4i,
```

THE DATABASE MANAGEMENT SYSTEM

The database management system, implemented in Ratfor, uses the PATRICIA data structure as the basis of its directory.[4] The simplest description of PATRICIA can be given as follows.[5]

Let the generalized K-formula αab indicate that nodes a,b are joined together by an α-link in that order, i.e. from a to b. An S-System is a symbol system or defining scheme which resembles a grammar. In distinction to the usual phrase structured grammar, however, the productions in an S-System frequently employ unique terminal symbols in each application of a given production. Symbols so treated are written for example, as $a_{[j]}$ where the brackets indicate an entire set from which a symbol is to be drawn (the drawing process has a memory; once a symbol has been removed from the set it may not appear in the production when applied later in the derivation). Accompanying this convention is one which says that if such symbols are enclosed in French quotes, ("<<", ">>") and have the same index then they are to be replaced by same symbol.

These conventions permit the following description of PATRICIA:

$$P \rightarrow \begin{bmatrix} \bar{\alpha}\ ar & \\ \alpha\ ar & \langle\langle LR\rangle\rangle \\ \alpha\ ar & \langle\langle RL\rangle\rangle \end{bmatrix} a_r$$

$$L \rightarrow \begin{bmatrix} \bar{\alpha}\ a_{[j]} & \\ \alpha\ a_{[j]} & \langle\langle LR\rangle\rangle \\ \alpha\ a_{[j]} & \langle\langle RL\rangle\rangle \end{bmatrix} \qquad R \rightarrow \begin{bmatrix} \bar{\beta}\ a_{[j]} & \\ \beta\ a_{[j]} & \langle\langle LR\rangle\rangle \\ \beta\ a_{[j]} & \langle\langle RL\rangle\rangle \end{bmatrix}$$

with α, β standing for the left and right links respectively. Hueristically, these productions have the graph grammar forms shown in Figure 2.

The structure is recognizable by a Push Down Automaton. Its description is as follows.

$$\begin{aligned} \underline{P}(\$):\quad & d(q_0, \bar{\alpha}, \$) \rightarrow (q_1, \$) \\ & d(q_1, a_j, \$) \rightarrow (q_2, a_j \&) \\ & d(q_2, a_j, a_j) \rightarrow (q_3, e) \\ & d(q_3, e, \$) \rightarrow (q_f, \$) \\ & d(q_o, \alpha, \$) \rightarrow (q_4, \$) \\ & d(q_4, a_j, \$) \rightarrow q_2{:}\ \underline{R}\ (a_j \$)\ \underline{\tilde{L}}\ (z) \\ & d(q_4, a_j, \$) \rightarrow q_2{:}\ \underline{L}\ (a_j \$)\ \underline{\tilde{R}}\ (z) \end{aligned}$$

■

$$\begin{aligned} \underline{L}(z):\quad & d(q_0, \bar{\ }, z) \rightarrow (q_1, z) \\ & d(q_1, a_j, z) \rightarrow (q_f, a_j z) \\ & d(q_0, \ , z) \rightarrow (q_2, z) \\ & d(q_2, a_j, z) \rightarrow q_f{:}\ \underline{R}(a_j z)\ \underline{\tilde{L}}(z') \\ & d(q_2, a_j, z) \rightarrow q_f{:}\ \underline{L}(a_j z)\ \underline{\tilde{R}}(z') \end{aligned}$$

■

---

[4] Knuth, D.E., The Art of Computer Programming, vol. 3, Addison-Wesley, Reading, Mass., 1973 pp. 490f.

[5] Thompson, J.C., and Atkins, G.E., A Syntactical Specification of the Data Structure PATRICIA, to appear in a technical journal.

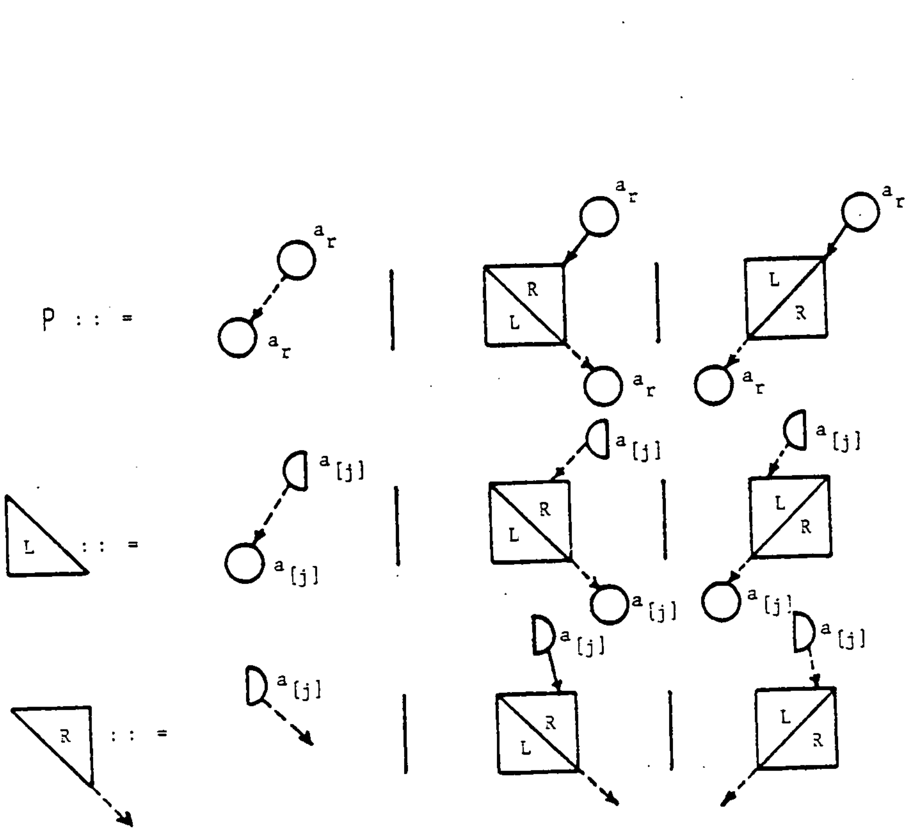


Figure 38

$\underline{L}$ (z) :	$d(q_0, \bar{\alpha}, z)$	(q_1, z)
	$d(q_1, a_j, a_j)$	(q_f, e)
	$d(q_0, \alpha, z)$	(q_2, z)
	$d(q_2, a_j, a_j)$	q_f: $\underline{R}(e)$ $\underline{\tilde{L}}(z')$
	$d(q_2, a_j, a_j)$	q_f: $\underline{\tilde{L}}(e)$ $\underline{R}(z')$

The notation q_f: $\underline{R}$ (...), appearing on the right side of a transition rule means that the present automaton returns to state q_f after invoking machine $\underline{R}$ and $\tilde{R}$ has terminated. R and $\tilde{R}$ are obtained from $\underline{L}$, $\underline{\tilde{L}}$ by replacing every occurrence of α by β and vice versa.

The fact that the structure is recognized by a PDA is very important because this means that basic integrity checking of the structure can be automated.

Node Relations

The description of PATRICIA is not complete without a careful statement of the defining relations among the data stored in the nodes. To make the definitions as general as possible, consider the following. Let the set $S = \{K_i\}$ of elements K_i be keys. In the sequel, a mapping will be defined between nodes in the structure and elements of S; this mapping will be written $\hat{K}a_j = K_i$. For the present, this mapping will be ignored. Let $\{\overset{\ell}{\equiv}\}$ be an indexed set of equivalence relations over S and let $\{\overset{m}{<}, \overset{m}{>}\}$ be an indexed set of precedence relations. Then consider the axioms:

(i) $K_p \overset{0}{\equiv} K_q \quad K_p, K_q \in S$

(ii) For $K_p, K_q \in S, \exists m$ such that $K_p \overset{\ell}{\equiv} K_f \quad 0 \leq \ell < m$ implies $K_p \overset{\ell'}{\not\equiv} K_q \quad m \leq \ell'$

(iii) Let m be the least such value appearing in (ii) for given K_p, K_q (and greater than zero); then either $K_p \overset{m}{<} K_q$ or $K_p \overset{m}{>} K_q$ exclusively

(iv) If $K_p \overset{m}{\equiv} K_q$ & $K_p \overset{m+1}{\not\equiv} K_q$ and $K_p \overset{m}{\equiv} K_r$ & $K_p \overset{m+1}{\not\equiv} K_r$ then $K_q \overset{m+1}{\equiv} K_r$ and if $K_p \overset{m+1}{<} K_q$ then $K_p \overset{m+1}{<} K_r$

(v) Iff for $K_p \overset{\ell_1}{\equiv} K_q$ & $K_q \overset{\ell_2}{\equiv} K_r$ is $K_p \overset{\ell_1}{\equiv} K_r \quad \ell_1 \leq \ell_2$

Select an arbitrary element $K_r \in S$. For any other element $K_q \in S$, we have

$$K_r \overset{\ell}{\equiv} K_q \quad 0 \leq \ell < m \tag{1}$$

In general, m depends on K_q. Fix m; label the subset of S for which (1) holds $E^m(k_r)$. Then S can be written

$$S = E^{m_1}(K_r) \cup E^{m_2}(K_r) \cup \ldots$$

It is clear that the $\bar{E}$'s are disjoint

$$E^{m_i}(K_r) \cap E^{m_j}(K_r) = \emptyset \quad m_1 \neq m_j$$

Each $E^{m_i}(K_r)$ which contains more than one element, in turn may be partitioned into sub-equivalence classes by first choosing a root $K_r' \in E^{m_i}(K_r)$. We may then write

$$E^{m_i}(K_r) = E^{m_i,m_1'}(K_r') \cup E^{m_i,m_2'}(K_r') \cup \ldots$$

Continuing this process, each key $K_p \in S$ may be considered the root of some equivalence class $E^{m_i,m_j\ldots}(K_p)$. It may be seen that this partitioning process does not generate a unique partition since the choice of roots is arbitrary within a given subclass.

Algorithms

The PATRICIA structure not only has the advantage of a precise mathematical description (and thus is subject to careful analysis) as well as being recognizable, but in addition, may be manipulated and implemented by simple algorithms. Algorithms for key search and key insertion are shown in Figures 39 and 40 respectively. This simplicity makes the task of implementing a practical system much easier than it might otherwise be.

As an example, a PATRICIA tree will be constructed for the text

HUMPTY DUMPTY SAT ON A WALL,

treating each of the phrases

HUMPTY DUMPTY SAT ON A WALL.

DUMPTY SAT ON A WALL.

SAT ON A WALL.

ON A WALL.

A WALL.

WALL.

as a key in the directory. To keep the example simple, assume that the data to be stored with the key is the part of speech of the leading word in the phrase. The relation thus has the form shown in Figure 41. The node structure of the PATRICIA tree has the form shown in Figure 42 where the fields have the following significance:[6]

KEY, a pointer to the text of the phrase. (Represented by *K* in the K-formulas.)

LLINK and RLINK, pointers within the tree. (Represented by α, β, respectively in the K-formulas.)

LTAG and RTAG, one-bit fields which tell whether or not LLINK and RLINK, respectively, are pointers to sons or ancesters of the node. (LTAG = 1, RTAG = 1, represented by α, β respectively, in the K-formulas.)

```
procedure P_SEARCH (HEAD, KEY, LLINK, RLINK, LTAG,
                    RTAG, SKIP, K);
pointer array LLINK, RLINK;
integer array KEY;
string        K;
  begin
     procedure P_PROBE (HEAD, KEY, LLINK, RLINK, LTAG,
                        RTAG, SKIP, K, n);
     pointer array LLINK, RLINK;
     integer array LTAG, RTAG, SKIP;
     string array  KEY;
     string        K;
     integer       n;
        begin  comment: p, q, ssum are global variables;
           integer dot;
           p := HEAD;  ssum := 0;  dot := 1;
           q := p;  p := LLINK [ q ]
           if LTAG [q] ≠ dot then
              begin
                 ssum := ssum + SKIP [p];
                 while  ssum ≤ n do
                    begin
                       if BIT (ssum, K) ≠ 0 then
                          begin
                             q := p; p := RLINK [q];
                             if RTAG [q] = dot then
                                                return
                          end
                       else
                          begin  comment: BIT = 0;
                             q := p;  p := LLINK [ q];
                             if LTAG [ q] = dot then
                                                return
                          end;
                       ssum := ssum + SKIP [p]
                    end
              end
        end P_PROBE;
  pointer p, q;
  integer n, ssum;
     n := LENGTH (K);
     P_PROBE (HEAD, KEY, LLINK, RLINK, LTAG, RTAG,
              SKIP, K, n);
     if SEQUELS (n, KEY [p], K) then return (p)
                                else return (Λ)
end P_SEARCH
```

Figure 39

SKIP, a number which tells how many bits to skip when searching, as explained in the algorithm. (SKIP corresponds to the value of m in the equivalence relations exhibited above.)

The PATRICIA tree for this example is shown in Figure 43. The reader will note that in the case where two or more attribute values are required to determine a key into the relation, the PATRICIA structure is particularly useful, since the character strings forming the desired key may be concatenated and the search performed on the resulting string. Thus the same algorithms serve to locate the desired tuple, regardless of the key structure.

Relational Algebra Language

As discussed in the section on Task 3.10, the query language is based on the relational algebra of Codd. The form adopted for this implementation is shown in Table 61. A prefix format was chosen to that command composition that would be straightforward. The query exhibited in the section on Task 3.10 can be written as

---

[6] Knuth, D.E., op. cit., p. 491.

```
procedure P_INSERT (HEAD, KEY, LLINK, RLINK, LTAG, RTAG,
                    SKIP, K);
pointer array LLINK, RLINK;
integer array LTAG, RTAG, SKIP;
pointer HEAD;
string array KEY;
string K;
  begin
      procedure P_PROBE (HEAD, KEY, LLINK, RLINK, LTAG,
                         RTAG, SKIP, K, n);
      pointer array LLINK, RLINK;
      integer array LTAG, RTAG, SKIP;
      string array  KEY;
      string        K;
      integer       n;
        begin  comment: p, q, ssum are global variables;
          integer dot;
          p := HEAD;  ssum := 0;  dot := 1;
          q := p;  p := LLINK [ q ]
          if LTAG [q] ≠ dot then
            begin
              ssum := ssum + SKIP [p];
              while  ssum ≤ n do
                begin
                  if BIT (ssum, K) ≠ 0 then
                    begin
                      q := p; p := RLINK [q];
                      if RTAG [q] = dot then
                                            return
                    end
                  else
                    begin  comment: BIT = 0;
                      q := p;  p := LLINK [ q ];
                      if LTAG [ q ] = doc then
                                            return
                    end;
                  ssum := ssum + SKIP [p]
                end
            end
        end P_PROBE;

    pointer p, q, r;
    integer ℓ, n, t, ssum, solid:
    string Kp;
    bit b;
      if HEAD = Λ then
        begin
          HEAD := AVAIL;
          KEY [HEAD] := K;
          LLINK [HEAD] := HEAD;
          LTAG [HEAD]  := 1;
          RLINK [HEAD]  := Λ
        end
      else
        begin
          solid := 0;
          n := LENGTH (K);
          P_PROBE (HEAD, KEY, LLINK, RLINK, LTAG, RTAG,
                   SKIP, K, n);
          comment:  search must be unsucessful...
                    no key is prefix of another;
          Kp := KEY [p];
          for ℓ := 1 to n do
            if BIT (ℓ, K) ≠ BIT (ℓ, Kp) then exit;
          b := BIT (ℓ,K);
          P_PROBE (HEAD, KEY, LLINK, RLINK, LTAG, RTAG,
                   SKIP, K,ℓ - 1);
          r := AVAIL;
          KEY [r]:= K;
          if LLINK [q ] = p then
            begin
              LLINK [q ]:= r;  t := LTAG [q];
              LTAG [q ]:= solid
            end
          else
            begin
              RLINK [q] := r;  t := RTAG [q];
              RTAG [q] := solid
            end
          if b = 0 then
            begin
              LTAG [r] := 1;  LLINK [r] := r;
              RTAG [r] := t;  RLINK [r] := p
            end
          else
            begin
              RTAG [r] := 1;  RLINK [r] := r;
              LTAG [r] := t;  LLINK [r] := p
            end;
          if t = 1 then SKIP [r] := ℓ - ssum
          else
            begin
              SKIP [r] := ℓ - ssum + SKIP [p];
              SKIP [p] := ssum - ℓ
            end
end P_INSERT
```

Figure 40

KEY	PART OF SPEECH
Humpty	Subject
Dumpty	Subject
Sat	Verb
On	Preposition
A	Article
Wall	Direct Object

Figure 41

KEY	SKIP
LLINK	RLINK

Figure 42

```
PROJECT(
        SELECT(
               JOIN(TRANSFORMER, SERVICE)
               FOR TRANSFORMER.SERIAL#=SERVICE.SERIAL#)
        WHERE  SERVICE.STATUS='IN' and
        SERVICE.STATUS DATE-CURRENT DATE < 180
OVER  SERIAL#, X-Y COORD.

    TRANSFORMER SERVICE J TRANSFORMER.SERIAL#
      =SERVICE.SERIAL# S STATUS = 'In" and
      STATUS DATE-CURRENT DATE < 180

                           P SERIAL#, X-Y COORD
```

Humpty Dumpty Sat On A Wall

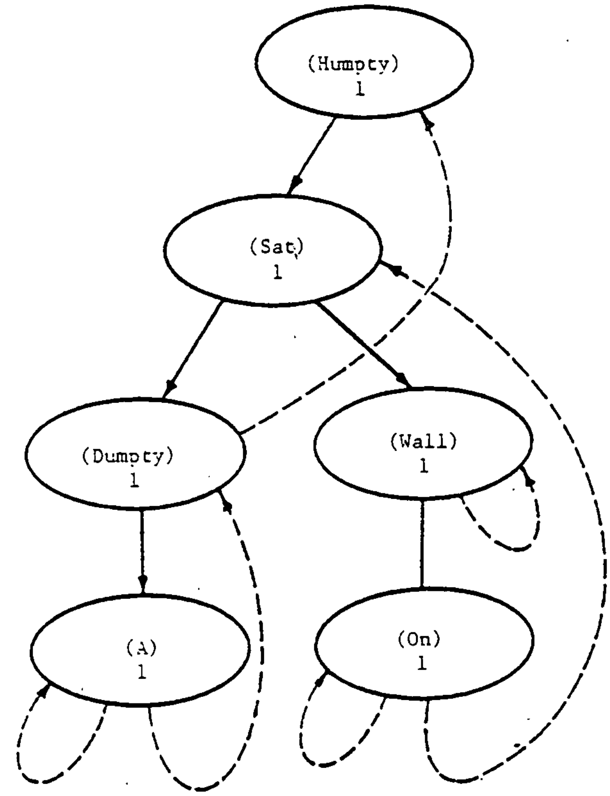


Figure 43

Table 61. Relational Algebra Statements

<select> function
syntax specification.
<input relation name> s <qualification>
<output relation name>
<qualification> ::= <tuple relation><qualification>
| <tuple relation>
<tuple relation>::= <attribute value><boolean operator> <tuple function>

Comments. <output relation name> is either a relation name or the symbol %. In the latter case, % stands for a temporary relation name. If % has been previously defined, it may be used as an <input relation name>.

Example.
transformer s serial# = 1775453 %

<join> function.
syntax specification.
<input relation name$_1$><input relation name$_2$> j
<bi-tuple boolean relation> <output relation name>
<bi-tuple boolean relation>::= <attribute value$_1$>
<boolean operator>
<attribute value$_2$>

<project>function.
syntax specification.
<input relation name> p <attribute name list>
<output relation name>
<attribute name list> ::= <attribute name>
<attribute name list>
| <attribute name>

<divide> function.
syntax specification.
<input divident relation name> <input unary divisor relation name> d <output unary quotient relation name>

<minus> function.
syntax specification.
<input subtrahend relation name><input minuend relation name> m <output result relation name>

<union> function
syntax specification
<input relation name$_1$><input relation name$_2$> u
<output relation name>

<intersect> function
syntax specification
<input relation name$_1$><input relation name$_2$> i
<output relation name>

<multiply> function
syntax specification
<input relation name$_1$><input relation name$_2$> mu
<output relation name>

<invoke database management system> function.
syntax specification
<database name> rodni <password>
comment
The name of the database management system is relation oriented database network information.

<destroy database> function
syntax specification.
<database name> dsdb

<create relation> function.
syntax specification.
<schema specification><access method>
C <relation name>
<schema specification> ::= (<attribute list> :
<key list>)
<attribute list> :: = <attribute name><domain specification><attribute list>
| <attribute name><domain specification>
<key list>:: = <attribute name><key list>
<attribute name>
<domain specification> ::= <attribute type> (<field length>)
<attribute type> ::= <int>| <float> | <char>

```
   <access method> ::=  pat
<destroy relation> function.
    syntax specification.
   <relation name> dsr
   <copy> function
    syntax specification
   <external database name><external relation name>
   <local access method> cp <local relation name>
<modify> function
syntax specification
    Not Implemented.
<output relation> function
syntax specification
   <relation name> wrr
comment.  Output is directed to DBMS buffer
Canonifier formats output for delivery to proper
output medium/device.
<output schema> function
Syntax Specification
<relation name> wrs
<input relation> function
syntax specification
      rdr <relation name>
Comment
    Data is placed in DBMS buffer in a format com-
patible with schema definition by the canonifier.
<delegate> function.
syntax specification
<relation name list> dl <access list>
<relation name list> ::= <relation name>
                         <relation name list>
                        |<relation name>
<access list> ::= <schema specification><permission>
                  <user id list><access list>
                 |<schema specification><permission>
                  <user id list>
<permission> ::=  rd | wr | ex
<user id list>::=  <user id><user id list>
                  |<user id>
<save> function
syntax specification
<retention status> SV <relation name>
<retention status> ::= <julian date>  %
Comment
    The retention status designated by % is temporary,
the relation being destroyed when the current session
with the DBMS terminates.
<purge> function
 syntax specification
           pu
<assign> function
```

```
Syntax Specification
<input relation name><assignment list> asg
  <output relation name>
<assignment list> ::= <attribute name> = <tuple
                       expression>
                       ; <assignment list>
                      |<attribute name> = <tuple
                       expression>
```

FORTRAN SOURCE OPTIMIZER

The portability requirements place a heavy burden on the system designer. The only suitable language which is truly portable at the present time is FORTRAN IV. However, FORTRAN's somewhat primitive control structures make the programming of an application such as a database management system extremely difficult. Such programming tends to be difficult even in a language such as Algol, which has much more powerful control structures. A glance at Figure 40 will illustrate why this is so. To obtain more powerful control structures within a FORTRAN context, the implementers on this project turned to Ratfor, a dialect of FORTRAN which has powerful control structures. Portability for Ratfor is obtained by translating those Ratfor statements which are not in themselves FORTRAN statements into FORTRAN statements. The major drawback to this translation process is that the resulting FORTRAN program is not nearly as efficient as an equivalent program coded by hand would be. Since the database software is extensive (it is comprised of more than forty subprograms), it is not feasible to do the translation by hand.

Thus, it is clear that what is needed is a tool to assist the programmer and that tool is the FORTRAN Source Optimizer.

Major Features

The information flow is illustrated in Figure 44. The optimizer accepts FORTRAN statements generated by the Ratfor processor and performs the following actions.

(1) Detects any syntax errors. The Ratfor processor does little or no syntax checking. Any statement which it does not recognize is assumed to be a FORTRAN statement (not a Ratfor statement) and is passed as it was encountered to the output file, which in turn is passed to the Optimizer.

(2) Finds common subexpressions. Many expressions can be "factored" so that certain arithmetic expressions need to be evaluated only once.

(3) Detects "dead" code. Some portions of a program may not be reachable during normal execution because the programmer has made errors in his algorithm design or because the algorithm was not faithfully translated into a correct Ratfor program.

(4) Eliminates redundant control transfers. The Ratfor control statements which are all single entrance-single exit constructs are implemented using FORTRAN goto statements. The implementation is done in a simple-minded way with the result that no testing is done for situations such as a transfer to statement number 10 which is itself a goto statement jumping to number 50. This kind of redundancy is eliminated.

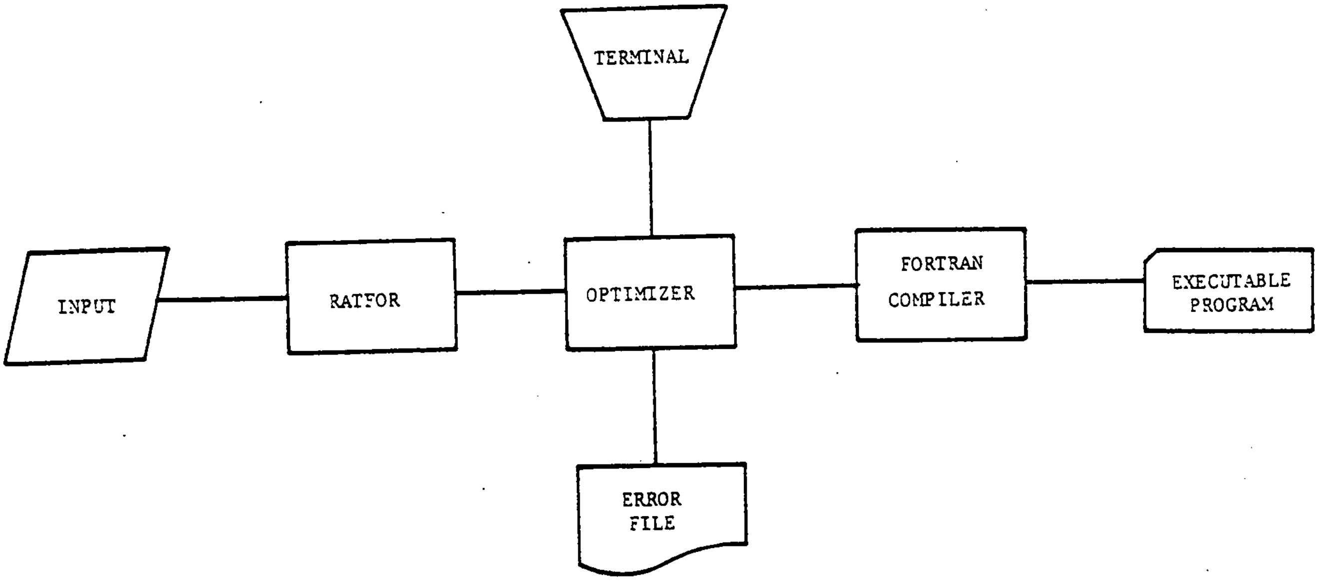


Figure 44

(5) Allows interactive program restructuring. The programmer executes the Optimizer interactively. After initial processing of the FORTRAN source, the optimizer is prepared to accept commands from the programmer and rearrangement of the source can be done. This rearrangement may result in more sophisticated code than the optimizer could produce unassisted. At the same time, the routine processing referred to above has already been done and is thus not a concern of the programmer.

Underlying Principles

Many of the optimizing steps depend on the generation of the flow graph which is a digraph whose edges represent transfers of control within the source program and whose vertices represent groups of source statements in which there are no transfers. The creation of such graphs is well-understood and standard methods are available.[7] These graphs permit the detection of loops, and unreachable (dead) code. Another graph construct which is important which the optimizer constructs is the DAG or Directed Acyclic Graph.[7] This graph is used to analyze the vertices of the flow graph. Such analysis shows how the values computed in one such basic block are used in subsequent blocks. This leads to the detection of common subexpressions and loop optimization.

The syntax analysis is done using simple precedence methods for expression validation and recursive descent parsing of statements. After the programmmer has modified the source program, he may again run both the syntax checker and the analyzer. This allows him to check for errors which might have inadvertantly crept in when the modifications were made.

SHELL IMPLEMENTATION

The Shell program implemented by the work reported here did not meet the portability goals originally set for it. The diversity and peculiarities of several operating systems under which the total system might be expected to function demanded of the staff far greater resources of time and effort than were available. This was recognized early in the life of the project and that recognition prevented any large expenditures of manpower which ultimately would have been wasted. Instead of tilting with such a windmill, the staff decided to implement the shell using the inherent features of Unix on an available PDP-11/70. It is possible for a programmer to write his own command processor (shell) to drive the underlying functions of Unix and this was what was done. It was felt that it was important to experiment with the features proposed in the original functional description and these were the ones implemented. The syntax for the Shell is shown in Table 62. They may be compared with their functional descriptions in Appendix F.

Table 62 Shell Commands

```
<execution> function
syntax specification
    <program name> <from> <to>

  <from> ::=← <input file name
              | <empty>
  <to> ::=  → <output file name>
              | <empty>
<pipeline> function
syntax specification
<program name><pipe descriptor>
<pipe descriptor> ::= :=:  <program name>
                                  <pipe descriptor>
| <empty>
<execute control file>
syntax specification
     sh ← <control file name>
<create> function
syntax specification
  <file pointer>  =  cr <file name>
```

---

[7] Aho, A.V., and Ullman, J.D., Principles of Compiler Design, Addison-Wesley, Reading, Mass., 1977.

```
<open> function
syntax specification
 <file pointer> =  op <file name>
<close> function
syntax specification
   cl  <file pointer>
  <unlink> function
   syntax specification
    rm <file name>
  <list files> function
      ls
```

Table 63 Major System Component Status

System	Implementation Status	Portability Status
Shell	✓	x
Analysis Programs	✓	✓
Network Editor	x	x
Canonifier	✓	✓
Optimizer	✓	x

✓ = implemented/portable   x = unimplemented/nonportable

SUMMARY

Table 63 summarizes the major components of the system which have been implemented. The Shell program itself is the only one which has not been implemented in a portable manner.

The canonifier program has proven to be the key to making all of the applications/analysis programs compatible without requiring any substantial modification of them.

The canonifier program works well as a front end to the database management system (DBMS) which was implemented using the PATRICIA data structure. Not only is PATRICIA a practical directory scheme but its behavior is susceptable to theoretical study which shows how to check the structure's internal consistency using techniques from Formal Language theory.

As part of the DBMS, a user language was implemented based on the relational algebra. This language allows one to write queries in prefix form so that nesting is straightforward.

Lastly, as part of the approach taken to achieve portability, a FORTRAN source optimizer was implemented which allows the system programmer to optimize the portable but inefficient code produced by the Ratfor translator. Using Ratfor allows one to program in a more comfortable language than standard FORTRAN.

TASK B.13: THE EVALUATION OF THE DISTRIBUTION PLANNING METHODOLOGIES DEVELOPED BY THE CONTRACTOR INCLUDING THE EFFECTS OF MAJOR ASSUMPTIONS IN THE MODELS AND TECHNIQUES AND THE AMOUNT OF EFFORT AND COST REQUIRED TO EXPAND THE IMPLEMENTED SOFTWARE PACKAGE INTO A PRODUCTION GRADE PROGRAM

A. THE EVALUATION OF THE DISTRIBUTION PLANNING METHODOLOGIES DEVELOPED BY THE CONTRACTOR INCLUDING THE EFFECTS OF MAJOR ASSUMPTIONS IN THE MODELS AND THE TECHNIQUES

A major accomplishment of this research design has been the use of the integrated data base, data base management system, representational approach, and functional definitions to achieve an optimal planning approach, in theory, which was not possible in previous approaches and to use this system to increase the dimension of the problems that can be solved. The key design elements which accomplish this system are:

The design switches which permit the planner to select possible feeder routes, substation sites, incremental capacities, conductor sizes, etc. The advantage gained here is that the planner can eliminate obviously non-optimal alternatives efficiently and thus significantly reduce the search tree in the MIP solution procedure.

The usage of the data base to reduce the number of variables required in the optimization model. This is accomplished by the design of parameter computation modules which make use of the structure and relevance of the data base. For example, the cost of the substation incremental capacity additions can be computed with knowledge of the current installed cost of transformers, and relevant hardware and the exchange possibilities present in the system. These exchange possibilities are decision variables in other apprаches. This approach assumes the optimal solution achieved will not involve more exchanges than were assumed by the parameter optimization program. It is possible that incremental capacity allocations will exceed the number of transformers in inventory and thus some errors will be made in the cost calculations as the program will not be able to account for all the interaction possibilities for a given inventory state and tradeout set. However, most of the small to medium size planning problems will be unaffected. This problem will be addressed in the implementation by using the overall demand increase as an estimate of the incremental capacity required and estimating the number of substations that will be involved. This will allocate to each substation of finite set of tradeout and inventory possibilities.

The cost assumptions made in the models by others in the literature on power loss values are eliminated in the models developed in this research by use of the actual power loss curves. The cost of power losses has a significant effect, and has much more importance than variations in the incremental cost figures. The completeness of the data base and its organization allow these curves to be computed and expressed as MIP parameters very efficiently. For large problems, the convex approximation used is still much more accurate than the linear approximations used in the literature. This is a particular example of what is meant by the data elements induced by the planning model. The cost of power losses calculation requires the cost of conductor per foot, grid points for the ends of the feeder, cost of labor, cost of line components, cost of energy, impedances, cost of taxes and maintenance, interest rate, and salvage values. All of these elements are contained in the data base.

Another major accomplishment of this research has been to integrate the siting and seizing, distribution, conductor size selection, routing and timing decisions all within one model. Current (1979) approaches still address these individually and then cycle through the calculations iteratively. It is an approach which precludes any optimality guarantee.

A further accomplishment has been to recognize, for the first time in the literature, a more realistic model of the electrical nature of the electrical distribution problem by explicitly including real and reactive power flows as variables. This allows capacitor and regulator allocation decisions to be modeled as mathematical programming decisions. In this approach, however, further research is needed into the solvability of the model and the determination of the degree of extension of the basic model.

In order to accomplish the integrated design, some unique features were developed for the database management, control and analysis system. The features are:

1. The graphical representation of the distribution network which uses line sections as nodes rather than edges.
2. The integrity constraints involving color coding to prevent designer errors.
3. The network editor whose design permit the distribution network to be searched for subnetwork configurations efficiently by treating subnetwork configurations as a text string.
4. The data base induced by the parameters of the optimization and analysis models.
5. A data model and form unique to the electric utility industry data systems. So far, no counter examples have been found as a result of the search in the literature or in the industry surveys.

Other than the assumptions previously discussed, no major assumptions are made. The assumptions implied in the analysis programs chosen which are the standard assumptions made to perform load flow, fault current, voltage profiles, etc., calculations. The thrust of the research has been to model more realistically and attempt to determine solution limitations. The major difference in this research work has been the use of the full modelling power of the MIP techniques as opposed to network flow models and dynamic programming approaches.

B. THE AMOUNT OF REQUIRED EFFORT TO EXPAND THE IMPLEMENTED SOFTWARE PACKAGE INTO A PRODUCTION GRADE PROGRAM

As was discussed in TASK B.12, only part of the conceptual model has been implemented. The major system components and their status is shown in Table 63. The effort to be spent upgrading the prototype system should be concentrated in four areas:

1. Making the shell program truly portable.
2. Implementing the network editor.
3. Restructuring some of the applications/analysis programs to take advantage of the network concepts built into the system.
4. Improving the performance of the database management system (DBMS).

The shell program functions as a minioperating system, providing file handling capabilities to the planner. Most operating system environments do not support such features. A notable exception is Unix, developed by Bell Laboratories. The availability of Unix to the research group permited the realization of the shell functions for the prototype system. However, transporting the shell to another environment such as the IBM 370 vs environment would be a difficult task but one which should be undertaken to meet the original goals of the methodology.

The network editor is perhaps the most innovative concept to emerge from the systems research done by this project. Lack of resources prevented its implementation but given the DBMS, the implementation would be straightforward.

Some of the algorithms used by the analysis programs should be changed to take advantage of the network orientation of the planning system. In particular, an investigation should be begun to determine if the network methods currently being studied are applicable to the planning problems discussed in TASK B.6. It appears that if they are applicable, speed improvements of as much as an order of magnitude could be anticipated.

Lastly, while the DBMS works efficiently with the prototype database, as the size of the database grows, the size requirements for real memory (as opposed to virtual memory) will also grow, resulting in slower access times. To counter this degradation, alternate access methods should be incorporated into the DBMS. This will improve performance and provide a more flexible system. Another feature which to date has not been included in the DBMS is an automatic restructuring capability to improve the way the data is organized internally. This feature should also be developed.

C. THE COST ESTIMATE TO EXPAND THE IMPLEMENTED SOFTWARE PACKAGE INTO A PRODUCTION GRADE PROGRAM

In order to design, develop and implement a production grade program, it is estimated that a two-year time schedule and a total estimated budget of $327,537 are required. The details of the estimated budget are given in the following tables.

Direct Labor	Estimated Hours	Rate/Hour	Estimated Cost	Total Estimated Cost
Principal Investigator	1,320	$13.88	$25,265	
2 Co-Principal Investigator	3,640	15.46	56,289	
2 Research Associates	2,080	15.08	31,359	
2 Research Assistants	6,240	5.77	36,000	
1 Secretary	2,080	4.29	8,929	
Total Direct Labor Cost				$157,842
Total Travel				$8,000
Other Direct Costs				
Fringe Benefits			$28,412	
Computer Costs (120 hrs. @ $250/hr)			30,000	
Supplies			6,000	
Communications			1,000	
Total Other Direct Costs				$65,412
Total Direct Costs				$231,254
Overhead				
61% of $157,842				$96,283
TOTAL ESTIMATED COST				$327,537

The following table presents the cost by individual fiscal years.

Cost Detail by Fiscal Year

	Year I	Year II	Total
Direct Labor	$76,020	$81,822	$157,842
Travel	$4,000	$4,000	$8,000
Other Direct Costs	$112,204	$119,050	$231,204
Overhead	$46,372	$168,961	$96,283
TOTAL ESTIMATED COSTS	$158,576	$168,961	$327,537

The details of the required budget for the first fiscal year are:

I. Salaries and Wages

A. Principal Investigator

9 months @ 0.25 FTE*	$5,156	
3 summer months @ 1.00 FTE	$6,375	

B. Co-Principal Investigator

9 academic months @ 0.25 FTE	$5,150	
3 summer months @ 1.00 FTE	$6,866	

C. Co-Principal Investigator

9 academic months @ 0.25 FTE	$6,338	
3 summer months @ 1.00 FTE	$8,450	

D. Research Associate - I

12 months @ 0.25 FTE	$7,300	

E. Research Associate - II

12 months @ 0.25 FTE	$7,633	

F. Research Assistant - I

12 months @ 0.50 FTE	$6,000	

G. Research Assistant - II

12 months @ 0.50 FTE	$6,000	

H. Research Assistant

12 months @ 0.50 FTE	$6,000	

I. Secretary ($4.12/hr)

12 months @ 0.50 FTE	$4,252	
Total		$76,020

II. Fringe Benefits

18% of the salaries and wages	$13,684

III. Indirect Costs

61% of the salaries and wages	$46,372

IV. Computer Usage

60 hours of CPU time @ $250/hr	$15,000

V. Miscellaneous

Supplies	$3,000	
Communication	$500	
Travel	$4,000	
		$7,500
TOTAL (For Year I)		$158,576

The details of the required budget for the second fiscal year are:

I. Salaries and Wages

A. Principal Investigator

9 academic months @ 0.25 FTE	$5,672
3 summer months @ 1.00 FTE	$7,562

B. Co-Principal Investigator

9 academic months @ 0.25 FTE	$5,665
3 summer months @ 1.00 FTE	$7,553

C. Co-Principal Investigator

9 academic months @ 0.25 FTE	$6,972
3 summer months @ 1.0 FTE	$9,265

D. Research Associate - I

12 months @ 0.25 FTE	$8,030

---

*FTE is used as an abbreviation for "full-time equivalent"

E. Research Associate - II
12 months @ 0.25 FTE $8,396

F. Research Assistant - I
12 months @ 0.50 FTE $6,000

G. Research Assistant - II
12 months @ 0.50 FTE $6,000

H. Research Assistant - III
12 months @ 0.50 FTE $6,000
$81,322

II. Fringe Benefits

18% of the salaries and wages $14,728

III. Indirect Costs

61% of the salaries and wages $49,911

IV. Computer Usage

60 hours CPU time @ $250/hr $15,000

V. Miscellaneous

Supplies $3,000
Communication $500
Travel $4,000
$7,500
TOTAL (For Year II) $168,961

APPENDIX A. BIBLIOGRAPHY OF DISTRIBUTION SYSTEMS PLANNING

INTRODUCTION

This section presents a bibliography of selected references pertaining to electrical distribution system planning. The objective is to encourage and facilitate broader use of automation concept in the distribution systems planning. References have been selected which deal predominantly with the distribution system. Emphasis is placed on references which illustrate practical as well as theoretical applications of distribution system planning techniques.

An extensive, but not necessarily complete, search has been made for the identification of papers and reports published in the open literature which describe the distribution planning methods and techniques. The listing of the titles is subdivided into three sections, depending upon the general substance of each article. However, a title may be listed in more than one section if the paper covers material in various sections.

The entries in each section are listed in alphabetical order. The last name of the first order author determines the alphabetical position. Many of the articles are available in abstract form in Science Abstracts, Section B, of the Engineering Index, and digesting or indexing periodicals, as well as in the original magazines listed. Only the more readily available foreign publications are included. A list of the periodicals which have been cited and their place of publication is given following the bibliography.

CLASSIFICATION OF LISTINGS

1. Analyses
2. Models
3. Techniques

1. ANALYSES

1. Baughman, M.L., and D.J. Bottaro, "Electric Power Transmission and Distribution Systems: Costs and Their Allocation," IEEE Trans. on PAS, Vol. PAS-95, No. 3, May/June, 1970, pp. 782-790.

2. Benson, R.U. and J.Z.D. Northcote-Green, "Development of a Data Base for Distribution Systems Planning," Conference Paper C73, pp. 374-376, IEEE PES Summer Meeting, July 15-20, 1973.

3. Billinton, R. and M.S. Grover, "Reliability Assessment of Transmission and Distribution Schemes," IEEE Trans. on PAS, Vol. PAS-94, No. 3, May/June, 1975, pp. 726, 732.

4. Billinton, R. and M.S. Grover, "Quantitative Evaluation of Permanent Outages in Distribution Systems," IEEE Trans. on PAS, Vol. PAS-94, No. 3, May/June, 1975, pp. 733-741.

5. Burstall, R.M., "Design of Electricity Distribution Networks--A Heuristic Approach," Interim Report of Univ. of Birmingham, Dept. of Engineering Production, March, 1964.

6. Campbell, H.E. and H.Z. Sinnott, "Economic Analysis of Residential Secondary Distribution Systems," AIEE Transactions Pt. IV (Power Apparatus and Systems), Vol. 76, Aug., 1957, pp. 574-585.

7. Campbell, H.E., "Economic Analysis of Distribution Systems," AIEE Transactions Pt. III (Power Apparatus and Systems), Vol. 79, Aug., 1962, pp. 423-440.

8. Chang, N.E., "Determination of Primary Feeder Losses," IEEE Trans. on PAS, Vol. PAS-87, No. 12, Dec., 1968, pp. 1991-1994.

9. Chang, N.E., "Determination and Evaluation of Distribution Transformer Losses of the Electric System Through Transformer Load Monitoring," IEEE Trans. on PAS, Vol. PAS-89, No. 6, July/Aug., 1970, pp. 1282-1284.

10. Chang, N.E., "Evaluation Distribution System Design by Cost Reliability Indices," Presented at the IEEE PES Winter Meeting, Feb., 1977.

11. Connor, R.A.W., et al., "Operational Statistics in the Management of Large Distribution Systems," Proc. IEE, Vol. 113, No. 11, Nov., 1966, pp. 1823-1834.

12. Converti, V., et al., "Optimized Distribution and Subtransmission System Planning by Digital Computer," I-System Philosophy, IEEE Trans. on PAS, Vol. PAS-83, April, 1964, pp. 341-347.

13. Cooper, J.M., "Optical Density Displays for Distribution in Planning," Rochester Gas and Electric Corporation Report, April, 1974.

14. Copland, F.G., "The Economics of Low-Voltage Electricity Supplies to New Housing Estates," Proc. IEE, Vol. 99, Paper No. 12365, Jan., 1952, pp. 95-101.

15. Eddy, F.N., "Communications Alternatives for Automated Distribution Systems," Control of Power Systems Conference, March 10-12, 1976, Oklahoma City, OK, pp. 128-133.

16. Eichenberger, Paul H. and F.M. Fine, "What Role Can Simulation Play in the Subtransmission and Distribution Planning Process?" Paper No. A754853 presented at the IEEE PES Summer Meeting, San Fransisco, Calif., July 20-25, 1975.

17. Ender, R.C., et al., "Optimized Distribution and Subtransmission Planning by Digital Computer," at the IEEE Summer Meeting, Toronto, Ontario, Canada, June 16-21, 1963.

18. Ender, R.C., G.G. Aver and R.A. Wylie, "Digital Calculation of Sequence Impedance and Fault Currents for Radial Primary Distribution Circuits," AIEE Transactions Pt. III (Power Apparatus and Systems), Vol. 79, Feb., 1961, pp. 1264-1277.

19. Gangel, M.W., et al., "Optimized Distribution and Subtransmission Planning by Digital Computer, Part IV-Primary Main Design," IEEE Trans. on PAS, Vol. PAS-84, December, 1975, pp. 1150-1164.

20. Gaussens, P., "Digital Computers in Distribution Network Studies," A Paper, Presented at Conference on Digital Computation for Electric Power Systems, Queen Mary College, London, September, 1963.

21. Gonen, T., et al., "A Research Progress Report for Electrical Energy Distribution Systems," University of Oklahoma, Norman, OK, Sept., 1978.

22. Gonen, T., F.E. Fagan, J.C. Thompson and C.M. Council, "Toward Automated Distribution Systems Planning," IEEE Control of Power Systems Conference, Texas A&M University, College Station, Texas, March 19-21, 1979.

23. Griscom, S.B., and R.Z. Lawrence, "Residential Distribution--An Analysis of Systems to Serve Expanding Loads," AIEE Transactions Part IV (Power Apparatus and Systems), Aug., 1956, pp. 535-542.

24. Grover, M.S. and R. Billinton, "Quantitative Evaluation of Overload Condition in Substation and Subtransmission Systems," Presented at the IEEE PES Winter Meeting, January, 1976.

25. Grover, M.S. and R. Billinton, "Quantitative Evaluation of Maintenance Policies in Distribution Systems," C75112-8, Presented at the IEEE PES Winter Meeting, New York, N.Y., Jan. 26-31, 1975.

26. Hassounch, Awwad S., "Load How Analysis of Power System Distribution Planning and System Modification," M.S. Thesis, The University of Texas at Arlington, May, 1974.

27. Hill, O.L. and W.A. Wittenberg, "Distribution Feeder Planning Technique," Transmission and Distribution, Vol. 26, October, 1974, pp. 66-69.

28. Hoestenbach, R.D., "Load Flow Analysis of Electrical Distribution Systems Using the Digital Computer," IEEE Trans. on Industry Applications, IA-12, No. 2, March/April, 1976, pp. 152-160.

29. Jomier, R., et al., "Concrete Experiment on the Administration of a Rural Network Using a Computer," Proc. Fifth Power System Computation Conference, Vol. 2, Paper 4.1/3, Cambridge, 1975.

30. Juricek, Michael J., Atsuhi Fukutome, and M.S. Chen, "Transportation Analysis of an Electric Power Distribution System," Paper No. A76052, presented at the IEEE PES Winter Meeting, New York, N.Y., January 25-30, 1976.

31. Kidder, A.H. and J.H. Neher, "Power Distribution System Parameters," AIEE Transaction (Power Apparatus and Systems), April, 1955, pp. 125-141.

32. Klinmann, Amikar H.R., "Maintenance of Voltage in Distribution Systems," Paper No. C75130-0, presented at the IEE PES Winter Meeting, New York, N.Y., Jan. 26-31, 1978.

33. Larson, P.E., "Economics of Distribution Systems," Electrical Engineering, Vol. 75, 1956.

34. Le Gal, "Recherche Automatique de Schemas de Secours Dans les Reseaux de Distribution Amoyonne Tension," Third Power Systems, Computation Conference, Rome, 1969.

35. Lemoine, J.C., "Planning the Development of High and Low Voltage Distribution Networks," Electricite de France, No. 1, 1975, pp. 5-18.

36. Mitchel, C.F., et al., "An Economic Analysis of Distribution Transformer Application," Paper No. 59-783, AIEE Summer Meeting, Seattle, Wash., June 21-26, 1959.

37. Morris, J.A., "Optimization of Power Line Conductor Through the Use of a Computerized Feeder Network Program," Modeling and Simulation Conference, University of Pittsburgh, PA, April 25-27, 1979.

38. Navas, F., H. Duran, M. Alvaro, "Studies in Power System Generation and Distribution Planning," Conference on Circuits, Systems, and Computers, Pacific Grove, CA, Nov. 27-29, 1973.

39. Northcote-Green, J.E.D., "Spatial Effects on Feeder Optimization in Distribution Planning," Modeling and Simulation Conference, University of Pittsburgh, PA, April 25-27, 1979.

40. Ossman, W.R., "Substation Expansion, Reliability and Transformer Loading Policy Analysis," IEEE Trans. on PAS, Vol. PAS-88, No. 8, August, 1969, pp. 1195-1205.

41. Rural Electrification Administration, "Voltage Drop Calculations," Bulletin 45-1, March, 1957.

42. Sarikas, R.H. and H.B. Thacker, "Distribution System Load Characteristics and Their Use in Planning and Design," AIEE Transactions Pt. III (Power Apparatus and Systems), Vol. 76, Aug. 1957, pp. 564-574.

43. Senupta, J.K., "A Generalization of Some Distribution Aspects of Chance—Constrained Linear Programming, International Econ. Review, June, 1970.

44. Smith, J.A., "Determination of Economical Distribution Substation Size," AIEE Transactions Part III (Power Apparatus and Systems), Vol. 80, Oct., 1961, pp. 663-669.

45. Smith, J.A., "Economics of Primary Distribution Voltages of 4.15 Through 34.5kV," AIEE Transactions Part III (Power Apparatus and Systems), Vol. 80, Oct., 1961, pp. 670-682.

46. Sommerfield, G., "Forecasting Load Growth," Transmission and Distribution, Vol. 26, Oct., 1974, pp. 44-46.

47. Stuart, T.A., "A Study of the Rochester District Based on Estimated Loads and Voltage Drops for 1977 and 1982," Rochester Gas and Electric Corporation Report, August, 1975.

48. Taylor, L.D., "The Demand for Electricity: A Survey," Bell J. Econ. Vol. 6, spring 1975, pp. 74-110.

49. Van Wormer, F.C., "Some Aspects of Distribution Load Area Geometry," AIEE Transactions Part IIIB (Power Apparatus and Systems), Vol. 3, pp. 1343-1349.

50. Wilde, R.J., J.T. Boardman, and D. Farguhar, "CAD System for Distribution Network Planning," IEE Conference Publication, No. 111, University of Southampton, Hamps, England, April 8-11, 1974, pp. 85-91.

51. Woodley, N.H., et al., "Subtransmission and Distribution Systems Expression by Simulation," A Paper, Presented at the 1969 American Power Conference.

52. Zimmerman, R.A. and H.E. Lokay, "Economic Comparison of Secondary Voltages 1-ϕ and 3-ϕ Distribution for Residential Areas," AIEE Transactions Part III, (Power Apparatus and Systems), Vol. 78, Aug., 1956, pp. 542-560.

2. MODELS

1. Adams, R.N., et al., "Mathematical Programming Systems in Electrical Power Generation, Transmission and Distribution Planning," in Proc. of 4th PSCC, Grenoble, France, 1972, Paper 1.1/3.

2. Adams, R.N. and M.A. Laughton, "Optimal Planning of Networks Using Mixed-Integer Programming,"

Proc. IEE, Vol. 121, No. 2, Feb., 1974, pp. 139-148.

3. Adams, R.N. and M.A. Laughton, "A Dynamic Programming/Network Flow Procedure for Distribution System Planning," Proceedings Power Industry Computer Applications Conference, June, 1975.

4. Anderson, D., "Models for Determining Least Cost Investments in Electricity Supply," Bell J. Econ., Vol. 3, spring 1972, pp. 267-301.

5. Armstrong, H.R., Z.A. Elliott, and H.O. Steolting, "Designing Distribution Lines for Better Performance," presented at the IEEE PES Winter Meeting, New York, NY, January 30-Feb. 4, 1966.

6. Avvakumov, V.G., A.I. Rybinnik, "Models of Multi-Purpose Optimization of the Quality of Electric Power in a Three-Phase Four-Conductor System." Izv Vyssh Uchebn Zaved Energy, No. 8, Aug., 1975, pp. 42-47.

7. Avvakumov, V.G., and A.M. Volotskii, "Concerning One Lexicographical Problem of Multiple-Target Optimization of Electric Power Quality," Izv Vyssh Uchebn Zaved Energy, Vol. 7, July, 1977, pp. 8-12.

8. Baldwin, C.J., "Techniques for the Simulation of Subtransmission and Distribution System Expansion," Proceedings Power Industry Computer Applications Conference, May, 1969.

9. Benson, R.U. and J.E.D. Northcote-Green, "Development of a Data Base for Distribution Systems Planning," Conference Paper No. C 73 374-6, IEEE PES Summer Meeting, July 15-20, 1973.

10. Boardman, J.T., and B.W. Hogg, "Computer Method for Design of Electricity-Supply Networks," Proc. IEE, Vol. 119, 1972, pp. 851-856.

11. Boardman, J.T. and R.J. Wilde, "Interactive Design of Electricity-Distribution Networks," Proc. IEE, Vol. 120, 1973, pp. 1006-1011.

12. Boess, M.R., Jr., et al., "Optimized Distribution and Subtransmission Planning by Digital Computer: V-The Total System," IEEE Trans. on PAS, October, 1966.

13. Burandt, R.P., "Optimizing Distribution Transformer and Secondary Design; Present-Worth Approach." AIEE Trans. Pt. III (Power Apparatus and Systems), August, 1961, pp. 345-355.

14. Burandt, R.P., et al., "Present Worth Approach for Optimizing Distribution Transformer and Secondary Designs to Serve Growing Loads," AIEE Transactions Part IV, (Power Apparatus and Systems), Vol. 80, August, 1961, pp. 345-350.

15. Burstall, R.M., "Computer Design of Electricity Supply Networks Using a Heuristic Method." Compt. J. Vol. 9, 1966, pp. 263-274.

16. Carson, M.J. and G. Cornfield, "Design of Low-Voltage Distribution Networks; Interactive Computer Method Based on the Calculus of Variations," Proc. IEE, Vol. 120, 1973, pp. 585-592.

17. Chen, M.S., W.D. Dickson and H.C. Cheng, "Simulation Technique for Rural Electric Power System Planning," IEEE Trans. on Industrial Applications, Vol. 1A-9, 1973, pp. 98-105.

18. Chia, C.S., "Optimized Distribution Design," A Ph.D. Thesis, University of Manchester, Manchester, England, 1971.

19. Converti, V., et al., "Optimized Distribution and Subtransmission System Planning by Digital Computer, I-System Philosophy," IEEE Trans. on PAS, Vol. PAS-85, April, 1974, pp. 341-347.

20. Cooper, J.M., et al., "Distribution Load Model of the Rochester Gas and Electric Corporation Report," July, 1973.

21. Cooper, J.M. and T.A. Stuart, "A Computer Graphics Approach to Distribution Planning," Paper No. C 74 345-4, Presented at the IEEE PES Summer Meeting, July, 1974.

22. Cornfield, G. and M.J. Carson, "Design of Low-Voltage Distribution Networks," Inst. Elec. Eng., Paper 6918 P. Nov., 1972.

23. Couch, G.H. and I.F. Morrison, "Substation Switching—An Approach to the Determination of Optimal Sequences," in Proc. 4th PSCC, Grenoble, France, 1972, Paper 2.1/1.

24. Crawford, Dale M. and S.B. Holt, Jr., "A Mathematical Optimization Technique for Locating and Sizing Distribution Substations and Driving Their Optimal Service Areas," IEEE Trans. on PAS, Vol. PAS-94, No. 2, March/April, 1975, pp. 230-235.

25. Denton, M.J. and D.N. Reps, "Distribution Substation and Primary Feeder Planning," AIEE Transactions Pt. III (Power Apparatus and Systems), Vol. 74, June 1955, pp. 484-498.

26. Garrett, G.P., "Expansion Planning of Radial Subtransmission Systems," Thesis, University of Texas at Arlington, May, 1974.

27. Garrett, Gary P., Atsushi Fukutome, and M.S. Chen. "Expansion Planning of Radial Subtransmission Systems," IEEE Trans. on PAS, Vol. PAS-96, No. 5, Sept./Oct., 1977, pp. 128-153.

28. Giovannini, M. Reynand, G., Salvaderi, L., "Selection of Substation Layout in Power Systems," IEE Reliability Conference of Power Supply Systems, London, England, Feb., 1977.

29. Gonen, T., et al., "A Research Progress Report for Electrical Energy Distribution Systems," University of Oklahoma, Norman, OK, Sept., 1978.

30. Gonen, T., J.C. Thompson, and R.M. Cubert, "An Interactive Distribution System Planning Model," Modeling and Simulation Conference, University of Pittsburgh, PA, April 25-27, 1979.

31. Gonen, T. and B.L. Foote. "Applications of Mixed-Integer Programming to Reduce Sub-Optimization in Distribution Systems Planning," Modeling and Simulation Conference, University of Pittsburgh, PA, April 25-27, 1979.

32. Grimsdale, R.L. and P.H. Sinclare, "Design of Housing Estate Distribution Systems Using a Digital Computer," Proc. Inst. Elec. Eng., Pt. A, Vol. 33, 1960.

33. Hassounch, A.S., "Load Flow Analysis of Power System Distribution Planning and System Modification," M.S. Thesis, The University of Texas at Arlington, May, 1976,

34. Hill, O.L., and W.A. Wittenberg, "Distribution Feeder Planning Technique," Transmission and Distribution, Vol. 26, October, 1974, pp. 66-69.

35. Hindi, K.S., "Distribution Networks Design," A M.S. Thesis, University of Manchester, Manchester, England, 1972.

36. Hindi, K.S., "Optimization of Gas Storage and L.V. Distribution System," A Ph.D. Thesis, University of Manchester, Manchester, England, 1975.

37. Hindi, K.S. and A. Brameller, "Optimal Cable Profile of L.V. Radial Distributors: Two Mathematical Programming Methods," Proc. IEE, Vol. 123, No. 4, 1976, pp. 331-334.

38. Hindi, K.S. and A. Bramellar, "Design of Low-Voltage Distribution Networks: A Mathematical Programming Method," Proc. IEE, Vol. 124, Jan., 1977, pp. 54-58.

39. Juricek, Michael J., "Application of the Digital Computer to Substation Planning," M.S. Thesis, The University of Texas at Arlington, August, 1974.

40. Juricek, M.J., Atsushi Fukutome, and M.S. Chen, "Transportation Analysis of an Electric Power Distribution System," Paper No. A 76 052-1, Presented at the IEEE PES Winter Meeting, New York, NY, Jan. 25-30, 1976.

41. Knight, U.G.W., "The Logical Design of Electrical Networks Using Linear Programming Methods," Proc. IEE, Vol. 107A, 1960, pp. 306-314.

42. Kujsyczyk, S., "Automatic Design of Housing-Estate Distribution Networks," Proc. IEE, Vol. 121, 1974, pp. 488-490.

43. Lawrence, R.F., et al., "Distribution System Planning Through Optimized Design, IV-Choice of Voltage Regulation Point in the System," Paper No. 60813, AIEE Meeting, 1960.

44. Lawrence, R.F., et al., "Automated Distribution System Planning," IEEE Trans. on PAS, Vol. PAS-83, April, 1964, pp. 311-316.

45. Lawrence, R.F., D.N. Reps, and A.D. Patton, "Distribution System Planning Through Optimized Design: Part I-Distribution Transformers and Secondaries," AIEE Transactions, Part III, Vol. 79, June, 1960, pp. 199-204.

46. Lawrence, R.F., D.N. Reps, and A.D. Patton, "Distribution System Planning Through Optimized Design: Part II-Comparative Economics of System Voltages," Ibid., pp. 204-223.

47. Le Gal, "Recherche Automatique de Schemes de Secours dans les Reseaux de Distribution a Moyenne Tension," Third Power Systems Computation Conference, Rome, Italy, 1969.

48. Lemoine, J.C., "Development Planning of High and Low Tension Distribution Networks, Optimal Investment Research Some General Ideas in Regard to Methods," EDF Bulletin De La Directions Des Etudes et Rechereches, Series B Resauss Electriques, ERDA-TR-231, 1975.

49. Lemus, J.L., H.F. Rutz and F.A. Viramontes, "Conductor Optimization and Optimal Capacitor Location in Distribution Systems," Mexicon 77: Conf. Int. y Expo Sobre Invest Desanollo y Apl de la Ing. Electr. y Electron., Mexico City, Mexico, Nov. 9-12, 1977, published by IEEE, New York, NY.

50. Lutz Bochanky, Dresden, "Optimum Electric Power Distribution System for New Residential Areas," ERDA-TR-228.

51. Masud, Enver, "An Interactive Procedure for Sizing and Timing Distribution Substations Using Optimization Techniques," Paper No. T 74142-6, Presented at the IEEE PES Winter Meeting, New York, New York, January 27-Feb. 1, 1974.

52. Masud, Enver, "Design for an Automatic Distribution Substation Planning Program," Proceedings, Frontiers of Power Technology Conference, Oklahoma State University, Oct., 1969.

53. Masud, Enver, "A Model for Determining the Optimal Expansion Policy for Power Distribution Substations," M.S. Thesis, University of Oklahoma, Norman, OK, 1972.

54. Merlin, A. and H. Back, "Search for a Minimal Loss Operating Spanning Tree Configuration in an Urban Power Distribution System," Power Systems Computation Conference, 1975.

55. Munasinghe, Mohan and W.G. Scott, "Long Range Distribution System Planning Based on Optimum Economic Reliability Levels," Paper No. A 78 576-1, Presented at IEEE PES Summer Meeting, Los Angeles, California, July 16-21, 1978.

56. Munasinghe, Mohan and W.G. Scott, "Engineering Management and Control of Distribution Systems via Mini-Computers," IEEE Control of Power Systems Conference, Oklahoma City, Oklahoma, March, 1978.

57. Muradyan, A.E., Y.S. Smotritskii, "Optimization of the Configuration of Power Distribution Networks with Obstacles at the Branching Points," Izv Akad Nauk Energy Transp. No. 5, Sept.-Oct., 1973, pp. 74-80.

58. Neal, C.W., "Long Range Distribution Planning on the OG&G System," Power Distribution Conference, University of Texas, Austin, Texas, October 20-22, 1958.

59. Okada, T. and C. Uenosona, "Long Term Development Planning of Distribution Systems," Electrical Engineering in Japan, Vol. 90, No. 12, December, 1970, pp. 273-283.

60. Okada, T. and C. Uenosona, "Mathematical Models for Optimal Design of Power Distribution Systems for Long Terms," in Proc. 4th PSCC, Grenoble, France, 1972, Paper 1.1/3.

61. Oldfield, J.V. and T. Lang, "The Long Term Design of Electrical Power Distribution Systems," IEEE Power Industry Computer Applications Conference, May, 1965.

62. Querubin, R. Guervo, G., and Navas, G., "On the Optimum Design of Radial Distribution Networks," PICS Conference, Bogota, Columbia, 1975.

63. Rhoten, G.P., "A General Program for Processing Distribution Design," Electrical World, Vol. 155, December 10, 1961, pp. 45-58.

64. Rural Electrification Administration, "System Planning Guide for Electric Distribution Systems," Bulletin 60-8, June, 1957.

65. Sander, B.M., "Forecasting Residential Energy Demand: A Key to Distribution Planning," Paper No. A77642-2, Presented at the IEEE PES Summer Meeting, Mexico City, Mexico, July 17-22, 1977.

66. Scott, W.G., "Computer Model Forecasts Future Loads," Electrical World, Vol. 178, Sept. 15, 1972, pp. 114-115.

67. Segner, Konstantin, D. Graf, and K.F. Peter, "Planning Techniques for Distribution Networks with the Optimization of Plant Parameters," Energietechnik, Vol. 25, No. 10, Oct. 1975, pp. 459-462.

68. Shapiro, I.Z., "Probalistic-Statistical Models for Determining and Predicting Power Losses in 6-10 kV Distribution Networks," Izv Vyssh Uchebn Zaved Energy, No. 4, April, 1978, pp. 15-19.

69. Shelton, S.E., "A Direct Optimization Approach to the Economic Expansion of Distribution Substations," M.S. Thesis, Iowa State University, Ames, Iowa, 1977.

70. Shelton, S.E. and A.A. Mahmoud, "A Model for Economic Expansion of Distribution Substations," A Paper, presented at the IEEE Region 5 Annual Conference, Tulsa, OK, April 16-18, 1978.

71. Shelton, S.E. and A.A. Mahmoud, "A Direct Optimization Approach to Distribution Substation Expansion," Paper No. A 78 592-8, Presented at the IEEE PES Summer Meeting, Los Angeles, CA, July 16-21, 1978.

72. Smith, J.A., "Determination of Economical Distribution Substation Size," AIEE Transactions Part III (Power Apparatus and Systems), Vol. 80, Oct., 1961, pp. 663-669.

73. Snelson, J.K. and M.J. Carson, "Logical Design of Branched L.V. Distributors," Proc. IEE, Vol. 117, No. 2, pp. 415-420, 1970.

74. Stover, Carl N. Jr., et al., "A Financial Forecasting Model for Rural Electric Distribution Systems," Paper No. C 74 352-1, Presented at the IEEE PES Summer Meeting, Anaheim, Calif., July 14-19, 1974.

75. Stover, Carl N., Jr., "A Planning Model Analysis of Long Range Distribution System Design Alternatives," Paper No. C 73 468-6, Presented at the IEEE PES Summer Meeting, Vancouver, Conn., July 15-20, 1973.

76. Stuart, T.A., "Electrical Distribution Planning Program VREG," Rochester Gas and Electric Corporation Report, July, 1973.

77. Stuart, T.A. and J.M. Cooper, "An Interactive Approach to Distribution Planning Using Computer Graphics," Paper No. C 75 095-5, Presented at IEEE PES Winter Meeting, New York, NY, Jan. 26-31, 1975.

78. Venikov, Va., Rv. Shnell, "Modeling and Automation in the Designing of the Electrical Power Equipment," Elektrichestvo, No. 8, Aug., 1970, pp.1-6.

79. Wall, D.L., G.L. Thompson, and J.E.D. Northcote-Green, "An Optimization Model for Planning Radial Distribution Networks," Paper No. F 78 653-8 presented at the IEEE PES Summer Meeting, Los Angeles, CA, July 16-21, 1978.

80. Weber, K., "Operations Research Models for Electric Distribution Systems," Bull Assoc Suisse Electr, Vol. 65, No. 9, May 4, 1974, pp. 659-664.

81. Webler, R.M., et al., "Secondary System Distribution Planning for Load Growth," AIEE Transactions Part IV (Power Apparatus Systems), December, 1963, pp. 908-927.

82. Woodley, N.H. and R.U. Benson, "Distribution System Planning Program, DSP," Westinghouse Electric Corporation Report No. 68-681, July 1, 1968.

3. TECHNIQUES

1. Adams, R.N. and M.A. Laughton, "Optimal Planning of Power Networks Using Mixed-Integer Programming," Proc. IEE, Vol. 121, No. 2, Feb., 1974, pp. 139-148.

2. Ambrosius, C.C. and R.H. Sarikas, "Planning Rural Systems for Continuing Growth," AIEE Transactions, Part III (Power Apparatus and Systems), Oct., 1960, pp. 623-628.

3. Baldwin, C.J., "Techniques for the Simulation of Subtransmission and Distribution System Expansion," Proc. Power Industry Computer Applications Conference, May, 1969.

4. Boardman, J.T. and B.W. Hogg, "Computer Method for Design of Electricity-Supply Networks," Proc. IEE, Vol. 119, No. 7, pp. 51-856, 1972.

5. Boardman, J.T. and R.J. Wilde, "Interactive Design of Electricity-Distribution Networks," Proc. IEE, Vol. 120, 1973, pp. 1006-1011.

6. Burstall, R.M., "Computer Design of Electricity Supply Networks Using a Heuristic Method," Comput. J., Vol. 9, 1966, pp. 263-274.

7. Carson, M.J. and G. Cornfield, "Design of Low-Voltage Distribution Networks," Proc. IEE, Vol. 120, May, 1973, pp. 585-592.

8. Chen, M.S., W.D. Dickson, H.C. Cheng, "Simulation Technique for Rural Electric Power System," IEEE Trans. on Industrial Applications, Vol. 212-9, 1973, pp. 98-105.

9. Connor, R.A.W., et al., "Operational Statistics in the Management of Large Distribution Systems," Proc. IEE, Vol. 113, No. 11, Nov., 1966, pp. 1823-1834.

10. Copland, F.G., "The Economics of Low-Voltage Electricity Supplies for New Housing Estates," Proc. IEE, Vol. 99, Paper No. 12365, pp. 95-101, January, 1952.

11. Couch, G.H. and I.F. Morrison, "Substation Switching--An Approach to the Determination of Optimal Sequences," in Proc. 4th PSCC, Grenoble, France, 1972, Paper 2.1/1.

12. Davies, M., "Design of L.V. Distributors from Standard Cable Sizes," Proc. IEE, Vol. 112, No. 5, pp. 949-956, 1965.

13. Economakos, E., "Distribution System Voltage Studies Using Simulation Techniques," Proc. IEE, Vol. 125, No. 2, Feb., 1978, pp. 151-152.

14. Ender, R.C., G.G. Auer, and R.A. Wylie, "Digital Calculation of Sequence Impedance and Fault Currents for Radial Primary Distribution Circuits,"

AIEE Transactions Part III (Power Apparatus and Systems), Vol. 79, Feb., 1961, pp. 1264-1277.

15. Erlandson, N.H., "Selection Design and Operation of 13.21/231 as a Distribution Voltage," AIEE Transactions Part III (Power Apparatus and Systems), Vol. 77, Oct., 1958, pp. 775-781.

16. Fitzgerald, J.P., et al., "Evaluation of Substation Loading Using Simulation Techniques," IEEE Trans. on PAS, April, 1964.

17. Gonen, T., F.E. Fagan, J.C. Thompson, and C.M. Council, "Toward Automated Distribution Systems Planning," IEEE Control of Power Systems Conference, Texas A&M University, College Station, Texas, March 19-21, 1979.

18. Gordon, H., "Maximum Utilization of Distribution Substation Capacity," Presented at the Missouri Valley Electric Association 48th Annual Engineering Conference, April 13-15, 1977.

19. Grimsdale, R.L. and P.H. Sinclare, "Design of Housing Estate Distribution Systems Using a Digital Computer," Proc. Inst. Elec. Eng., Pt. A, Vol. 35, 1960.

20. Hill, O.L. and W.A. Wittenberg, "Distribution Feeder Planning Technique," Transmission and Distribution, Vol. 26, Oct., 1974, pp. 66-69.

21. Hindi, K.S., and A. Brameller, "Optimal Cable Profile of L.V. Radial Distributors: Two Mathematical Programming Methods," Proc. IEE, Vol. 123, No. 4, 1976, pp. 331-334.

22. Hindi, K.S. and A. Brameller, "Design of Low-Voltage Distribution Networks: A Mathematical Programming Method," Proc. IEE, Vol. 124, No. 5, Jan., 1977, pp. 54-58.

23. Hopkins, D.C. and D.R. Samson, "Useful Methods for Determining Primary Feeds Points in Future Distribution System Planning," AIEE Transactions, Part IV, (Power Apparatus and Systems), August, 1954, pp. 856-863.

24. Juricek, Michael J., "Application of the Digital Computer to Substation Planning," M.S. Thesis, The University of Texas at Arlington, Aug., 1974.

25. Klein, K.W., "Computer Gives Economic Loading for Transformer and Secondary," Electrical World, New York, NY, Nov. 30, 1959, pp. 83-85.

26. Knight, U.G.W., "The Logical Design of Electrical Networks Using Linear Programming Methods," Proc. IEE, Vol. 107A, 1960, pp. 306-316.

27. Knight, U.G.W., "Economic Design and Operation of Power Systems," Electrical Revue, 1963, pp. 124-130.

28. Knight, U.G.W., "Economic Design and Operation of Power Systems," Electrical Revue, 1963, pp. 163-166.

29. Kujszczyk, S., "Automatic Design of Housing-Estate Distribution Networks," Proc. IEE, Vol. 121, 1974, pp. 488-490.

30. Ladd, R.D., "How to Determine Distribution Transformer Utilization," Transmission and Distribution, Cos Cob, Conn., Jan., 1960.

31. Light, John, "An Economic Approach to Distribution Conductor Size Selection," Presented at the Missouri Valley Electric Association 49th Annual Engineering Conference, April 12-14, 1978.

32. Lonsdale, J.G. and G.B. Hitchen, "Reliability Evaluation in the Planning of Distribution Systems," IEE Conference on Reliability of Power Supply System, London, England, February 21-23, 1977.

33. Ludot, J.P. and M.C. Rubinstein, "Methods pour la Pionification a Court Terme des Reseaux de Distribution," in Proc. 4th PSCC, Grenoble, France, 1972, Paper 1.1/12.

34. Lutz Bochanky, Dresden, "Optimum Electric Power Distribution System for New Residential Areas," ERDA-TR-228.

35. McNab, N.A. and A.J. Bode, "Lower Cost Urban Electricity-Distribution Systems by Automatic Control Application to 11 kV and 33kV Systems," IEE Proc. Vol. 120, 1973, pp. 445-446.

36. Morrison, C., "A Linear Approach to the Problem of Planning New Feed Points Into a Distribution System," AIEE Transactions Part III (Power Apparatus and Systems), 1963, pp. 819-832.

37. Myslin, D.A., "Certain Questions of Automating the Design of Regional Substations," Isvestiva Akademii Nauk SSR, No. 5, ERDA-TR-233, 1973.

38. Orth, E.J., Jr. and J.L. Redmond, "A Digital Tool for Planning Radial Distribution Systems," Proceedings Power Industry Computer Applications Conference, April, 1963.

39. Paradis, C. and D. Mukhedkas, "Computer Aided Distribution Design," The Institution of Engineers, Australia, 1974.

40. Rhoten, G.P., "A General Program for Processing Distribution Design," Electrical World, Vol. 156, Dec. 16, 1961, pp. 45-48.

41. Rural Electrification Administration, "System Planning Guide for Electric Distribution Systems," Bulletin 60-8, June, 1957.

42. Sander, B.M., "Forecasting Residential Energy Demand: A Key to Distribution Planning," Paper No. A 77 642-2, Presented at the IEEE PES Summer Meeting, Mexico City, Mexico, July 17-22, 1977.

43. Sarikas, R.H. and H.B. Thacker, "Distribution System Load Characteristics and Their Use in Planning and Design," AIEE Transactions, Part III (Power Apparatus and Systems). Vol. 76, Aug., 1957, pp. 564-574.

44. Segner, Konstantin, D. Graf, and K.F. Peter, "Planning Techniques for Distribution Networks with the Optimization of Plant Parameters," Energietechnik, Vol. 25, No. 10, Oct., 1975, pp. 459-462.

45. Snelson, J.K. and M.J. Carson, "Logical Design of Branched L.V. Distributors," Proc. IEE, Vol. 177, 1970, pp. 415-420.

46. Strausser, E.A., "Digital Computer Applied to Electrical Distribution System Planning," AIEE Transactions Part III (Power Apparatus and Systems), Vol. 77, 1959, pp. 1244-1252.

47. Stuart, T.A., "Electrical Distribution Planning

Program VREG," Rochester Gas and Electric Corporation, on Report, July, 1975.

48. Stuart, T.A. and J.M. Cooper, "An Interactive Approach to Distribution Planning Using Computer Graphics," Paper No. C 75 095-5, Presented at IEEE PES Winter Meeting, New York, NY, Jan. 26-31, 1975

49. Woodley, N.H. and R.U. Benson, "Distribution System Planning Program, DSP," Westinghouse Electric Corporation Report No. 68-681, July 1, 1968.

INDEX OF (FIRST-NAMED) AUTHORS

INDEX OF SOURCES

AIEE Transactions	New York, N.Y.
ASEA Journal	Sweden
Distribution (General Electric Co.)	Schnectady, N.Y.
Electrical Engineering in Japan	Tokyo, Japan
Electric Light and Power	Chicago, IL
Electricite de France	Paris, France
EPRI Journal	Palo Alto, CA

Elektrichestvo	Moscow, USSR
Electrotechniek	Netherlands
Electricity Council	London, England
Electrical Review	London, England
Electric Technology	Moscow, USSR
Electric Times	London, England
Electric World	New York, N.Y.
Hidroelectrica Espanola	Madrid, Spain
Institute of Electrical Engineers	London, England
IEEE Transactions on PAS	New York, N.Y.
Institute of Engineers	New Delphi, India
Isvestiya Akademii Nauk	Moscow, USSR
Ontario Hydro Research Quarterly	Toronto, Canada
Osterreichische Zeitschrift fur Elektrizitatswirtschaft	Bonn, West Germany
Proceedings of the American Power Conference	Chicago, IL
Proceedings of the IEE	London, England
Revue Generale de L'Electricite	Paris, France
Transmission and Distribution	Cos Cob, Conn.
Westinghouse Engineer	Pittsburgh, PA

APPENDIX B. A SAMPLE COPY OF THE QUESTIONNAIRE

DEVELOPMENT OF ADVANCED METHODS FOR PLANNING ELECTRIC ENERGY DISTRIBUTION SYSTEMS

1. Do you plan for long range? Yes________ No________

   Is long range planning defined in terms of load levels? Yes________ No________

   If yes, what % above present levels? ________

   What other factors determine if the plan is long range? ________

2. Do you plan for short range? Yes________ No________

   What time span would be considered short range?

   ____

   What other factors besides years, determine if a plan is short range? ________

3. What factors do you consider in determining boundaries for distribution service areas?

   ____

4. What factors are considered in load level projects?

   ____

   How are the various factors considered in load projections quantitatively? ________

5. What size of areas is used for load projections:

   In urban areas? ________

   In rural areas? ________

6. How are future substation sites picked? ________

7. How many years before a substation is constructed is the site purchased?

   Average time ________

   Shortest time ________

8. If a substation exists, what are the major factors considered to determine ultimate expansion capability?

   A.

   B.

   C.

   D.

   E.

   F.

   G.

9. What transmission voltages are the substation planned for? ________

10. What distribution voltages are considered in economic expansion studies? ________

11. When comparative plans are made for serving developing areas, what are the major alternatives considered?

    A.

    B.

    C.

    D.

    E.

    F.

    G.

12. How are results of Question 11 presented for:

    A. Engineering Review ________

    B. Operating Review ________

    C. Implementing Plan ________

13. How do the developing load patterns in an area feed back to the planner to cause changes in plans considered in Question 11? ________

____

____

14. Do you have available, computer time for distribution planning? Yes__________ No__________

15. Would an interactive distribution planning computer model improve your planning?
Yes__________ No__________ If yes, why?

____

____

____

____

If no, why? ______________________________

____

____

16. Is reliability considered in distribution planning? Yes__________ No__________

If yes, how is it factored into plans? ________

____

____

____

17. If the answer to Question 16 is yes, what limitations by:

A. $

B. Customers

C. Load Levels

D. Others

are placed on reliability in distribution planning?

18. Do you value the worth of service (revenue loss) for customers? Yes __________ No__________

If yes, do you value the worth of service differently by class of customer? Yes______ No______

If valued differently, what are typical values assigned per:

Customer ______________________________

KVA ______________________________

KWH ______________________________

Other ______________________________

19. What models do you use in distribution planning?

____

____

APPENDIX C

RESEARCH ADVANCES IN SOLUTION METHODOLOGY FOR DETERMINING MINIMUM COST EXPANSION OF SUBSTATION CAPACITY FOR POWER DISTRIBUTION SYSTEMS

The problem of least cost expansion of substation capacity for power distribution systems requires the determination of three critical factors:

(1) The identification of the particular substations that should be given additional capacity.

(2) The determination of how much additional capacity each substation should receive.

(3) The specification of when the additional capacity should become available.

These determinations must be based on constraints of cost, load, voltage, and reserve requirements. They also depend importantly on the decision concerning the distribution of loads among substations.

The complexity and interacting effects of these considerations have been captured in a mathematical model by Masud (13), which characterizes the problem as 0-1 "Multiple Choice" mathematical optimization problem with generalized upper bounding (GUB) constraints. In order to realize the potential benefits from this model, it is necessary to identify a way to solve it effectively, thereby determining the implications of its irrelated underlying factors for real-world decisions. Specifically, as noted in our proposal*, an efficient tailored, computer solution routine is required that operates on the model structure in a manner to determine the optimal (least cost) plan for expanding substation capacity to meet electrical power and energy needs.

This report presents the innovations that have been made in solution technology under this proposal as a foundation for the full scale effort of developing a completely integrated, operational, computer solution routine. These innovations succeed in advancing the state-of-the-art in mathematical programming solution techniques. In particular, we report a new computer solution methodology for the linear programming (LP)/generalized upper bound (GUB) knapsack problem.

The LP/GUB knapsack problem is a major subproblem of the least cost substation capacity expansion model, upon which the solution of the total problem crucially depends. More precisely, the LP/GUB knapsack problem is the mathematical relaxation of the substation capacity expansion problem generated by means of a surrogate constraint solution strategy. The requirement of repeated solution of this subproblem (many hundreds or even thousands of times) in the surrogate strategy dictates that it be solved in the most efficient possible manner.

Our research provides a solution procedure for the LP/GUB knapsack problem that is notably superior to any previously known. This claim is established by the derivation of computational bounds (Theorem 3 and Corollaries 4 and 5, following), that dominate the best bounds in the literature. The possibility of obtaining an improved method, foreshadowed by a result reported in the initial proposal, becomes manifest in this new development. In fact, for the case in which each GUB set is the same, the order of complexity bound for our procedure has the same form as that for the ordinary knapsack problem. This surprising result means that the LP/GUB knapsack problem is susceptible to solution with unprecedented efficiency.

The main body of the report details the background and significance of the LP/GUB knapsack problem. The mathematical characterizations of this problem and our results for solving the problem by a new specialization of the dual simplex algorithm are then elaborated. We develop the order of complexity bounds for two variants of our method and show that both of these variants dominate previous approaches over practical ranges of the parameters.

The compelling application of these new results to the substation capacity expansion problem, as previously noted, arises by making them the driving mechanism behind the implementation of surrogate constraint relaxation strategies. More specifically, since the generation and solution of surrogate constraint relaxations is identified in our proposal as a major component of a specialized solution algorithm for the substation capacity expansion problem, the ability to solve LP/GUB knapsacks efficiently will materially improve the ability to solve the substation capacity expansion problem effectively. In addition, these innovations will lead to improved applications of branch and bound exclusion tests.

Certainly, the complete effectiveness of these results must depend on the process of generating the surrogate constraints themselves, and on the manner in which the LP/GUB knapsack solutions are integrated with the other main components of the total solution procedure. This involves the development of specialized branch and bound fathoming techniques and associated list processing and labeling algorithms. In general, the algorithmic and programming procedures described in our proposal emerge as the essential efforts to take maximum advantage of the new LP/GUB knapsack solution methodology developed here. By means of these specialized algorithmic and software development efforts, the envisioned possibility of a solution method 3 to 7 times more efficient than the best available commercial integer programming software stands to be realized or even surpassed, thereby yielding important cost savings in the computer solution of the substation capacity expansion problem. These savings will be further magnified by the ability to obtain comparable gains in efficiency in post-optimality analyses designed to ask "what if" questions about changes in options and cost structures relating to substation capacity expansion. In this fashion, the present results provide a foundation for an integrated solution routine of particular importance to the broader goals of planning to meet future energy needs in the generation and distribution of electrical power.

1. HISTORICAL BACKGROUND

A good deal of attention has been given to standard LP knapsacks for their role as relaxations in branch and bound methods for solving integer knapsack problems [2, 5, 9]. Such problems have been studied as an end in themselves, and also as surrogate constraint relaxations for more general 0-1 integer programming

---

*Proposal for Electrical Energy Distribution Systems in response to ERDA RFP No. EC-77-R-02-0034 submitted by University of Oklahoma, Office of Research Administration, 1000 Asp Avenue, Room 314, Norman, OK 73019.

(IP) problems.

Many 0-1 IP problems, however, are of the "multiple choice" variety, attended by the requirement that the variables of partitioned subsets sum to one. Specialized IP methods for problems involving such generalized upper bound (GUB) constraints have been proposed in settings of varied generality (e.g., [3, 4, 6]), and recently some attention has been given to integer knapsacks with GUB constraints [14, 15]. To solve these and more general problems using LP and surrogate relaxations, it is important to be able to solve LP/GUB knapsacks efficiently. It is also valuable to be able to solve LP/GUB knapsack problems to accelerate the solution of ordinary LP/GUB problems by the dual simplex method, as pointed out by Witzgall [16]. Consequently, the goal of this paper is to develop an algorithm for the LP/GUB knapsack problem that is both easily implemented and highly efficient.

Two earlier papers dealing with this problem (in slightly less general form than treated here) are worthy of special note. The paper by Sinha and Zoltners [15] is the first to identify the characteristics of the undominated solution space for the case in which the knapsack is an inequality constraint. These authors then develop a method that is reported to speed the branch and bound solution of the integer GUB knapsack problem. The second paper, due to Witzgall [16], examines the case where the knapsack is an equality constraint spanned by the GUB sets. Witzgall's work is especially notable for its geometric characterizations and the specification of "worst case" computational bounds for his algorithm. In particular, the algorithm of [16] is shown to be of complexity O(n log n) + O(m(n-m)), where n is the number of variables and m is the number of GUB sets. This is the first result that bounds the complexity of the LP/GUB knapsack problem in this manner.

In this paper we use an alternative framework that focuses directly on properties of the dual simplex method applied to the LP/GUB knapsack problem. After specifying necessary and sufficient conditions for dual feasible bases, we identify relationships that hold automatically in the application of the dual simplex method. These relationships are then utilized to develop a specialized version of this method which is shown to be of complexity at most O(n(log n + log m)), or in the case where each GUB set contains the same number of elements, O(n log n). These bounds are interesting not only because they reduce the previous estimate of the order complexity of the LP/GUB knapsack problem, but also because they reduce to the same form as one of the standard algorithmic bounds for the ordinary LP knapsack problem without GUB constraints, thereby establishing a connection between these more and less general problems.

2. PROBLEM NOTATION

The LP/GUB knapsack problem may be written

$$\text{Minimize} \sum_{j\in N} c_j x_j \tag{1}$$

$$\text{subject to} \sum_{j\in N} a_j x_j = a_o \tag{2}$$

$$\sum_{j\in J_k} x_j = 1;\ k \in M = \{1,\ldots,m\} \tag{3}$$

$$x_j \geq 0;\ j \in N = \{1,\ldots,n\}$$

where $J_p \cap J_q = \emptyset$ for $p \neq q$ and $J = \bigcup_{k\in M} J_k \subseteq N$.

There are no restrictions on any of the problem coefficients (a_o, a_j, c_j), except that we exclude the trivial situation in which $a_j = 0$ for $j \in N - J$.

Two subcases of interest included by our results are for $N = J$ (as in Witzgall [16]) and for $N - J = \{n\}$, where x_n is a slack or surplus variable (as in Sinha and Zoltner [15]). We will comment on the specializations of our results to these subcases at appropriate points.

To begin, we make a simple and well known observation concerning the structure of basic solutions for this problem.

Remark 1. In every basic solution to the equations (2) and (3), m - 1 of the sets J_k, $k \in M$ will have exactly one basic variable. The remaining J_k set will have one basic variable if there is a basic variable in N-J, and otherwise will have two basic variables. (By convention we refer to a variable as "in" a set if its subscript is in the set.)

To facilitate the subsequent development, we will introduce notational conventions that will be useful for depicting the form of a typical basic solution within the framework of the dual simplex method. Throughout this paper we will let J_q denote the exceptional set that has two basic variables, when this situation applies, and in general, let x_{k*} denote the basic variable (or one of the basic variables) in set J_k, $k \in M$. We will suppose that k* is unique for each set J_k, and call x_{k*} the starred basic variable for J_k. In the case of J_q, we will denote the basic variable other than x_{q*} by $x_{q'}$. As will be seen, this convention will allow us to associate different formulas with x_{q*} and $x_{q'}$, though of course these formulas yield equivalent expressions when q* and q' are interchanged. Additionally when there exists a basic variable in N-J it is denoted by x_p. Finally, we introduce the objective function variable $x_o = -\sum_{j\in N} c_j x_j$ whose maximization achieves the minimization of (1), and let NB denote the index set of current nonbasic variables.(4)

Basic solution forms

Case 1. x_p is basic in N - J.

$$x_o + \sum_{j\in NB} u_j x_j = u_o \tag{5}$$

$$x_p + \sum_{j\in NB} v_j x_j = v_o \tag{6}$$

$$x_{k*} + \sum_{j\in NB\cap J_k} x_j = 1 \qquad k \in M \tag{7}$$

(Note, $NB\cap J_k = J_k - \{k*\}$.)

Case 2. No variables are basic in N - J; $x_{q'}$ and x_{q*} are basic in J_q.

$$x_o + \sum_{j\in NB} u_j x_j = u_o \tag{8}$$

$$x_{q'} + \sum_{j\in NB} v_j x_j = v_o \tag{9}$$

$$x_{q*} + \sum_{j \in NB \cap J_q} (1-v_j)x_j$$

$$+ \sum_{j \in NB - J_q} (-v_j)\ x_j = 1 - v_o \qquad (10)$$

$$x_{k*} + \sum_{j \in NB \cap J_k} x_j = 1 \qquad k \in M - \{q\} \qquad (11)$$

For our subsequent development, we need to identify the precise connections between the coefficients of the basis representations in Case 1 and Case 2 and the coefficients of the original problem representation (1)-(4).

To reduce all formulas to the same notation for both Case 1 and Case 2 when $N \neq J$, we define $J_{m+1} = (N-J)U\{n+1\}$, where x_{n+1} is a "fictitious" variable, unrestricted in sign, with $a_{n+1} = c_{n+1} = 0$. We further specify that x_{n+1} is always the starred basic variable for the set J_{m+1}, i.e., $n+1 = (m+1)*$. Although we are completely unconcerned about the value of x_{n+1}, we may view x_{n+1} as difinitionally equal to $1 - \sum_{j \in N-J} x_j$ and indeed x_{n+1} will receive this value by the prescriptions we will specify. Upon defining $M_o = MU\{m+1\}$ when $N \neq J$ and $M_o = M$ otherwise, the GUB equations of (3) therefore hold with M replaced by M_o. (That is, the existence of x_{n+1} would make case 1^oequivalent to Case 2 except for the fact that x_{n+1} is unrestricted.) In particular, then, the preceding equations for the Case 2 basic solution may be regarded as also applicable to Case 1, for $q = m + 1$, $q' = p$ and $q* = n + 1$, enabling subsequent formulas to be simplified. However, we will on occasion find it useful to discuss Case 1 and Case 2 on separate terms (when the unrestricted value of x_{n+1} has special implications).

By these conventions, the connections between current basis coefficients and the original problem coefficients are expressed in the following remark.

Remark 2. Let $\alpha = a_o - \sum_{k \in M_o} a_{k*}$, $\delta = a_{q'} - a_{q*}$

and $d_j = c_j - v_j(c_{q'} - c_{q*})$

then the coefficients of (8) - (11) (with M replaced by M_o) may be expressed in terms of those of (1) - (4) by:

$$v_o = \alpha/\delta$$

$$v_j = (a_j - a_{k*})/\delta \text{ for } j \in J_k,\ k \in M_o$$

$$u_j = d_j - c_{k*} \text{ for } j \in J_k,\ k \in M_o$$

The derivation of the remark is immediate by the application of Gaussian elimination. It may be noted, incidentally, that the arbitrary designation of $x_{q'}$ and x_{q*} implies that the coefficients of equation (10) can alternately be obtained from the expression for the v_j coefficients in Remark 2 by interchanging q' and $q*$ in this expression.

3. PROPERTIES OF BASIC DUAL FEASIBLE SOLUTIONS

The goal of this section is to identify special properties of basic dual feasible solutions to (1)-(4), as a foundation for initialing a dual method. The following theorem (which slightly generalizes results of [15,16]) accomplishes this by providing necessary and sufficient conditions for a basis to be dual feasible--i.e., to yield $u_j \geq 0$, $j \in NB$, in the expression for x_o in (5) and (8). For this result we keep Case 1 and Case 2 separate.

Theorem 1. A basic solution is dual feasible for (1)-(4) if and only if:

$$k* \text{ is selected so that } d_{k*} = \underset{j \in J_k}{\text{Minimum}} \{d_j\}$$
$$k \in M - \{w\},$$

(where d_j is as defined in Remark 2 and in Case 2, $\{w\} = \{q\}$ and $\{w\} = \emptyset$, otherwise).

Case 1. $c_h/a_n \leq c_i/a_i \qquad h \in H,\ i \in I$

where $H = \{h \in N - J: a_h < 0\}$, $I = \{i \in N - J: a_i > 0\}$ and p is selected to be an $h \in H$ that yields the maximum c_h/a_n or to be an $i \in I$ that yields the minimum c_i/a_i.

Case 2. q' is selected so that $c_{q'} \leq c_j$ for all $j \in J_q$ such that $a_j = a_{q'}$, and

$$\left.\begin{array}{ll} (c_{q'}-c_r)/(a_{q'}-a_r), & r \in R \\ c_h/a_h; & \in H \end{array}\right\} \leq \left\{\begin{array}{ll} (c_{q'}-c_s)/(a_{q'}-a_s), & s \in S \\ c_i/a_i, & i \in I \end{array}\right.$$

where $R = \{r \in J_q: a_r < a_{q'}\}$, $S = \{s \in J_q: a_s > a_{q'}\}$.

Then $q*$ is selected to be an $r \in R$ that yields the maximum value of all terms on the right of the inequality, provided this is possible in consideration of the terms c_h/s_h and c_i/a_i. (Otherwise, the choice of q' does not allow dual feasibility. Also, whenever H or I is empty, the inequalities involving the corresponding $h \in H$ or $i \in I$ are not applicable.)

Proof. The stipulations about $k*$ and Case 1 are immediate from Remark 2, noting that $c_{q*} = a_{q*} = 0$ for Case 1. The stipulations about Case 2 are derived as follows. When $x_{q'}$ and x_{q*} are both basic, then there are dual multipliers Θ for equation (2) and π for the J_q equation of (3) such that $u_j = c_j - (\Theta a_j + \pi)$ for $j \in J_q$. These multipliers must be selected to yield $u_{q'} = u_{q*} = 0$. From $u_{q'} = 0$ we obtain $\pi = c_{q'} - \Theta a_{q'}$, and hence $u_j = c_j - c_{q'} + \Theta(a_{q'} - a_j)$. The dual feasibility requirement $u_j \geq 0$ yields

$$\Theta(a_{q'} - a_j) \geq c_{q'} - c_j$$

Thus, if $a_j = a_{q'}$, then $c_{q'} \leq c_j$, as first stipulated under Case 2. The alternatives $a_j < a_q$, and $a_j > a_{q'}$, identified by $j \in R$ and $j \in S$, respectively, yield

$$(c_q-c_r)/(a_{q'}-a_r) \leq \Theta \leq (c_{q'}-c_s)/(a_{q'}-a_s) \quad r \in R,\ s \in S$$

Dual feasibility requirements $u_j = c_j - \Theta a_j \geq 0$ for $j \in N - J$ further yield

$$c_h/a_n \leq \Theta \leq c_i/a_i \qquad h \in H,\ i \in I$$

leading to the full set of inequalities stipulated for Case 2. Finally, the condition $u_{q*} = 0$ requires that $u_{q*} = 0$ requires that $q*$ be selected so that $\Theta = (c_{q'} - c_{q*})/(a_{q'} - a_{q*})$. This completes the proof.

Theorem 1 discloses what may also be argued by simple dominance considerations--first, that we may throw out all elements of a set J_k with tied a_j values except for one with the smallest c_j value, and second, that all elements of H and I may be discarded except those yielding the maximum c_h/a_h and the minimum c_i/a_i. Thus $N - J$ can be restricted to at most two elements. If both these elements exist, and $c_h/a_h > c_i/a_i$, then the problem has an unbounded optimum. Otherwise, Case 1 of Theorem 1 provides an immediate starting duel feasible basic solution whenever $N - J$ is nonempty, by selecting either x_h or x_i as a basic variable (according to which of these variables exist). This observation also applies when $N = J$, because it is possible to add an artificial variable x_n (for n increased by 1), yielding $N - J = \{n\}$, with $a_n = 1$ and c_n large. (This variable is not to be confused with the "fictitious" x_{n+1}.)

However, Theorem 1 also makes it possible to obtain starting dual feasible solutions without resorting to the elementary Case 1 situation. The following corollary indicates an easy way to do this when $N = J$ and $N - J = \{n\}$. We assume for this setting that $a_n=1$ for $N - j = \{n\}$. In addition, we will suppose $c_n = 0$ for $N - J = \{n\}$, using Gaussian elimination on the objective function to achieve this if necessary.

Corollary 1. When $N = J$ or $N - J = \{n\}$, a Case 2 starting basic dual feasible solution can be obtained by designating any J_k to be J_q, selecting q' so that

$$a_{q'} = \underset{j \in J_q}{\text{Minimum}} \{a_j\}, \qquad c_{q'} = \underset{j \in J_q:\ a_j = a_q}{\text{Minimum}} \{c_j\}$$

and selecting $q^* \in S$ so that

$$(c_{q^*} - c_{q'})/(a_{q^*} - a_{q'}) = \underset{s \in S}{\text{Minimum}}\ (c_s - c_{q'})/(a_s - a_{q'})$$

If $S = \emptyset$, then $x_{q'} = 1$ (and the problem shrinks). If $S \neq \emptyset$, but $N - J = \{n\}$ (with $a_n = 1$ and $c_n = 0$), then $c_{q^*} < c_{q'}$ or else, again $x_{q'} = 1$. (For this case $c_j \geq c_{q'}$ for $j \in J_q$ allows $x_j = 0$.)

When $N = J$ in Corollary 1, replacing (2) by its negative leads to an alternative application of the corollary, equivalent to picking $a_{q'}$ to be a maximum and selecting $q^* \in R$ to yield a maximum ratio.

We now turn to the main results of this paper, characterizing the relationships of the dual simplex method applied to (1) - (4), and developing an efficient specialization for this problem. As a by-product we will also identify ways to generate other starting basic solutions that accord with the conditions of Theorem 1.

4. SPECIALIZATION OF THE DUAL SIMPLEX METHOD

For convenience in the following development, we outline the steps of the dual simplex method as follows.

Step 0. Begin with a dual feasible basis.

Step 1. Select any equation, other than the x_o equation, with a negative constant term. (If none exists, the current basic solution is optimal.) Represent this equation in the form of (9) (thereby identifying the outgoing variable as $x_{q'}$):

$$x_{q'} + \sum_{j \in NB} v_j x_j = v_o \qquad (v_o < 0)$$

Step 2. Let $NB^- = \{j \in NB: v_j < 0\}$. If NB^- is empty, the problem has no feasible solution. Otherwise, select the incoming variable x_i, $i \in NB^-$ to yield

$$u_i/v_i = \underset{j \in NB^-}{\text{Maximum}} \{u_j/v_j\}$$

where the u_j coefficients are those of the current x_o equation (8).

Step 3. Execute a basis exchange (pivot) step that replaces $x_{q'}$ by x_i in the basis. The updated form of the pivot equation (9), which becomes the new x_i equation, is

$$x_i + \sum_{j \in NB^*} (v_j/v_i) x_j = v_o/v_i$$

where NB* is the new set of nonbasic variables (replacing i by q') and $v_{q'} = 1$ (as implicit in (9)). The updated form of all remaining equations is obtained by Gaussian elimination (or equivalently, direct substitution) using the x_i equation to remove x_i from the other equations. Then return to Step 1.

The foregoing description of the dual method is entirely general and not specific to the LP/GUB knapsack problem except for the notation linking the current pivot equation to (9) and the x_o equation to (8). By means of this notational link, however, we may now make additional observations concerning the solution path of the dual simplex method for this problem.

Note first of all that the convention of representing the pivot equation in the form of (9) is entirely permissible in the restricted setting of the LP/GUB knapsack problem since we may always interchange the roles of $x_{q'}$ and x_{q^*} as necessary to allow this representation. Clearly, too, at most one of the two equations (9) and (10) can have a negative constant term and thereby qualify as the pivot equation. Thus, representing the pivot equation in the form of (9) serves to uniquely identify the indexes q' and q*. In fact, using the connections of Remark 2, we may immediately express the conditions for identifying $v_j < 0$ and the maximum ratio of Step 2 of the dual method in terms of the original problem coefficients.

Remark 3. If $a_{q'} > a_{q^*}$, then

$$v_j < 0 \text{ if and only if } a_j < a_{k^*} \qquad (12)$$

and if in addition $v_j \neq 0$, $v_h \neq 0$ for $j \in J_r$, $h \in J_u$ (possibly r = u), then

$$u_j/v_j \leq u_h/v_h \text{ if and only if } \Theta_{jr^*} \leq \Theta_{hu^*} \qquad (13)$$

where $\Theta_{fg} = (c_f - c_g)/(a_f - a_g)$. If $a_{q'} < a_{q^*}$, then the direction of the second inequality in (12) and in (13) is reversed.

Although this remark follows directly by substituting the coefficient identities of Remark 2 into Remark 3, its implications are quite useful. This is due to the somewhat surprising fact that the application of the dual simplex method assures that if $a_{q'} > a_{q^*}$ holds at one iteration, then $a_{q'} > a_{q^*}$ (for other indexes q' and q*) at all iterations. This relationship and others associated with it are expressed in the following main result of this section.

Theorem 2. Let J_t denote the set containing the incoming variable x_i determined in Step 2 of the dual simplex method. If $t = m + 1$ (i.e., if $i \in N-J$), then the pivot must yield an optimal solution. If $t \leq m$, and if the pivot does not yield an optimal solution, then upon representing the next pivot equation also as

(9), all of the following hold:

(a) J_t becomes the new J_q

(b) x_{t*} becomes the new outgoing variable $x_{q'}$

(c) x_i becomes the new x_{q*}

(d) the ratio values Θ_{jk*}, $j \in J_k$, remain unchanged for all $k \in M_o - \{t\}$

(e) $a_{q'} > a_{q*}$ before the pivot if and only if $a_{q'} > a_{q*}$ (for the new q' and $q*$) after the pivot.

(f) Over a series of pivots, as the index k is periodically selected as t, the elements a_{k*} will only change in descending sequence if $a_{q'} > a_{q*}$ and will only change in ascending sequence if $a_{q'} < a_{q*}$.

Proof. Each of the assertions is a direct outcome of applying the dual simplex method. First, the x_i equation of Step 3 of the dual method must have a positive constant term (since both v_o and v_i are negative), and cannot qualify as the new pivot equation. However, this equation currently has the form of (9) (since x_{t*} and not x_i, is the current starred basic variable for the set J_t). Thus, equation (10) is the only possibility for the new pivot equation, in which case it may be put in the form of (9) by interchanging the roles of i and t*. The interchange of i and t* is unnecessary if $i \in N-J$ because x_{t*} is the unrestricted variable x_{n+1}, and an optimal solution is already obtained. Otherwise, if the current solution is not feasible (the solution value of x_i exceeds 1), the interchange immediately establishes (a), (b) and (c) of the theorem. Next, since J_t is the only set J_k in which the identity of x_{k*} changes by the pivot, it also follows that the values Θ_{jk*} change only for k=t, establishing (d). The condition $a_{q'} > a_{q*}$ before the pivot is equivalent to stipulating $a_i < a_{t*}$ in consideration of the fact that $v_i < 0$ (Remark 3). But since t* becomes the new q' and i becomes the new q*, this yields (e). Finally, (f) follows directly from (e) and Remark 3, completing the proof.

We will henceforth suppose for simplicity that $a_{q'} > a_{q*}$ on all iterations, understanding that the directions of inequalities specified in the following discussion may have to be reversed if this is not the case. (Alternatively, it is always possible to assure $a_{q'} > a_{q*}$ by the device of replacing equation (2) by its negative in case $a_{q'} < a_{q*}$.) With this understanding, Theorem 2 directly implies

Corollary 2. (For $a_{q'} > a_{q*}$): If the maximum ratio R_k, given by

$$R_k = \underset{\substack{j \in J_k \\ a_j < a_{k*}}}{\text{Maximum}} \{\Theta_{jk*}\} \tag{14}$$

is known for each set J_k, $k \in M_o$, together with the index i(k) such that $R_k = \Theta_{jk*}$ for $j = i(k)$, then the incoming variable x_i is identified by

$$i = i(t) \text{ where } R_t = \underset{k \in M_o}{\text{Maximum}} \{R_k\} \tag{15}$$

and the pivot step leaves all R_k except R_t unchanged for the determination of the new R_t by (15) at the next pivot. (If $a_{q'} < a_{q*}$, the maximum in (14) is replaced by a minimum over $a_j > a_{k*}$.)

The significance of Corollary 2 is twofold. First of all, it allows the dual simplex method to be implemented for the LP/GUB knapsack problem without ever explicitly calculating the u_j and v_j coefficients. Secondly, it allows the R_k values to be efficiently stored in a heap, with the maximum R_t at the top. Then as R_t is removed, and replaced with a new value, the unchanged values of the remaining R_k enable the heap to be reconstituted with minimal computation (on the order of O(log m)).

The issue remaining before giving a detailed specification of the steps of a specialized dual algorithm, is the efficient determination of R_k by (14). Since each time a new R_k is found, the variable $x_{i(k)}$ will become the new x_{k*} (the next time k is selected as t by (15)), all of the $j \in J_k$ such that $a_j \geq a_{i(k)}$ may immediately be dropped, since they will be of no further interest. This approach by itself, as will be shown, leads to a specialized method whose worst case computational bound is superior to that of [16] when the number of GUB sets exceeds the number of elements in each set. (This generally occurs in practical applications of an "assignment" nature, where the number of items to be assigned generally far exceeds the number of possible assignments per item.) However, an even better approach from the standpoint of worst case bounds results by a simple preliminary pass through each set J_k, eliminating in advance the elements that do not qualify to be selected as k*. Since the elements that are left will be visited in descending order of the a_j values (for $a_{q'} > a_{q*}$), it follows that each successively smaller a_j will be the a_{k*}, and the task of identifying a maximum by (14) is eliminated.

Specifically, then, we seek to identify a subset J_k^o of J_k whose elements are linked by a predecessor/successor ordering, where the immediate successor s(j) of an index $j \in J_k^o$ identifies the next element that qualifies to serve as k* after j, and the immediate predecessor p(j) of j identifies the element of J_k^o that qualifies to serve as k* immediately before j. Initially, of course, s(j) and p(j) just arrange the elements of J_k in descending (ascending) order and we will suppose that in the process of creating such a linking that duplicate a_j values are removed by retaining only the one associated with the smallest c_j value. The process of dropping an element from J_k in the construction of J_k^o can be accomplished simply by linking its immediate predecessor to its immediate successor.

Under this predecessor/successor linking, (14) can be written

$$R_k = \Theta_{ik*} \geq \Theta_{jk*} \text{ for all successors } j \text{ of } i = i(k)$$

Then i will become the new k* (except for the first i selected as k*). Thus, in particular, since we may eliminate the situation of tied maximum ratios by selecting the one with the smallest a_i coefficient (which has no tied successors), and since dropping superfluous elements will yield $k* = p(i)$, the identifying characteristic of J_k^o becomes

$$\Theta_{ip}(i) > \Theta_{jp}(i) \tag{16}$$

for all successors j of i and for all $i \in \overline{J}_k^o$, where $\overline{J}_k^o$ is j_k^o stripped of its first and last elements, which respectively have no predecessors or successors. The task of weeding out elements of J_k to assure this relationship is made easy by the following.

Remark 4. The inequality (16) holds for all $i \in \overline{J}_k^o$ and for all successors j of i if and only if it holds for all $i \in \overline{J}_k^o$ and for $j = s(i)$.

Proof. We need only show that for any h, i, j, r (taking the roles $h = p(i)$, $j = s(i)$ and $r = s(j)$) such that $a_h > a_i > a_j > a_r$, the two "successive" inequalities $\Theta_{ih} > \Theta_{jh}$ and $\Theta_{ji} > \Theta_{rj}$ imply $\Theta_{ih} > \Theta_{rh}$. First, for the coefficients as ordered, we note that $\Theta_{ih} > \Theta_{jh}$ is equivalent to $\Theta_{jh} > \Theta_{ji}$, since both of these inequalities reduce to $c_h a_i + c_i a_j + c_j a_h > c_i a_h + c_j a_i + c_h a_j$. Similarly, $\Theta_{ji} > \Theta_{rj}$ is equivalent to $\Theta_{rj} > \Theta_{ri}$. Hence we obtain $\Theta_{ih} > \Theta_{jh} > \Theta_{ji} > \Theta_{rj} > \Theta_{ri}$ and in particular $\Theta_{jh} > \Theta_{rj}$, which is equivalent to $\Theta_{ri} > \Theta_{rh}$. Consequently, $\Theta_{ih} > \Theta_{rh}$, completing the proof.

To make convenient use of this observation we introduce a dummy index 0 to "start" and "terminate" the predecessor/successor linking, where 0 is treated as immediate predecessor of the largest a_j and the immediate successor of the smallest a_j. The procedure for modifying the initial linking on J_k so that it becomes a linking on J_k^o is then as follows.

0. To start, let h, i and j be the "first three" elements of J_k, that is, $h = s(0)$, $i = s(h)$, $j = s(i)$. (If J_k has less than three elements, then $J_k^o = J_k$ and nothing is to be done.)

1. Compare Θ_{ih} to Θ_{jh}.
   (a) If $\Theta_{ih} > \Theta_{jh}$, set $h = i$ and go to Step 2.
   (b) If $\Theta_{ij} = \Theta_{jh}$ or if $\Theta_{ih} < \Theta_{jh}$ and $p(h) = 0$, drop i and go to Step 2.
   (c) If $\Theta_{ih} < \Theta_{jh}$ and $p(h) \neq 0$, drop i, set $i = h$, and $h = p(i)$. Then return to the start of Step 1.
2. Set $i = j$ and $j = s(i)$. If $j = 0$, the procedure stops and the linking correctly identifies the ordered elements of J_k^o. Otherwise, return to Step 1.

The validity of the foregoing procedure is an immediate consequence of Remark 4. Note that the index j never "backs up" to a predecessor value, but remains unchanged in Step 1 and set to its successor at Step 2. Consequently Step 2 will always be executed $n_k - 2$ times, where n_k is the number of elements in J_k. Whenever the method does not go to Step 2, the index i is dropped at 1(c), which can occur at most $n_k - 2$ times (since i is never the first or last element), for a total number of iterations of the procedure equalling at most $2(n_k - 2)$. This procedure is patterned after one due to Witzgall [16] (who obtains a different iteration count) except that Witzgall's approach is based upon a geometric determination of the locations of points on or below line segments, rather than on a direct comparison of ratios as afforded by Remark 4.

It should also be noted, in contrast to the less general situation examined in [16], that the elements of J_k^o may not all qualify to be basic in a dual feasible solution. If $N \neq J$, it is additionally necessary that the ratios $\Theta_{ip(i)}$ be bounded by the limiting ratios from N - J, as shown in Theorem 1. This means that some of the initial and final elements of J_k^o (under the predecessor/successor linking) may also drop out of consideration. Rather than bothering to check for this situation in advance, however, the first and last relevant elements of J_k^o can be determined automatically by starting from some initial basic dual feasible solution and simply executing the specialized dual algorithm.

In general, these observations lead to the following Corollary as an extension of the options available from Corollary 1 for obtaining an initial dual feasible basis.

Corollary 3. The set of Case 2 dual feasible bases, any one of which provides an acceptable starting basis for the specialized dual simplex method, can be generated by selecting an arbitrary J_k^o to be J_q^o, and selecting any element i from this set (other than the first element) such that $\Theta_{ip(i)}$ satisfies the limiting bounds from N - J (identified in Theorem 1). Then i and p(i) may respectively serve as q' and q*. If no such element i exists, then some other set must serve as J_q^o, and whatever element of the "unacceptable" J_k^o thereby enters the basis in the starting solution is compelled to be basic in all dual feasible bases (hence, the associated variable may be fixed at the value 1).

The elements q' and q* found in Corollary 3 may need to be interchanged, so that the first pivot equation can be represented by (9). (In this case, the $a_{q'} > a_{q*}$ assumption must be replaced by the $a_{q*} > a_{q'}$ assumption, reversing the roles of the predecessor/successor links.) If a Case 1 basis is used as the start, then $a_p > 0$ (for x_p the basic variable in (5)) implies $a_{q'} > a_{q*}$ on all iterations (since p takes the initial role of q' with $a_{q*} = a_{n+1} = 0$), whereas an artificial start (with $p = n$, $a_n = 1$ and c_n large) will select the first ("largest a_j") element of each J_k as the initial a_{k*}.

The specialized dual simplex method based on the foregoing results may now be described as follows.

The Specialized Dual Simplex Method

1. Initialization.
   (a) Create the predecessor/successor linkings and the J_k^o sets, $k \in M_o$. (For $N \neq J$, define J_{m+1}^o to be the set containing the elements (at most two in number) with limiting ratios identified by Theorem 1.) (This step can be deferred or applied in conjunction with Step 1(b), using the starting basis there to reduce the range of elements considered for inclusion in the J_k^o sets.)
   (b) Create a starting dual feasible basis (as by Theorem 1 and Corollary 1 or Corollary 3). Compute the initial v_o value by computing

$$\alpha = a_o - \sum_{k \in M} a_{k*} \text{ and } v_o = \alpha/(a_{q'} - a_{q*})$$

If $v_o \geq 0$ and either $q' \in N - J$ or $v_o \leq 1$, then the current basic solution ($x_{q'} = v_o$, $x_{q*} = 1 - v_o$ and $x_{k*} = 1$, $k \in M_o - \{q\}$) is optimal. Otherwise, interchange q' and q* if necessary so that $v_o < 0$. For what follows we suppose $a_{q'} > a_{q*}$. (If not, the word "maximum" should be replaced by "minimum," and the successor symbol s() should be replaced by the predecessor symbol p().)

(c) Identify the ratios $R_k = \Theta_{s(k*)k*}$ for each $k \in M_o$. (If $s(k*) = 0$, the ratio R_k does not exist, and is bypassed. For the case $k = m + 1$ where by convention $k* = n + 1$, we define $s(k*)$ to be the first element of j^o_{m+1} excluding the current q' (if $q' \in J_{m+1}$). Hence $R_{m+1} = c_j/a_j$ for $j = s(k*)$, if this element j exists.) Put these ratios in a heap, with the maximum at the top.

2. Identify the incoming basic variable and the new basis composition.

Pick the maximum ratio from the top of the heap and denote it R_t. (If the heap is empty, there is no feasible solution.) The current variable $x_{q'}$ leaves the basis and $x_{s(t*)}$ enters the basis. If $s(t*) \in N - J$, the current basic solution is optimal for $q' = s(t*)$ and $v_o = \alpha/(-a_{q'})$ (where α is unchanged from its previous value). Otherwise, the current x_{t*} becomes the new $x_{q'}$, while $x_{s(t*)}$ is the new x_{q*}; i.e., set $q' = t*$ and $q* = t* = s(t*)$.

3. Update the current basic solution.

Update α and v_o by setting $\delta = a_{q'} - a_{q*}$, $\alpha = \alpha + \delta$ and $v_o = \alpha/\delta$. If $v_o \geq 0$, the current basic solution is optimal. Otherwise, identify the new value of $R_t = \Theta_{s(t*)t*}$ (for the new t*). If the ratio does not exist ($s(t*) = 0$), reform the heap for the ratios still in it. Otherwise, add R_t back to the heap. Then return to Step 2.

An analysis of the maximum amount of computation required by this method is as follows. The creation of the predecessor/successor linkings (that initially arrange the a_j coefficients in descending/ascending order for each J_k) requires on the order of $O(n_k \log n_k)$ computation for each set, or an effort of at most $\sum_{k \in M_o} O(n_k \log n_k) \leq O(n \log n)$. (For the case where each GUB set has the same number of elements, n/m, we may refine this to $O(n(\log n - \log m))$.)

The work to modify the linking to identify the J^o_k set involves at most $2n_k - 4$ iterations of the procedure based on Remark 4, or $2n - 4m$ iterations over all sets, requiring computation or order $O(n - m) \leq O(n)$ (including the effort of generating and selecting the minimum d_j values) while computing the initial v_o value is $O(m)$. Finally, computing the R_k ratios requires $O(m)$ computation, while putting them in a heap is an effort of order $O(m \log m)$. Thus, the total initialization effort of Step 1 can be expressed as $O(n \log n) + O(m \log m)$.

For Steps 2 and 3, at most $n - m - 1$ elements (the successors of the k* elements, excluding the initial q') remain to be examined in the J^o_k sets, $k \in M_o$ and so these steps will require at most $n - m - 1$ iterations. Exclusive of reforming the heap, these two steps require a handful of "if checks," assignments, a couple of additions and 1 division. Reforming the heap requires an effort of $O(\log m)$, hence in total the amount of effort required at Steps 2 and 3 is $O((n - m)\log m)$. Putting these together with the effort required at initialization we can state

Theorem 3. The computational complexity of the LP/GUB knapsack problem is of order at most

$$O(n \log n) + O(n \log m)$$

or

$$O(n(\log n + \log m))$$

We have stated these order bounds separately instead of simply giving the $O(n(\log n + \log m))$ bound, because of the overestimate involved in the $O(n \log n)$ term. In particular, as previously noted, this term can instead be expressed as $O(n(\log n - \log m))$ for the situation in which each GUB set has n/m elements. Thus, in this case we have

Corollary 4. When each GUB set contains the same number of elements, the computational complexity of the LP/GUB knapsack problem is at most $O(n \log n)$.

The bound of Witzgall is given in [16] as $O(n \log n) + O((n - m)m)$, where the $O(n \log n)$ term is essentially the same as that of Theorem 3, and also can be replaced by $O(n(\log n - \log m))$ for GUB sets with n/m elements. The primary difference between the bound of [16] and that of Theorem 3 is therefore the contrast between $O((n - m)m)$ and $O(n \log m)$. For easier comparison, let $g = n/m$ (so that g is the number of elements in each GUB set if each set has the same cardinality). Then these terms can be respectively written $O((g - 1)m^2)$ and $O(g(m \log m))$. Since $g \geq 2$ in any meaningful problem (or else there are GUB sets with only 1 element), the latter term clearly represents a smaller order of effort than the former, particularly as m or g (hence n) becomes larger. This difference appears to stem from the fact that our procedure specializes the dual simplex method directly, whereas Witzgall's instead carries out preliminary "topological reductions" (corresponding to those obtained via Remark 4) but otherwise leaves the dual method primarily to its own devices (for the case J = N). (Sinha and Zoltner's procedure and Witzgall's procedure appear closely related in this respect.)

It is interesting to note the type of order bound that results for our method when the initialization effort of setting up the predecessor/successor links and adapting them to the J_k^o sets is not employed.

The modifications for this approach are as follows.

Alternative method. (Omitting the initial ordering of the a_j coefficients by the predecessor/successor links, and the creation of the J_k^o sets.)

1. Initialization.
   (a) Deleted
   (b) As in the previous method, except that Corollary 3 is not used as a strategy for creating an initial basis. In addition, drop the index k^* from each J_k.
   (c) Instead of setting $R_k = \Theta_{s(k^*)k^*}$, examine each $j \in J_k$ (for J_k as currently constituted). If $a_j \geq a_{k^*}$ then drop j from J_k, and if not compute the ratio Θ_{jk^*}, saving the minimum of these computed ratios as R_k. (Then $s(k^*)$ denotes the j that gives this minimum ratio.)
2. Identify the incoming basis variable and the new basis composition.

   As in the previous method.
3. Update the current basic solution.

   As in the previous method, except for setting $R_t = \Theta_{s(t^*)t^*}$. Instead, first drop t^* from J_t, and for each remaining $j \in J_t$ (as currently constituted) carry out the operation indicated in 1(c) for $k = t$.

The type of analyses applied to the computations for the previous method allows us to state

Corollary 5. When each GUB set has the same number of elements $g = m/n$, the computational effort required by the Alternative Method for the LP/GUB knapsack problem is of order

$$O(n) + O((n - m)\log m) + O((n - m)g)$$

or

$$O(n) + O((n - m)(g + \log m)$$

Again we have written the bound in different ways to facilitate comparison with the other bounds. The $O(n)$ term here is comparable to a $O(n)$ term that was previously assimilated into $O(n \log n)$ in both our approach and in Witzgall's. Thus, for a clearer comparison, the bound of Corollary 4 can be rewritten

$$O(n) + O(n) \log g) + O((n - m)m).$$

While the worst case bound of Corollary 4 appears generally superior to the other two, note that the bound for the Alternative Method appears more attractive than that of [16] for $g = m$, and becomes increasingly attractive as m becomes larger relative to g, due to the fact that increases in log m are dwarfed by increases in m. (The value of m is often several fold greater than g in practical applications. For example, in the applications of [8, 11, 12, 13], m ranges from 4g to 50g.) Coupling this with the fact that the Alternative Method requires less "set up" effort than the other methods makes it an appealing alternative for problems in which worst case bounds are expected to be overly pessimistic. In this context, any attempt to consider "likely" cases instead of worst cases must also account for the advantages that may derive from initiating a specialized dual algorithm from an advanced starting basis, rather from an "extreme end" of the dual feasible region (as in [15] and [16]).

Finally, it is interesting to consider the specialization of these bounds to the ordinary knapsack problem. In this problem, the number of variables before adding slacks to give GUB constraints is $m = n/2$ (i.e., the addition of slacks yields $g = 2$). Bounds of both previously indicated versions of the Specialized Dual Simplex Method (from Corollary 4 and Corollary 5) reduce to $O(m \log m)$ in this case, which is a standard bound for algorithms for the knapsack problem. Recently, however, Balas and Zemel [1] have developed an improved bound of $O(m)$ for m variable knapsacks. This raises the interesting question of whether it is possible to find a method for the general LP/GUB knapsack problem whose worst case computational effort specializes to $O(m)$, yet that maintains advantages for the general case. We conjecture that this is not possible.

REFERENCES

1. Balas, E. and Zemel, E. "Solving Large 0-1 Knapsack Problems," *Management Sciences Report No. 408*, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, June 1976, Revised July 1977.
2. Barr, Richard S. and Ross, G. T. "A Linked List Data Structure for a Binary Knapsack Algorithm," Research Report CCS 232, Center for Cybernetic Studies, The University of Texas at Austin, August 1975.
3. Beale, E. M. L. and Tomlin, J.A., "Special Facilities in a General Mathematical Programming System for Non-Convex Problems Using Special Ordered Sets of Variables," in *Proceedings of the 5th IFORS Conference*, London, 1970, 447-454.
4. Beale, E. M. L. and Tomlin, J. A., "An Integer Programming Approach to a Class of Combinatorial Problems," *Mathematical Programming*, 3, 3, (1972), 339-344.

5. Fayard, D. and Plateau, G., "Resolution of the 0-1 Knapsack Problem: Comparison of Methods," *Mathematical Programming*, (1975), 272-307.

6. Glover, F., "Convexity Cuts for Multiple Choice Problems," *Discrete Mathematics*, 6 (1973), 221-234.

7. Glover, F., "Surrogate Constraint Duality in Mathematical Programming," *Operations Research*, 23 (1975), 434-451.

8. Glover, F. and Klingman, D., "Network Applications in Industry and Government," *AIIE Transactions*, December 1977.

9. Greenberg, H. and Hegerick, R. L., "A Branch Search Algorithm for the Knapsack Problem," *Management Science*, 16 (1970), 327-332.

10. Greenberg, H. and Pierskalla, W. P., "Surrogate Mathematical Programming," *Operations Research*, 18 (1970), 924-939.

11. Gross, D., and Pinkus, C. E., "Optimal Allocation of Ships to Yards for Regular Overhauls," *Technical Report 63095*, George Washington University, Washington, D.C., 1972.

12. Klingman, D., Ross, G. T., and Soland, R. M., "Optimal Lot-Sizing and Machine Loading for Multiple Products," Proceeding of Conference on the Disaggregation Problem, Ohio State University, Columbus, Ohio, March 1977.

13. Masud, Enver, "An Interactive Procedure for Sizing and Timing Distribution Substations Using Optimization Techniques," *Paper T75 142-6*, Presented at the IEEE PES Winter Meeting, New York, January 27-February 1, 1974.

14. Nauss, Robert M., "The 0-1 Knapsack Problem with Multiple Choice Constraints," Working Paper, University of Missouri, St. Louis, March 1975, revised May 1976.

15. Sinha, P. and Zoltners, A. A., "The Multiple Choice Knapsack Problem," *Research Report No. 8-76*, School of Business Administration, University of Massachusetts, September 1976.

16. Witzgall, Christoph, "On One-Row Linear Programs," Applied Mathematics Division National Bureau of Standards, Washington, D.C., May 1973.

APPENDIX D
ANALYSIS OF DISTRIBUTION FEEDER COSTS

D.1. Introduction

In general, the distribution feeder costs can be classified as: (1) the cost of investment, (2) the cost of lost energy due to I^2R losses in the feeder, and (3) the cost of demand lost (i.e., the cost of lost capacity) due to the I^2R losses. The cost of investment is the largest cost component and it includes material costs involving feeders. The total cost of a given distribution feeder can be summarized by the following equation,

$$TAFC = AIC + ADC + AEC \qquad (D.1)$$

where

TAFC = total annual feeder cost in dollars per mile,
AIC = annual investment cost in dollars per mile,
ADC = annual demand cost due to lost capacity as a result of energy (or copper) losses in the feeder in dollars per mile,
AEC = annual energy cost due to I^2R losses in the feeder in dollars per mile.

D.2. Annual Investment Cost

The annual investment cost of a given feeder is the installed cost of the feeder times the fixed cost rate of the feeder. This fixed cost rate or so-called the carrying charge rate, of the feeder includes the cost of capital, taxes, insurance, operation and maintenance, and depreciation, etc. The annual investment costs need to be considered for rural and urban areas, separately. Tables D.1 and D.2 present some of the typical ACSR conductors used, at 12.5kV and 24.9kV voltage levels, in the rural areas, respectively. Tables D.3 and D.4 give some of the typical ACSR conductors used in the urban areas at 12.5kV and 34.5kV voltage levels, respectively.

D.3. Annual Energy Cost

The following formula has been used to calculate the annual energy cost due to I^2R losses in feeders,

$$AEC = 3I^2 \times R \times S_E \times F_{LL} \times F_{LS} \times F_{LSA} \times 8.76 \times 10^{-3} \qquad (D.2)$$

where

AEC = annual energy cost due to I^2R losses in feeders in dollars per mile,
I = current in Amperes,
R = resistance of a given conductor in ohms per conductor per mile,
S_E = energy cost in mills per kWh,
F_{LL} = load location factor,
F_{LS} = loss factor,
F_{LSA} = loss allowance factor.

The load location factor is a per unit value and is considered to be that point on a feeder having distributed loading where the total feeder load can be assumed to be concentrated for the purpose of I^2R loss calculations.

The loss factor is the ratio of the average power of the average power loss over a year's period to the peak loss occurring in that period. It can also be defined as the ratio of the actual total kWh losses to what the kWh losses would have been if the peak losses had continued throughout the 8760 hours in the year. The loss factor can be approximated by the following equation:

$$F_{LS} = 0.3F_{LD} + 0.7\,F_{LD}^2 \qquad (D.3)$$

where

F_{LD} = load factor

As losses are supplied to the primary distribution system through preceeding power system elements they incur additional losses. Therefore, not only the primary feeder losses must be supplied but also the additional losses incurred in the transmission and transformer systems must be supplied due to the flow through of primary feeder losses. The factor taking these losses into the consideration is called the loss allowance factor.

The annual energy cost due to I^2R losses in the feeders have been calculated for both copper and ACSR conductors at 4.16kV, 12.5kV, 24.9kV, and 34.5kV voltage levels. The developed nomographs are shown in Figures D.1 - D.20.

D.4. Annual Demand Cost

The following formula has been used to calculate the annual demand cost,

$$ADC = 3I^2 \times R \times F_{LL} \times F_{PR} \times F_{LSA} \times [S_p \times T_p + S_t \times T_t + S_{DS} \times T_{DS}] \qquad (D.4)$$

where

ADC = annual demand cost due to loss capacity as a result of energy losses in feeders in dollars per mile,
I = current in ampers,
R = resistance of a given conductor in ohms per conductor per mile,
F_{LL} = load location factor,
F_{PR} = peak responsibility factor,
F_R = reserve factor,
F_{LSA} = loss allowance factor,
S_p = production cost,
T_p = production fixed cost rate,
S_t = transmission cost,
T_t = transmission fixed cost rate,
S_{DS} = distribution substation cost,
T_{DS} = distribution substation fixed cost rate.

The reserve factor is the ratio of total generation capability to the total load and losses supplied. The peak responsibility factor is a per unit value of peak feeder losses which are coincident with the system demand. But not all distribution feeders have their peaks at the same time as the system peak. Since at times other than at the system peak, excess capacity is available, it is only valid to charge for the actual capacity necessary to serve the feeder losses which occur at the time of the system peak. The annual demand cost analysis included both copper and ACSR conductors at 4.16kV, 12.5kV, 24.9kV and 34.5kV voltage levels. The developed nomographs are shown in Figures D.21-D.40.

D.5. Total Annual Equivalent Cost

The total annual equivalent feeder cost in dollars per mile per MVA is calculated as,

$$TAEFC = AEIC + ADC + AEC \qquad (D.5)$$

where

TAEFC = total annual equivalent feeder cost in dollars per mile,
AEIC = annual equivalent investment cost in dollars per mile.

The AEIC can be calculated from

$$AEIC = (B-V)\,(a/p)_n^i + Vi \qquad (D.6)$$

where

B = installed first cost of the feeder in dollars per mile,
V = net salvage value of the feeder at the end of the nth year in dollars per mile,

or

Table D.1. Typical ACSR Conductors Used at 12.5kV in Rural Areas

Conductor size	Ground Wire size	Conductor weight	Ground Wire weight	$/lb	Installation & Hardware Cost	Total Installed Feeder Cost
#4	#4	356	356	0.6	6945.6	7200
1/0	#2	769	566	0.6	7176.2	8900
3/0	1/0	1223	769	0.6	7737.2	10400
4/0	1/0	1542	769	0.6	8563	11800
266.8MCM	1/0	1802	769	0.6	9985	13690
477 MCM	1/0	3462	769	0.6	10967	17660

Table D.2. Typical ACSR Conductors Used at 24.9kV in Rural Areas

Conductor size	Ground Wire size	Conductor weight	Ground Wire weight	$/lb	Installation & Hardware Cost	Total Installed Feeder Cost
#4	#4	356	356	0.6	7605.6	8460
1/0	#2	769	566	0.6	7856.2	9580
3/0	1/0	1223	769	0.6	8217.2	10880
4/0	1/0	1542	769	0.6	8293	11530
266.8MCM	1/0	1802	769	0.6	9615	13320
477 MCM	1/0	3462	769	0.6	11547	19240

Table D.3. Typical ACSR Conductors Used at 12.5kV in Urban Areas

Conductor size	Ground Wire size	Conductor Weight	Ground Wire Weight	$/lb	Installation & Hardware Cost	Total Installed Feeder Cost
#4	4	356	356	0.6	21,145.6	22000
1/0	4	769	356	0.6	22,402.2	24000
3/0	4	1223	356	0.6	24,535	27000
477MCM	1/0	3462	769	0.6	28,307	35000

Table D.4. Typical ACSR Conductors Used at 34.5kV in Urban Areas

Conductor size	Ground Wire size	Conductor Weight	Ground Wire Weight	$/lb	Installation & Hardware Cost	Total Installed Feeder Cost
4	4	356	356	0.6	21,375.6	22230
1/0	4	769	356	0.6	22,632.2	24230
3/0	4	1223	356	0.6	24,815	27230
477MCM	1/0	3462	769	0.6	28,537	35230

V = gross salvage value of the feeder at end of the nth year less cost of removal and restoration in dollars per mile,
n = useful life of the feeder in years,
i = carrying charge rate in percent,
$(a/p)^{i}_{n}$ = capital recovery factor for i percent carrying charge rate and feeder useful life of n year,

or

$(a/p)^{i}_{n}$ = uniform series worth of a percent sum for i percent carrying charge rate and feeder useful life of n year.

The capital recovery factor is calculated from

$$(a/p)^{i}_{n} = \frac{i(1+i)^{n}}{(1+i)^{n}-1} \qquad (D.7)$$

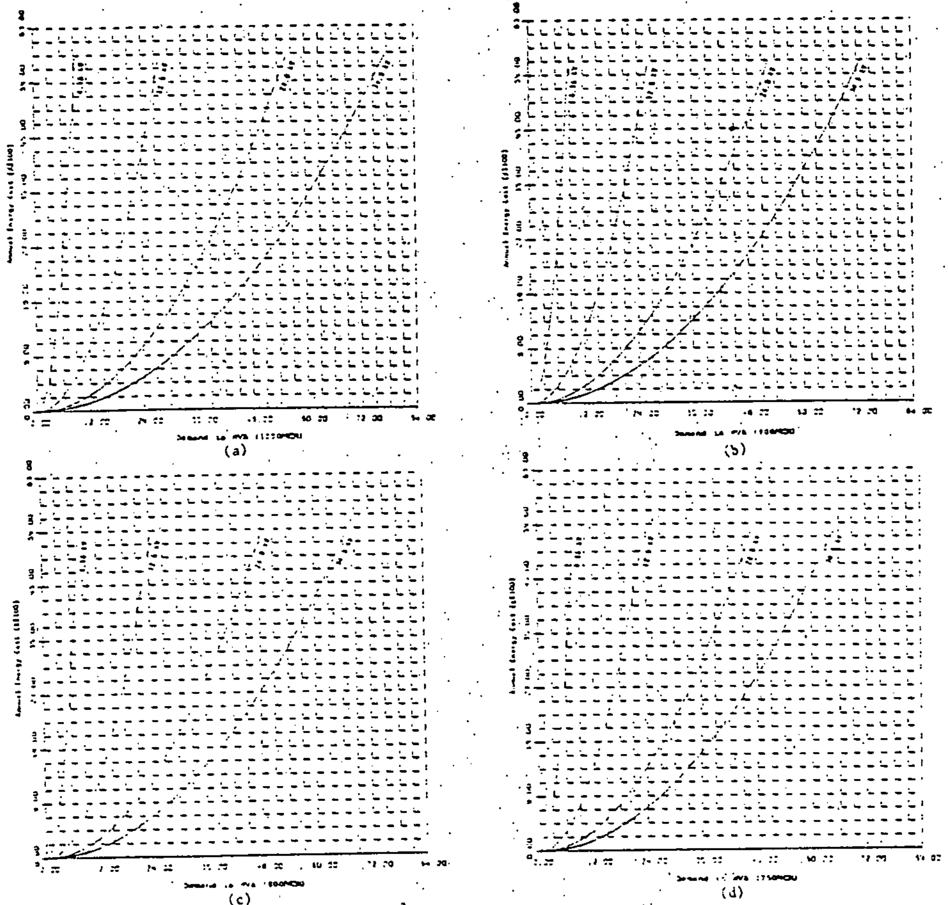
Therefore,

$$AEIC = (B-V)\frac{i(1+i)^{n}}{(1+i)^{n}-1} \qquad (D.8)$$

also

$$TAEFC = (B-V)\frac{i(1+i)^{n}}{(1+i)^{n}-1} + Vi + ADC + AEC \qquad (D.9)$$

Using a carrying charge rate of 12 percent and a useful feeder life of 30 years, the total annual equivalent feeder costs have been calculated for urban and rural areas. The developed nomographs are shown in Figures D.41- D.43.



(a) (b) (c) (d)

Figure D.1. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) 1000 CM, (b) 900 CM, (c) 800 CM, (d) 750 CM

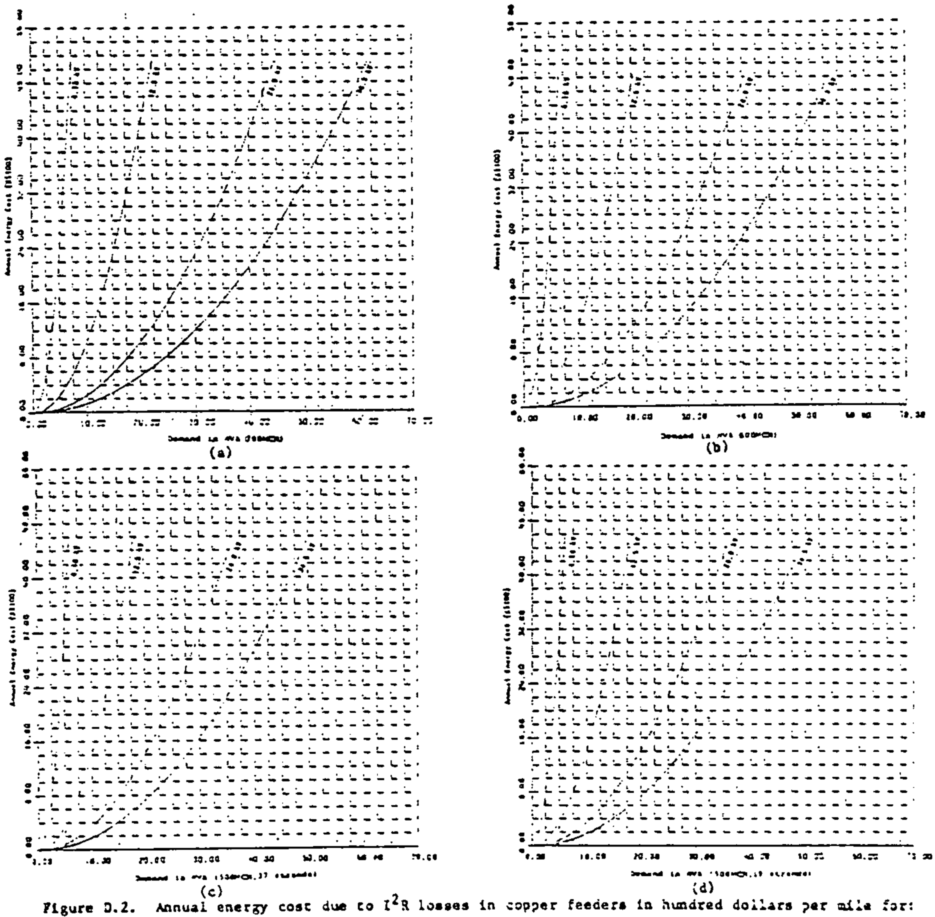


Figure D.2. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) 700 CM, (b) 600 CM, (c) 500 CM, 37 strands, (d) 500 CM, 19 strands

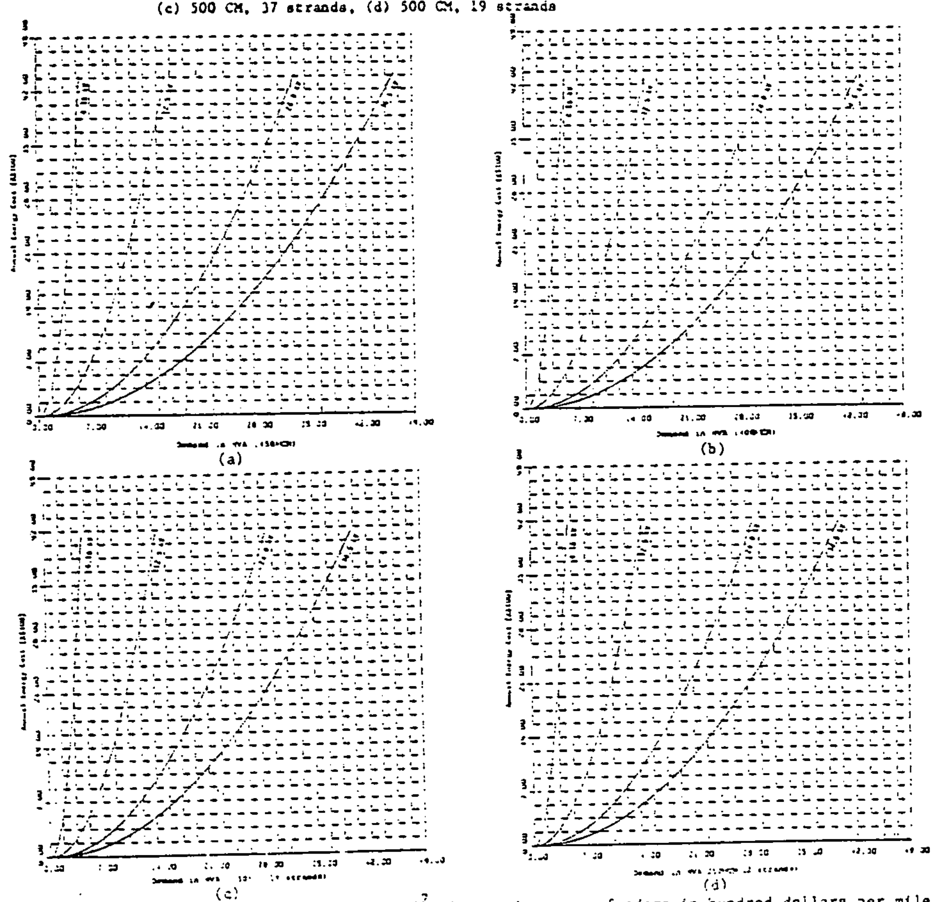


Figure D.3. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) 450 CM, (b) 400 CM, (c) 350 CM, 19 strands (d) 350 CM, 12 strands

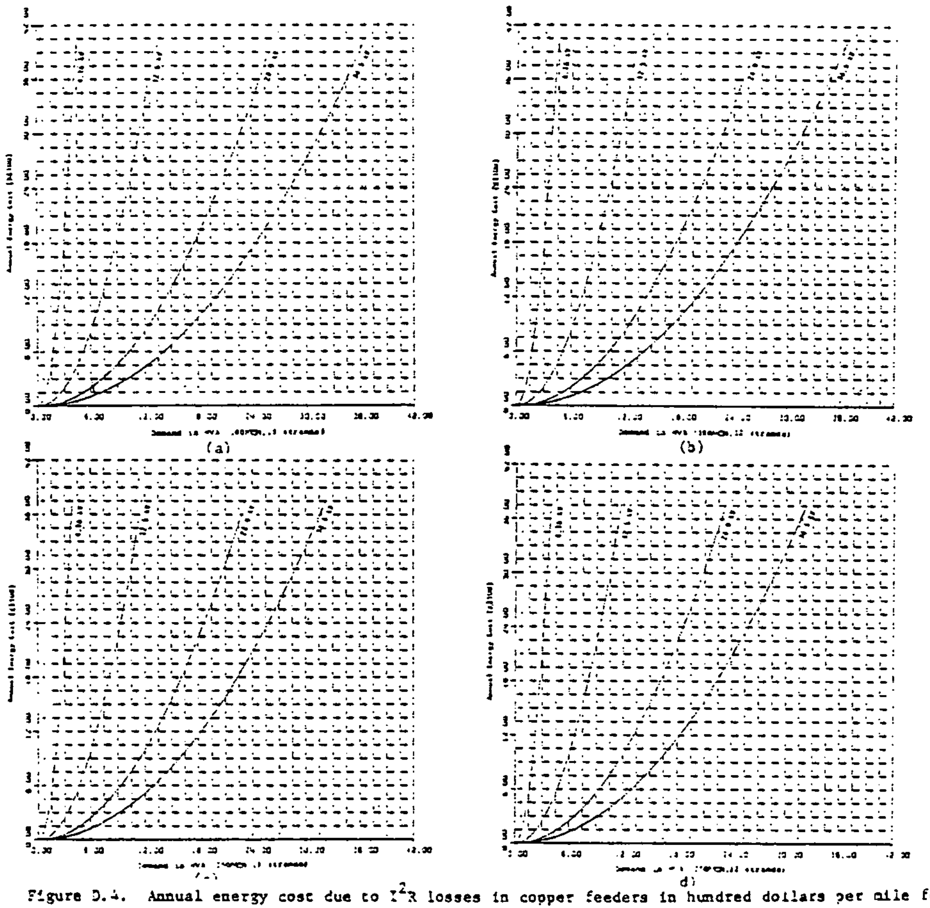


Figure D.4. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for:
(a) 300 CM, 19 strands, (b) 300 CM, 12 strands, (c) 250 CM, 19 strands, (d) 250 CM, 12 strands

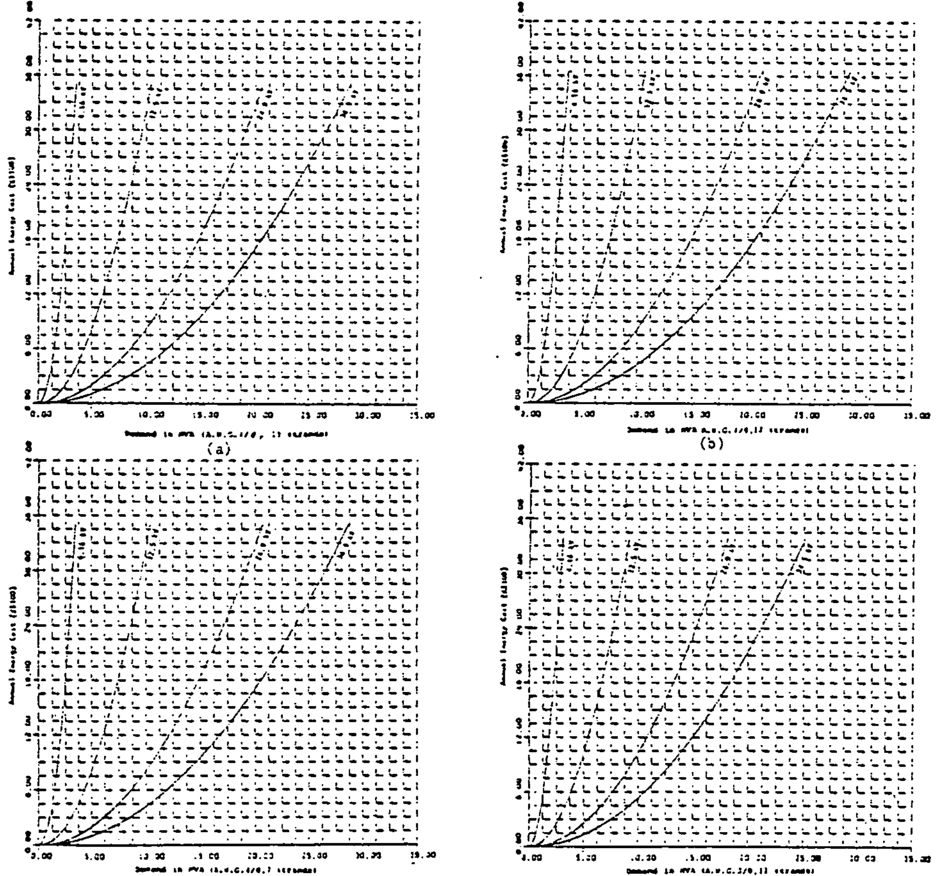
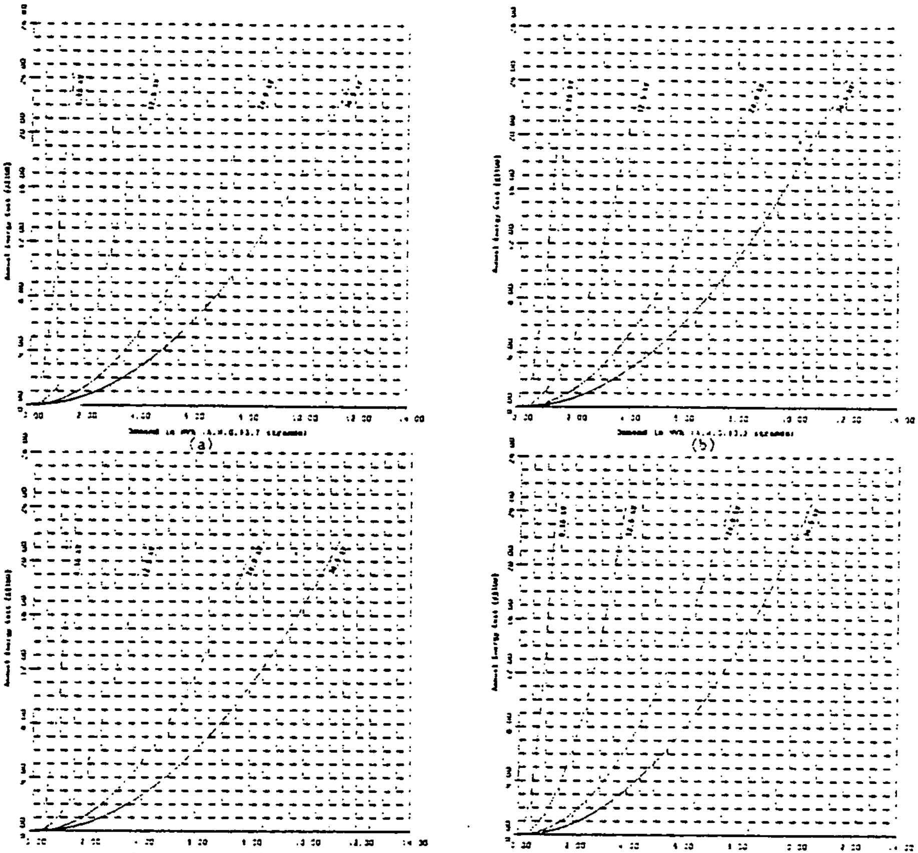
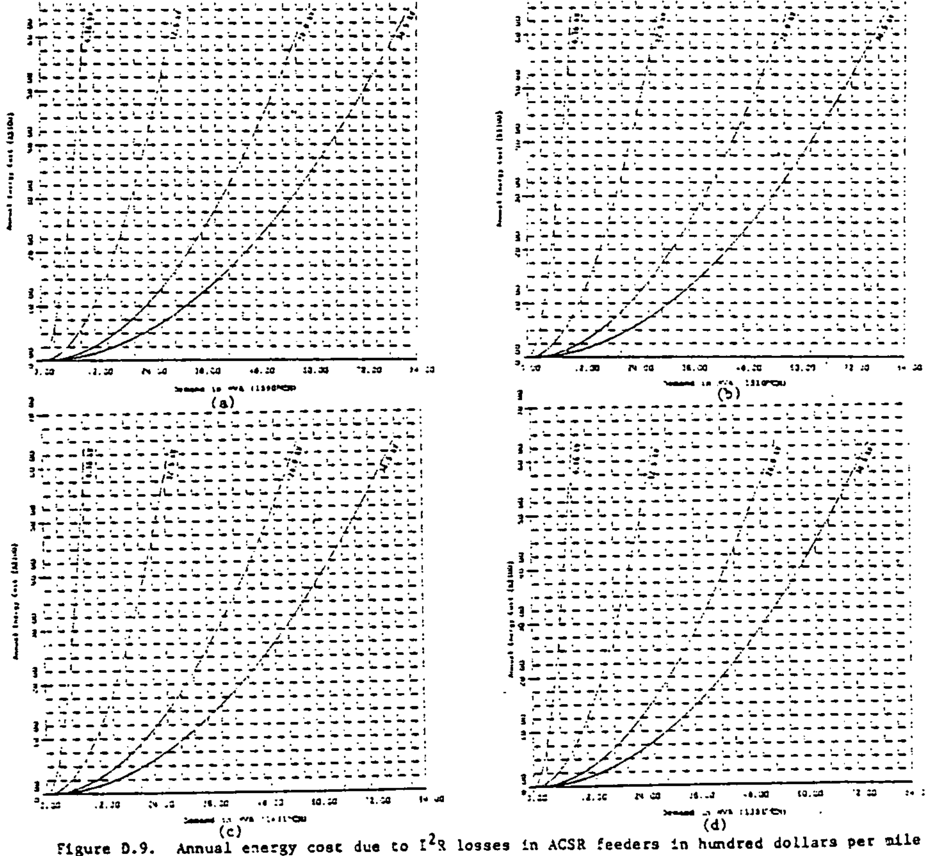


Figure D.5. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for:
(a) A.W.G.4/0,19 strands (b) A.W.G.4/0, 12 strands (c) A.W.G.4/0, 7 strands (d) A.W.G.3/0, 12 strands



(a) (b) (c) (d)

Figure D.8. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) A.W.G.#3, 7 strands (b) A.W.G.#3, 3 strands (c) A.W.G.#3, 1 strand (d) A.W.G.#4, 3 strands



(a) (b) (c) (d)

Figure D.9. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 1590 CM, (b) 1510 CM, (c) 1431 CM, (d) 1351 CM

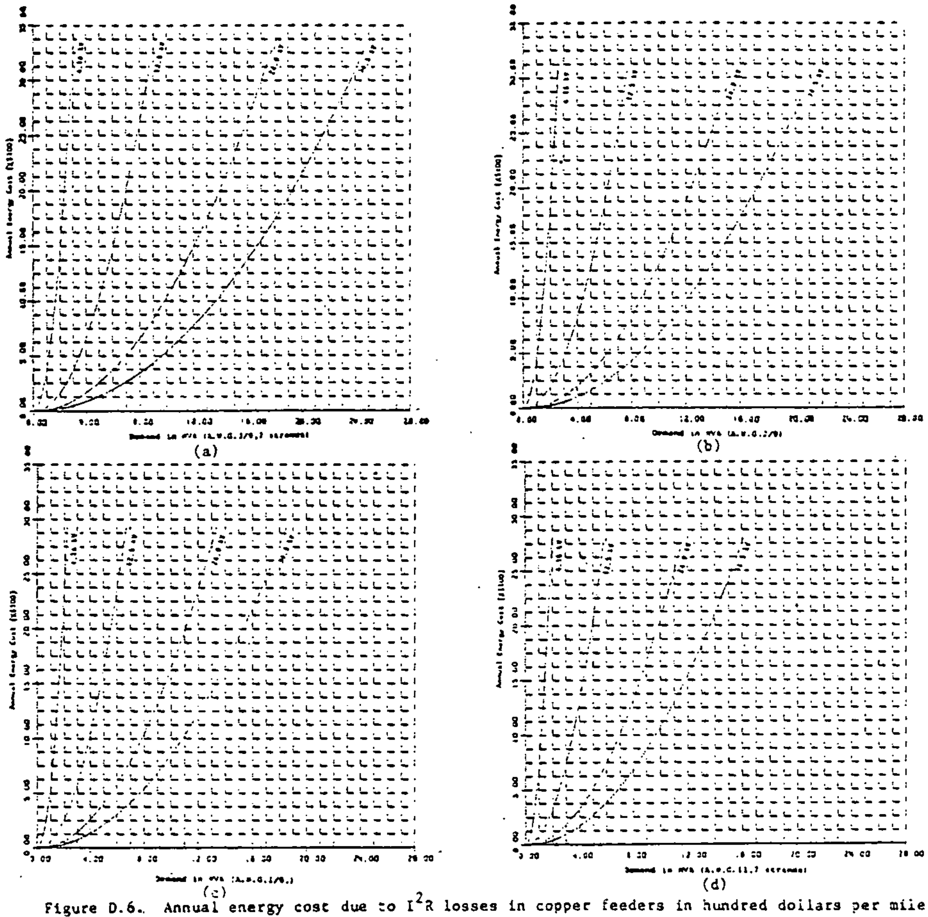


Figure D.6. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) A.W.G.3/0, 7 strands (b) A.W.G.2/0 (c) A.W.G.1/0, (d) A.W.G.#1, 7 strands

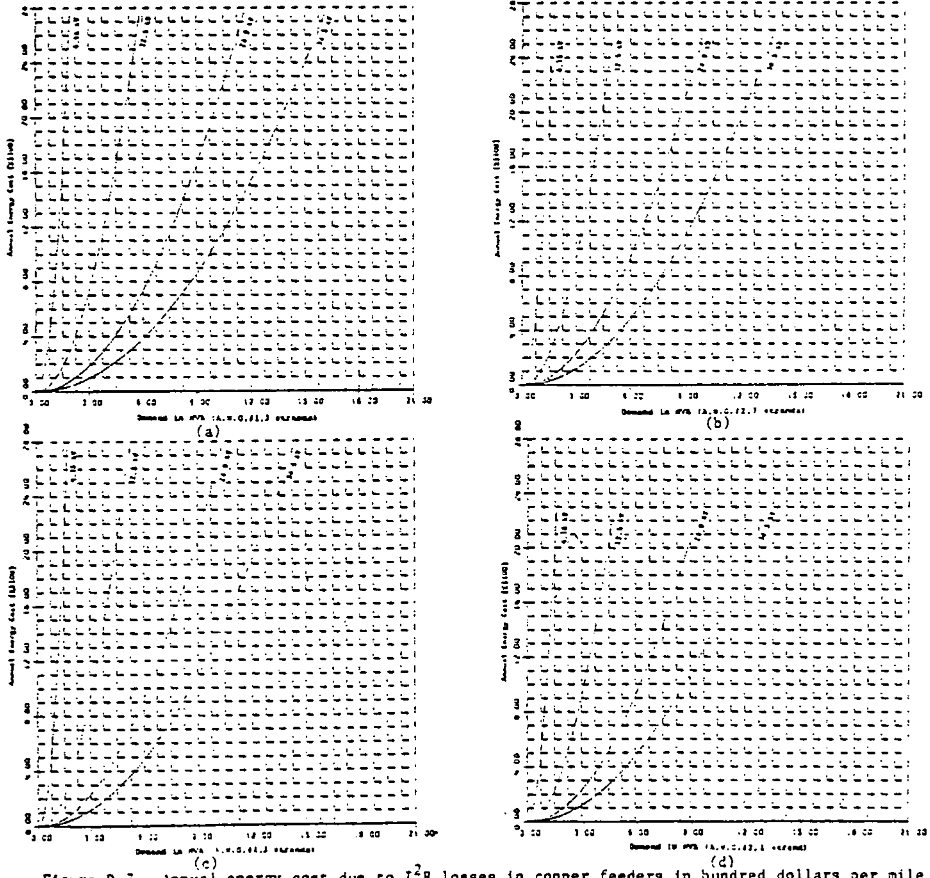
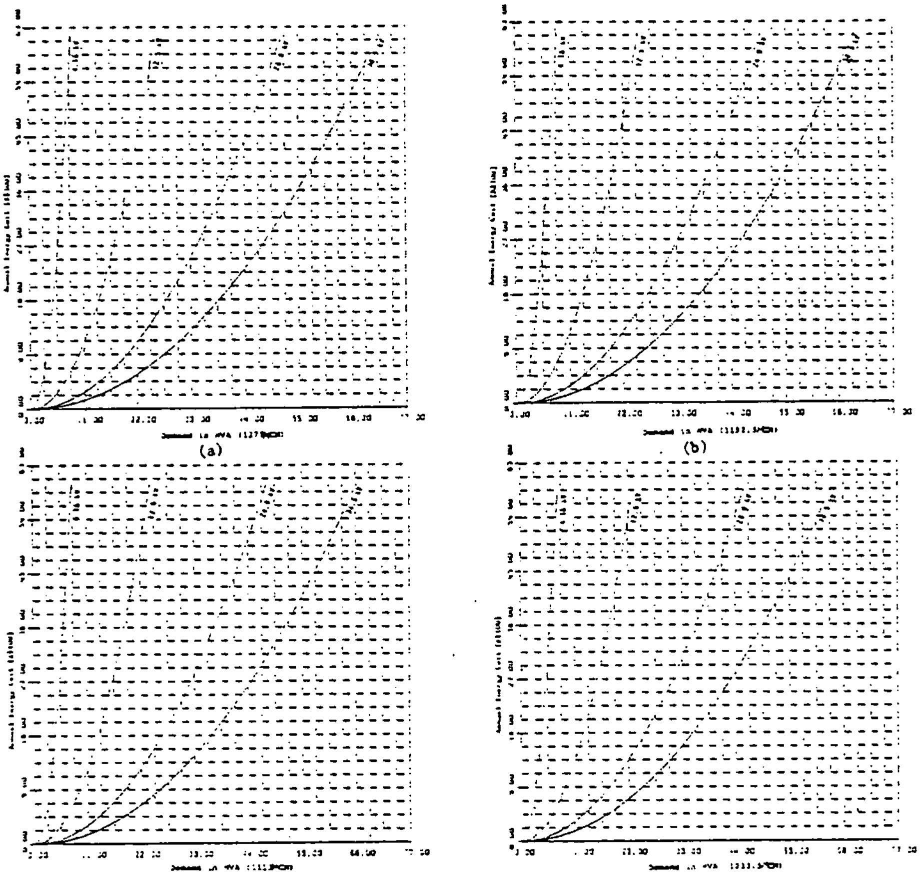
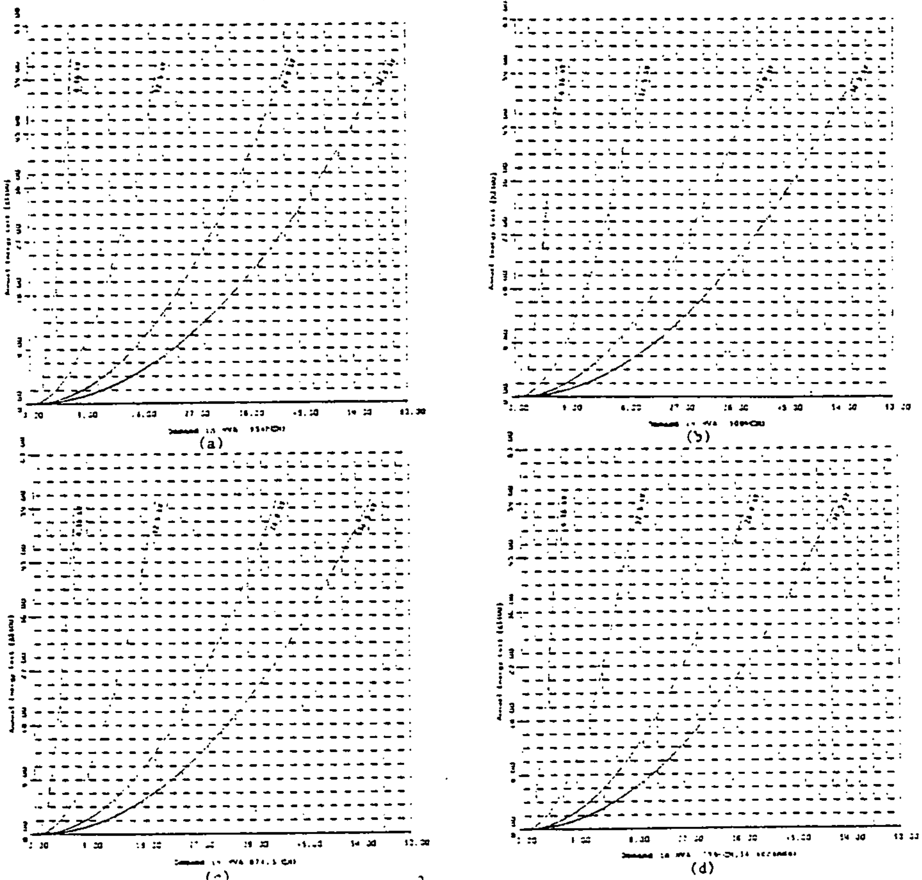


Figure D.7. Annual energy cost due to I^2R losses in copper feeders in hundred dollars per mile for: (a) A.W.G.#1, 3 strands (b) A.W.G.#2, 7 strands (c) A.W.G.#2, 3 strands (d) A.W.G.#2, 1 strand



(a) (b) (c) (d)

Figure D.10. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 1272 CM, (b) 1192.5 CM, (c) 1113 CM, (d) 1033.5 CM



(a) (b) (c) (d)

Figure D.11. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 954 CM, (b) 900 CM, (c) 874.5 CM, (d) 795 CM, 54 strands

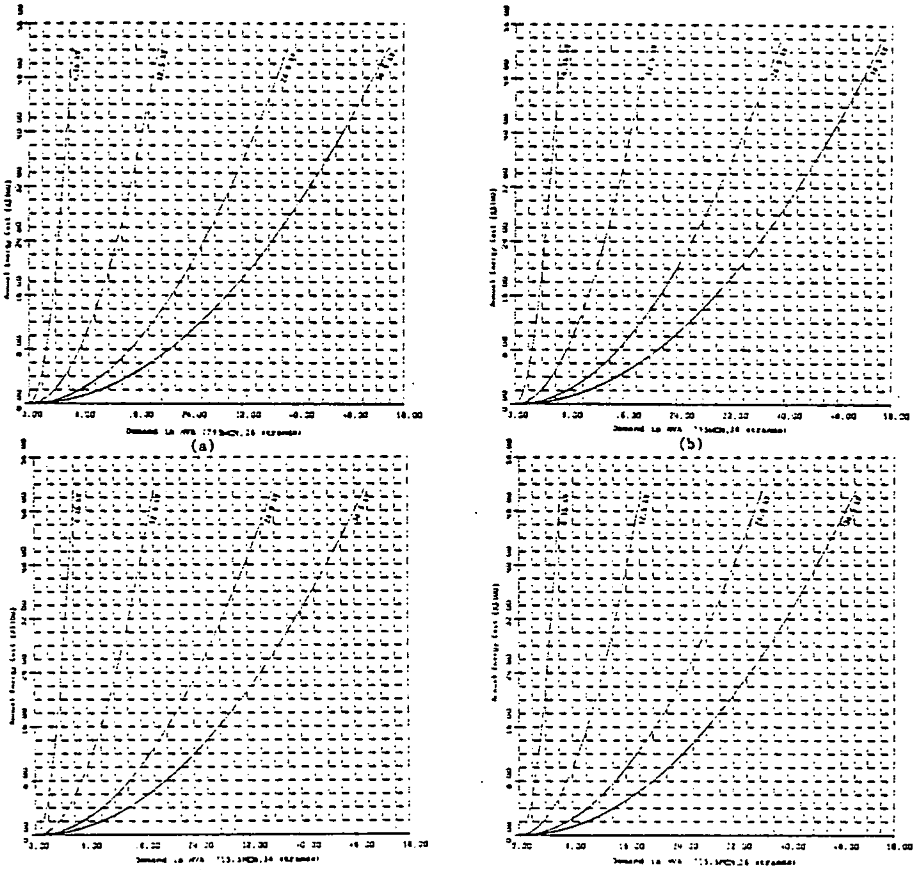

(a) (b) (c) (d)

Figure D.12. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 795 CM, 26 strands, (b) 795 CM, 30 strands, (c) 715.5 CM, 54 strands, (d) 715.5 CM, 26 strands

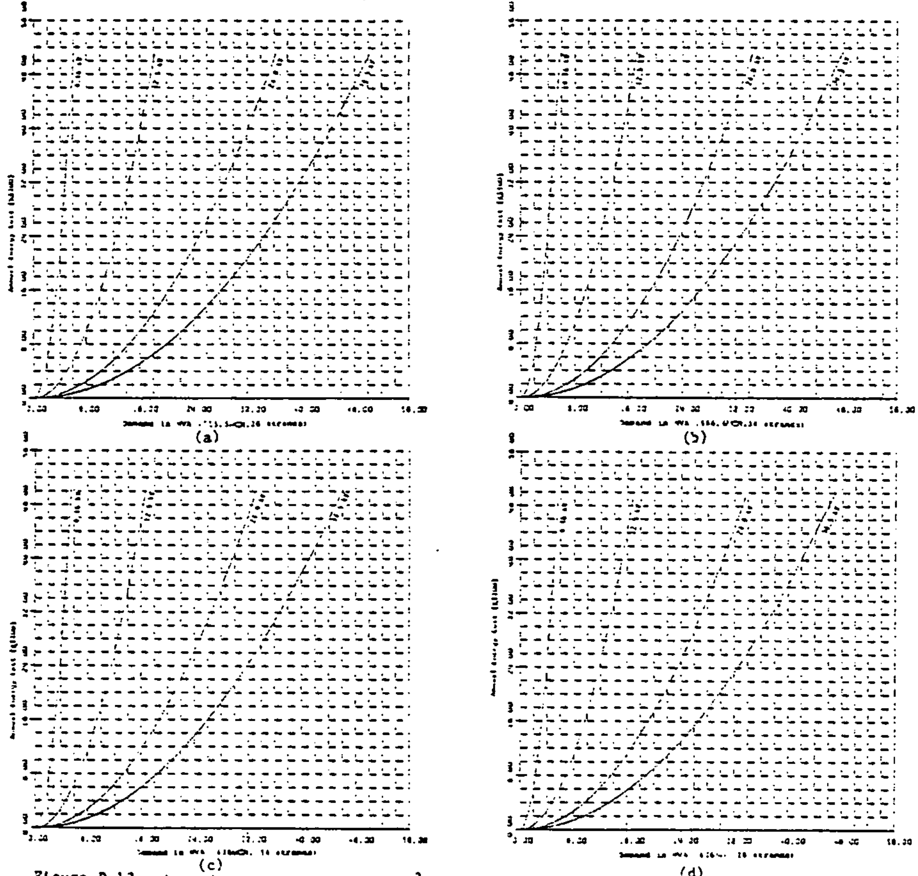

(a) (b) (c) (d)

Figure D.13. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 715.5 CM, 30 strands, (b) 566.6 CM, 54 strands, (c) 636 CM, 54 strands, (d) 636 CM, 26 strands

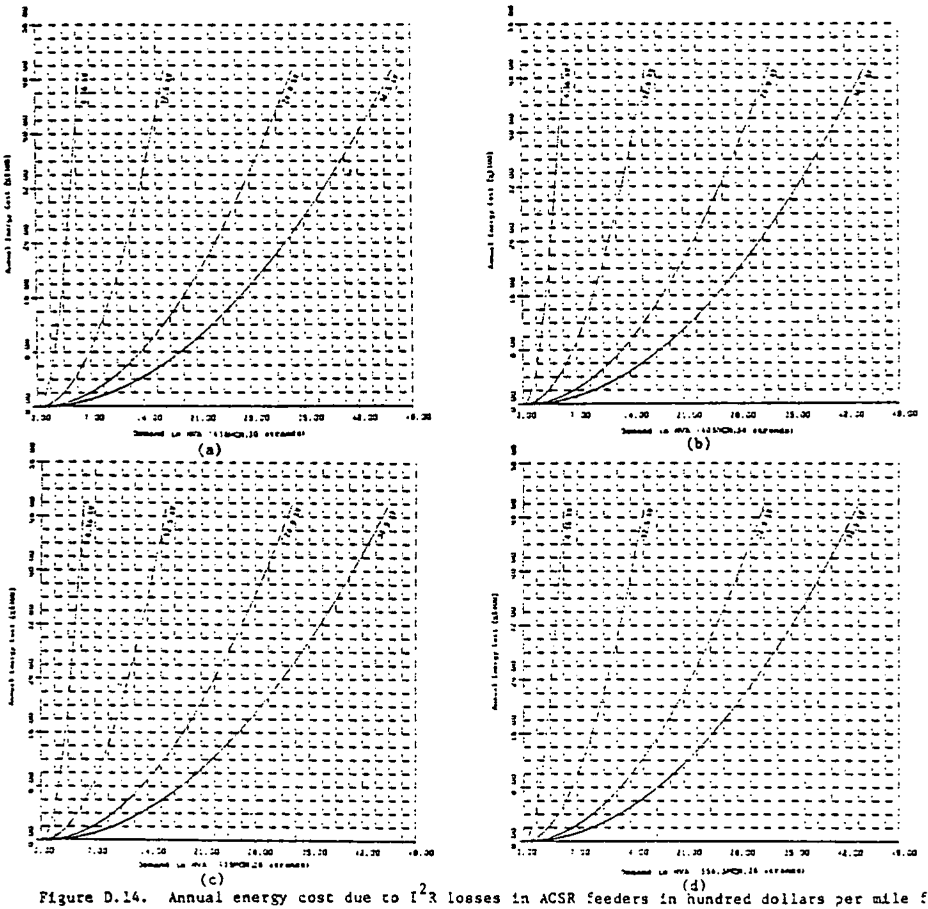


Figure D.14. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 636 CM, 30 strands, (b) 605 CM, 54 strands, (c) 605 CM, 26 strands, (d) 556.5 CM, 26 strands

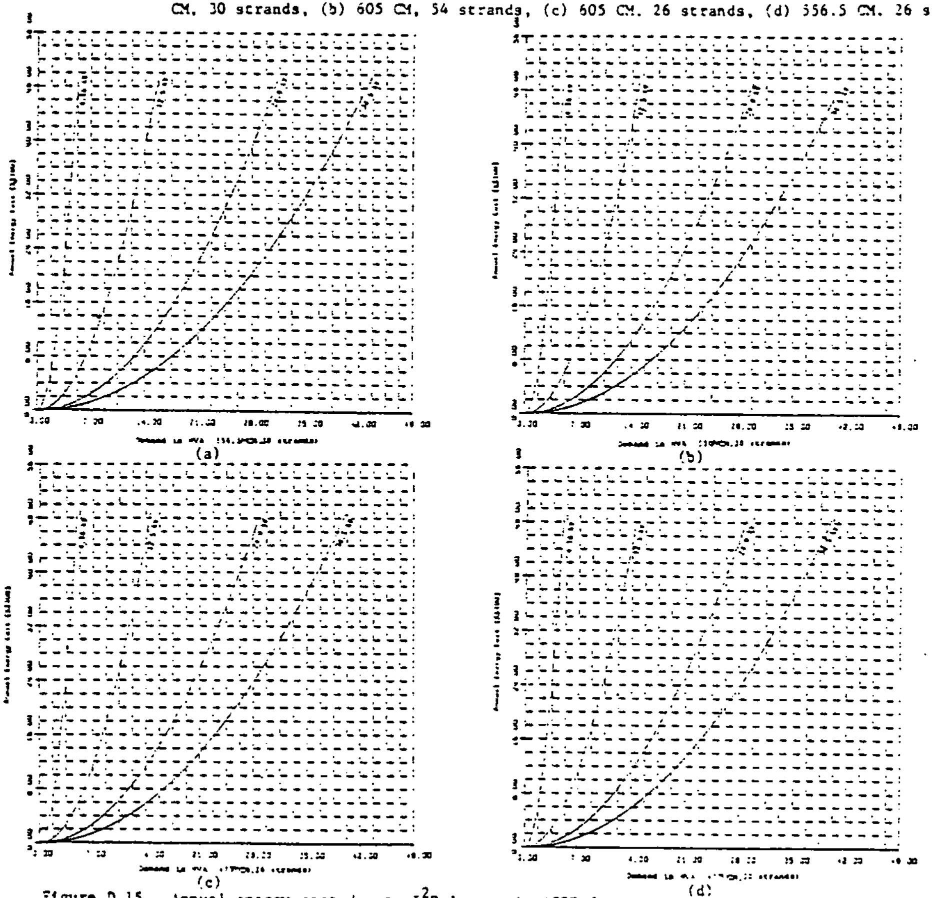


Figure D.15. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 556.5 CM, 30 strands, (b) 500 CM, 30 strands, (c) 477 CM, 26 strands, (d) 477 CM, 30 strands

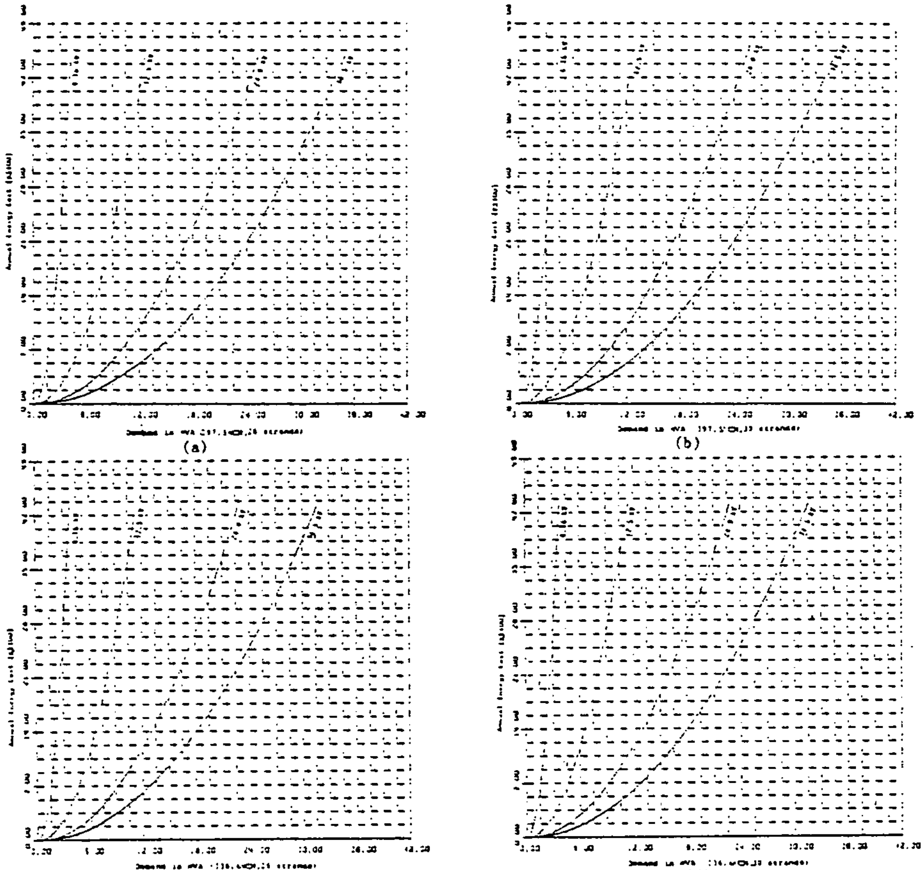


Figure D.16. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 397.5 CM, 26 strands (b) 397.5 CM, 30 strands (c) 336.4 CM, 26 strands (d) 336.4 CM, 30 strands

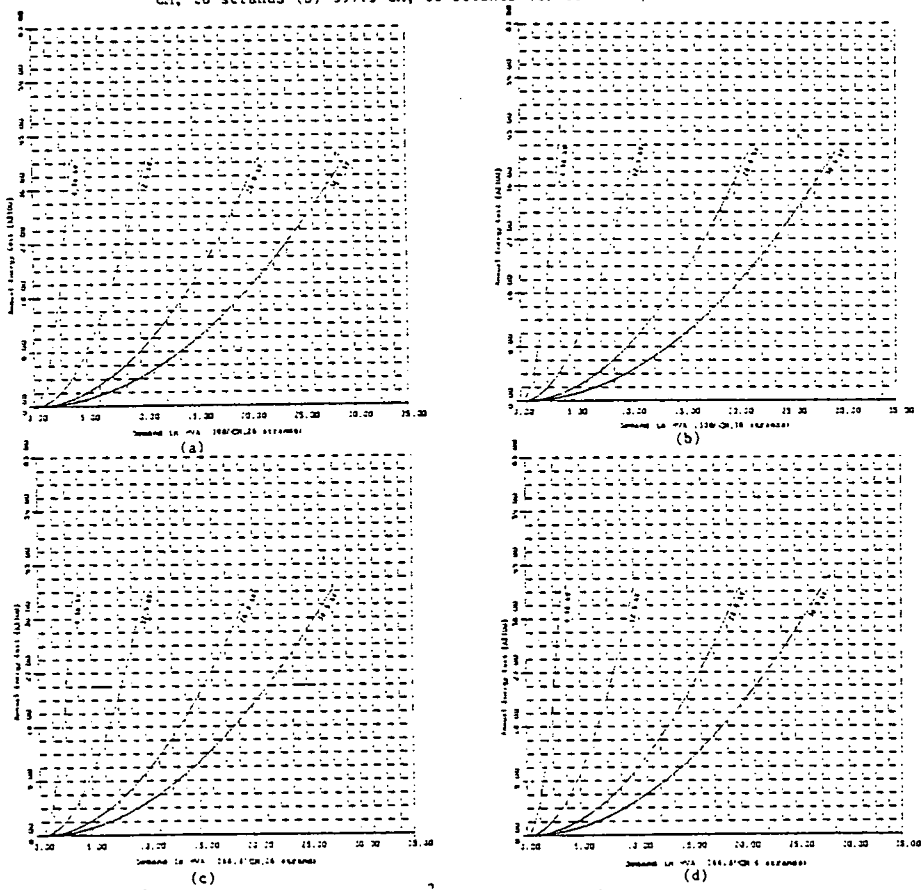
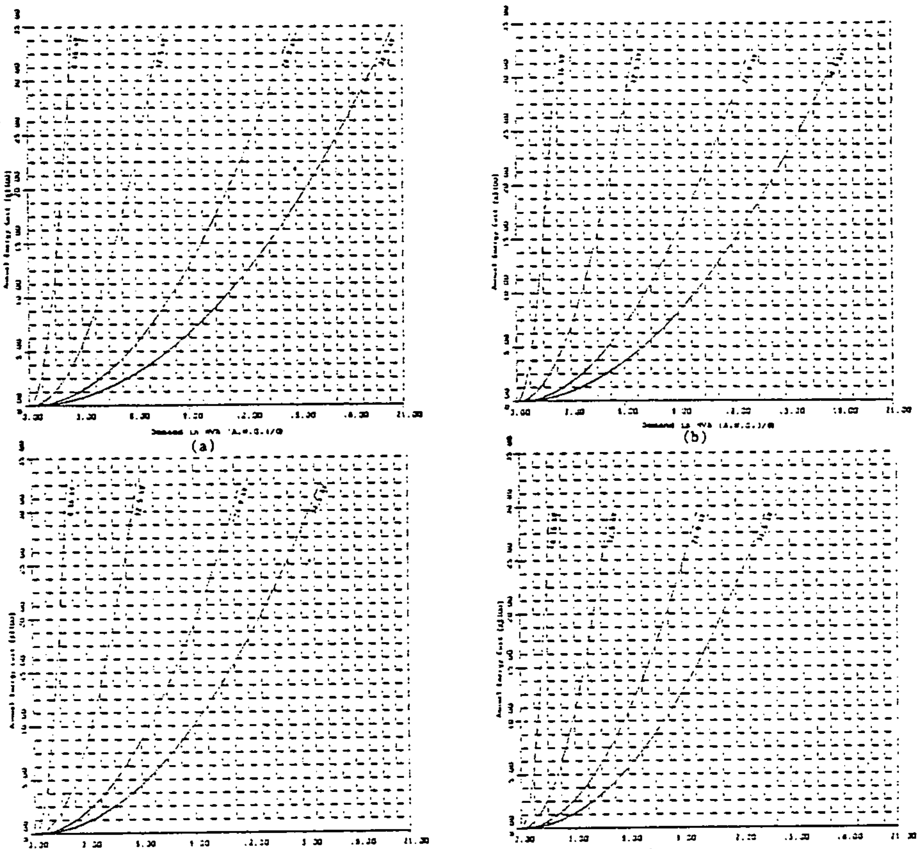
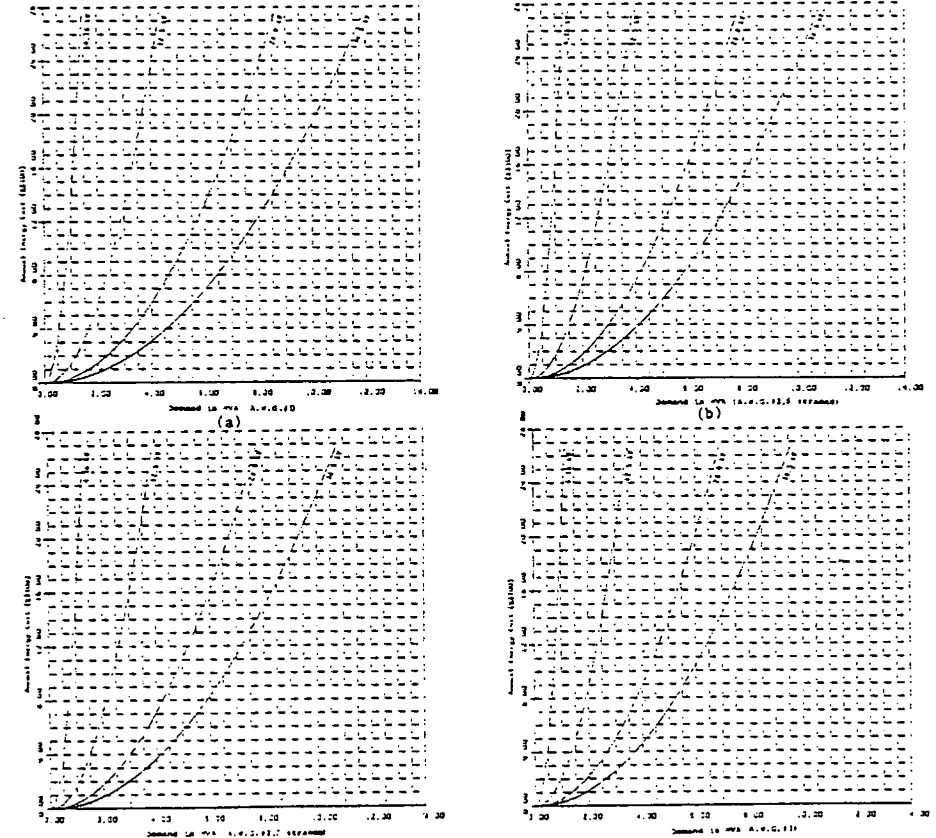


Figure D.17. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) 300 CM, 26 strands (b) 300 CM, 30 strands (c) 266.8 CM, 26 strands (d) 266.8 CM, 6 strands



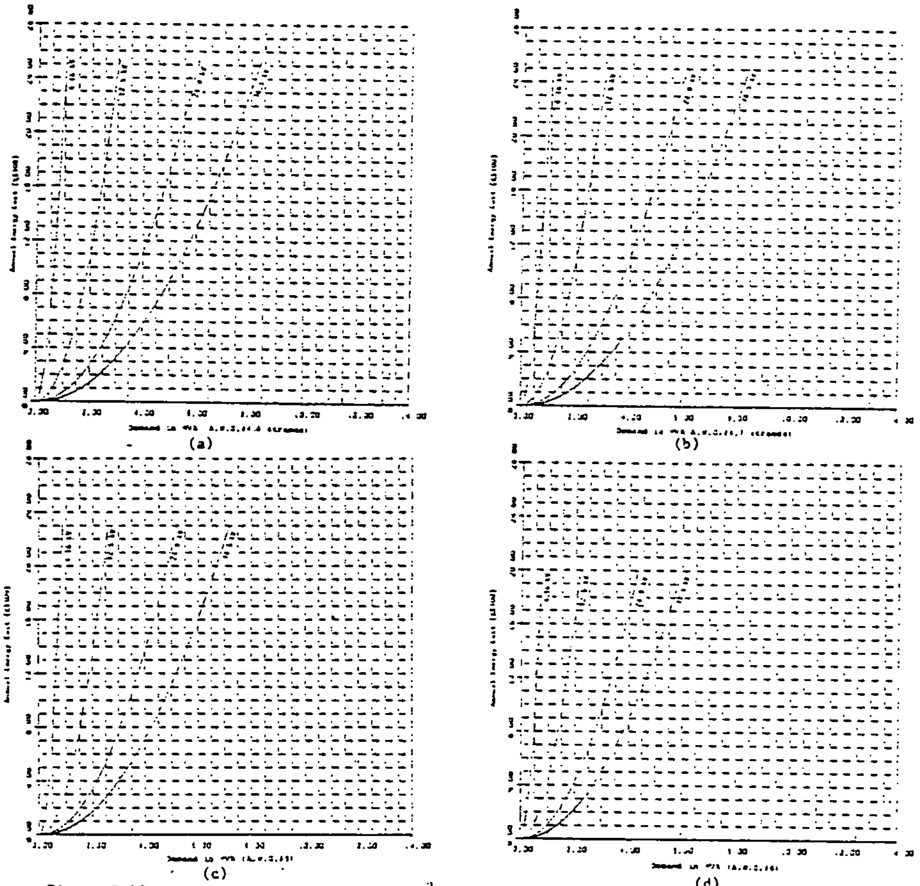
(a) (b) (c) (d)

Figure D.18. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.4/0, (b) A.W.G.3/0, (c) A.W.G.2/0, (d) A.W.G.1/0



(a) (b) (c) (d)

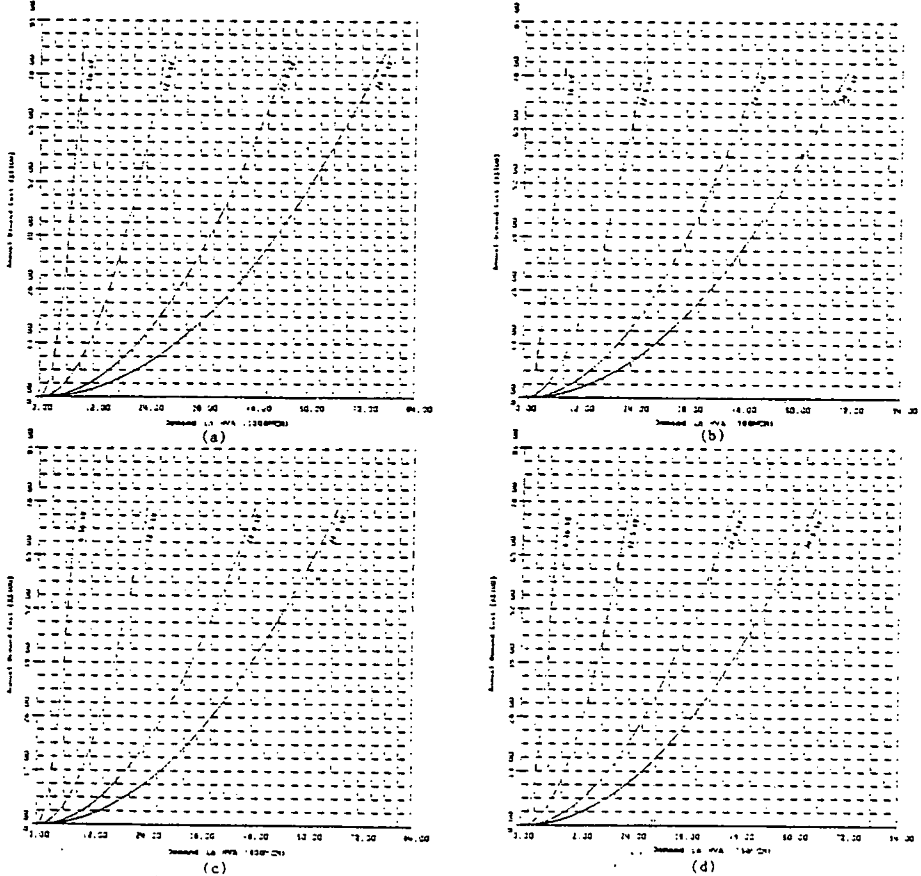
Figure D.19. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.#3, (b) A.W.G.#2, 6 strands, (c) A.W.G.#2, 7 strands, (d) A.W.G.#3



(a) (b)

(c) (d)

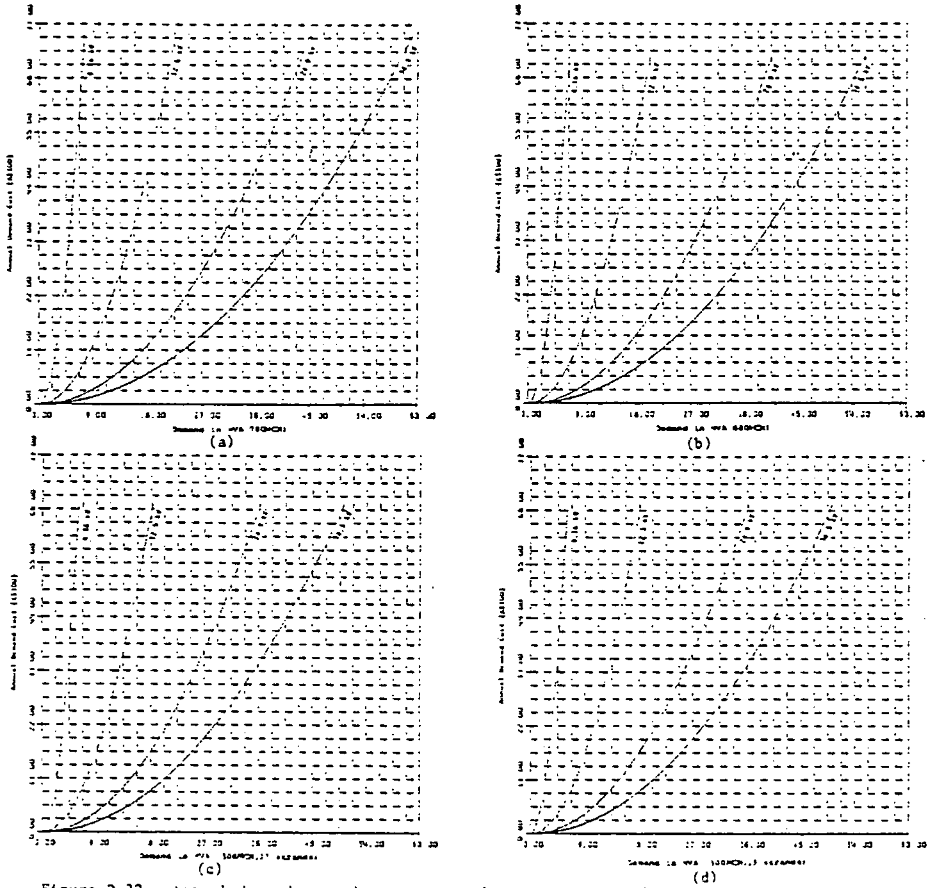
Figure D.20. Annual energy cost due to I^2R losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.#4, 6 strands, (b) A.W.G.#4, 7 strands, (c) A.W.G.#5, (d) A.W.G.#6



(a) (b)

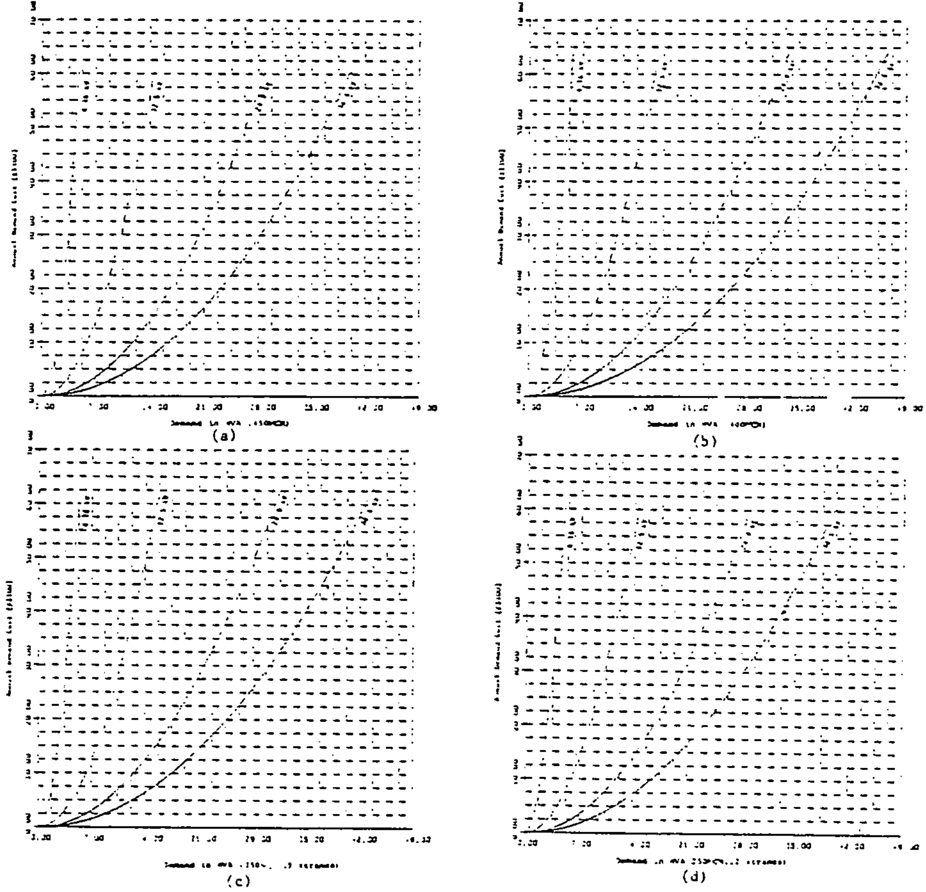
(c) (d)

Figure D.21. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) 1000 CM, (b) 900 CM, (c) 300 CM, (d) 750 CM



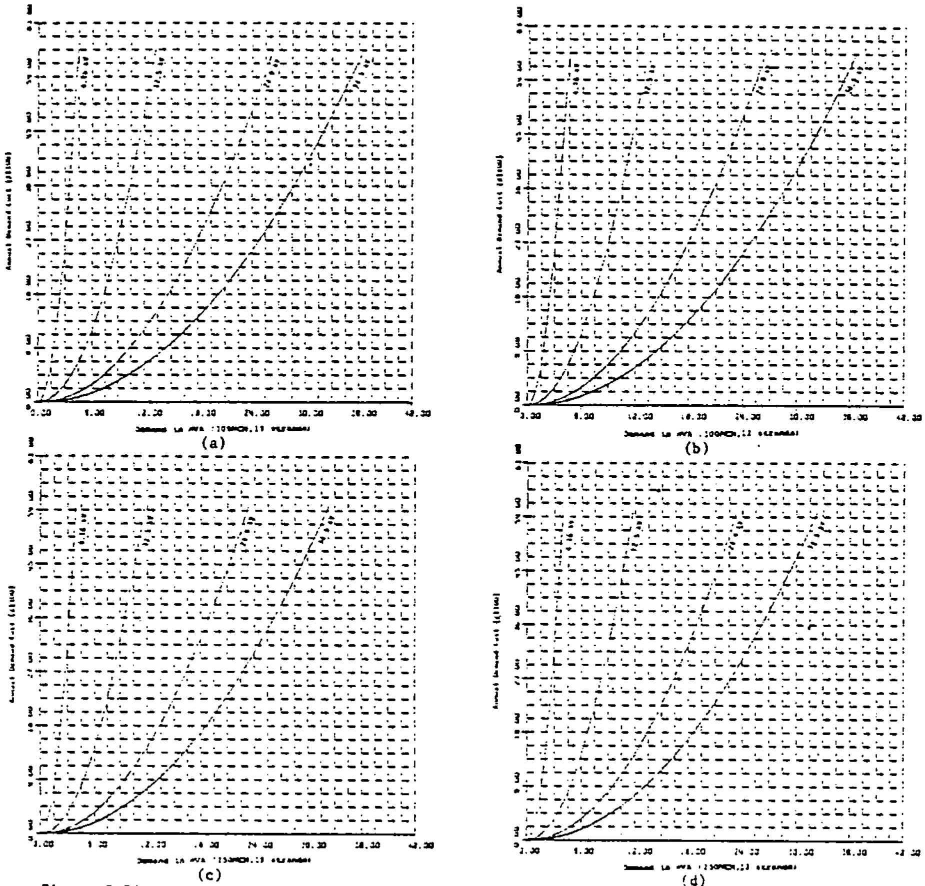
(a) (b) (c) (d)

Figure D.22. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) 700 CM, (b) 600 CM, (c) 500 CM, 37 strands, (d) 500 CM, 19 strands



(a) (b) (c) (d)

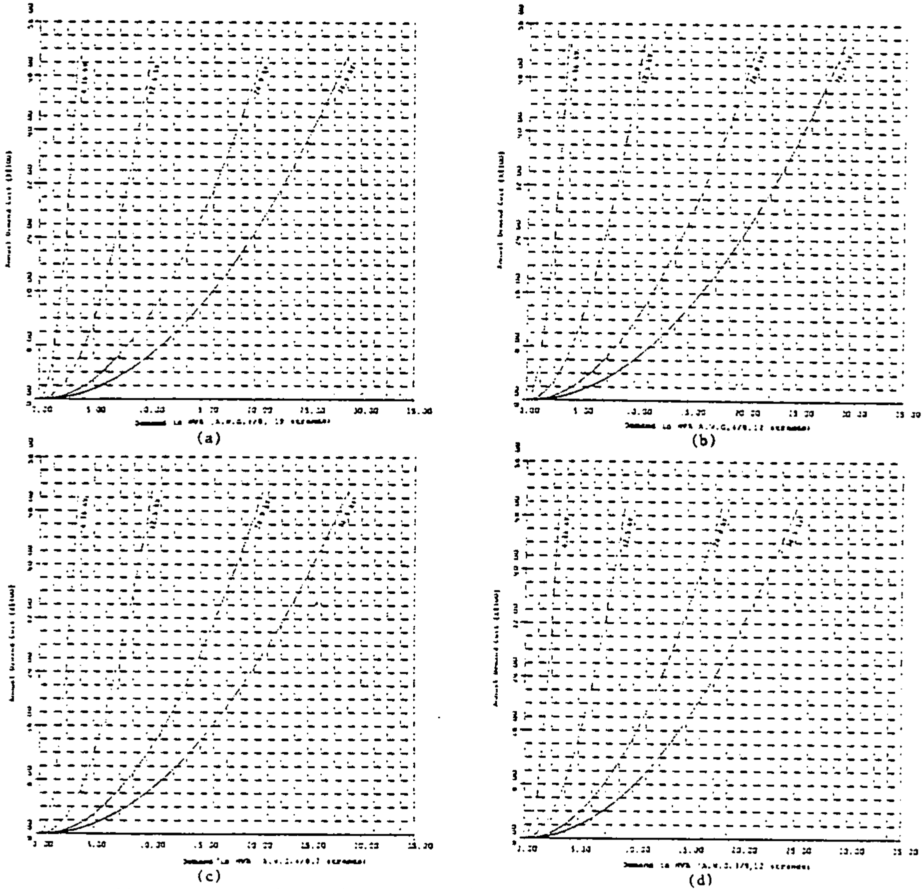
Figure D.23. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) 450 CM, (b) 400 CM, (c) 350 CM, 19 strands, (d) 350 CM, 12 strands



(a) (b)

(c) (d)

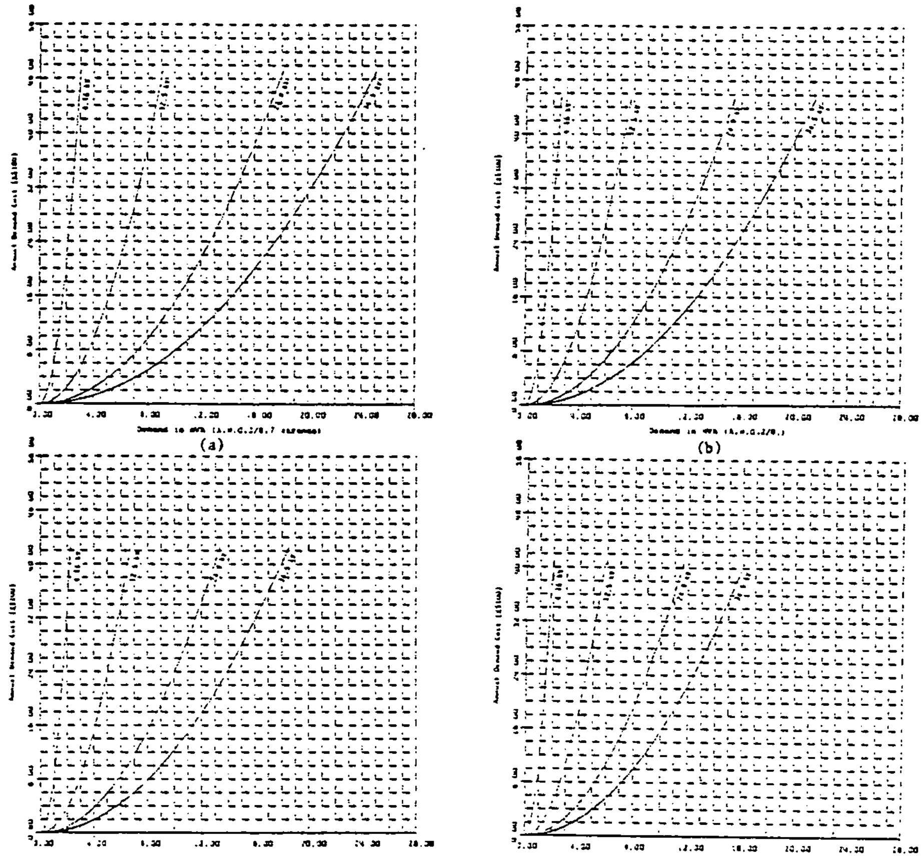
Figure D.24. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) 300 CM, 19 strands, (b) 300 CM, 12 strands, (c) 250 CM, 19 strands, (d) 250 CM, 12 strands



(a) (b)

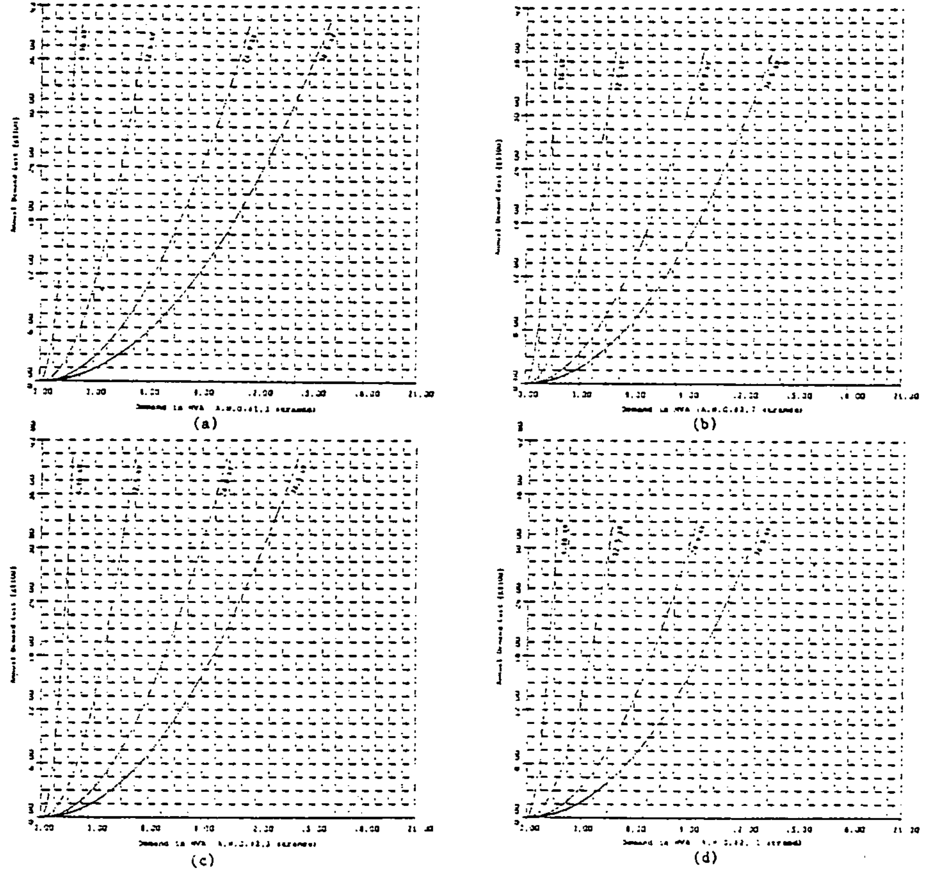
(c) (d)

Figure D.25. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) A.W.G.4/0, 19 strands, (b) A.W.G.4/0, 12 strands, (c) A.W.G.4/0, 7 strands, (d) A.W.G.3/0, 12 strands



(a) (b) (c) (d)

Figure D.26. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) A.W.G.3/0, 7 strands, (b) A.W.G.2/0, (c) A.W.G.1/0, (d) A.W.G.#1, 7 strands



(a) (b) (c) (d)

Figure D.27. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) A.W.G.#1, 3 strands, (b) A.W.G.#2, 7 strands, (c) A.W.G.#2, 3 strands, (d) A.W.G.#2, 1 strand

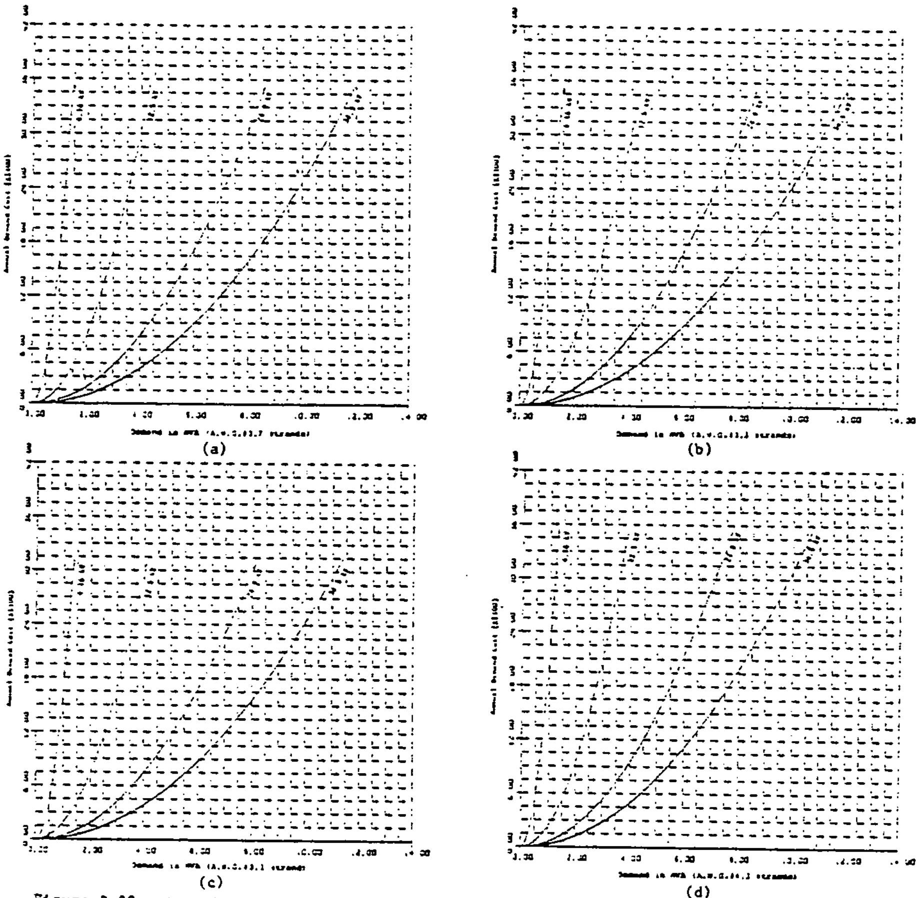


Figure D.28. Annual demand cost due to energy losses in copper feeders in hundred dollars per mile for: (a) A.W.G.#3, 7 strands, (b) A.W.G. #3, 3 strands, (c) A.W.G.#3, 1 strand (d) A.W.G.#4, 3 strands

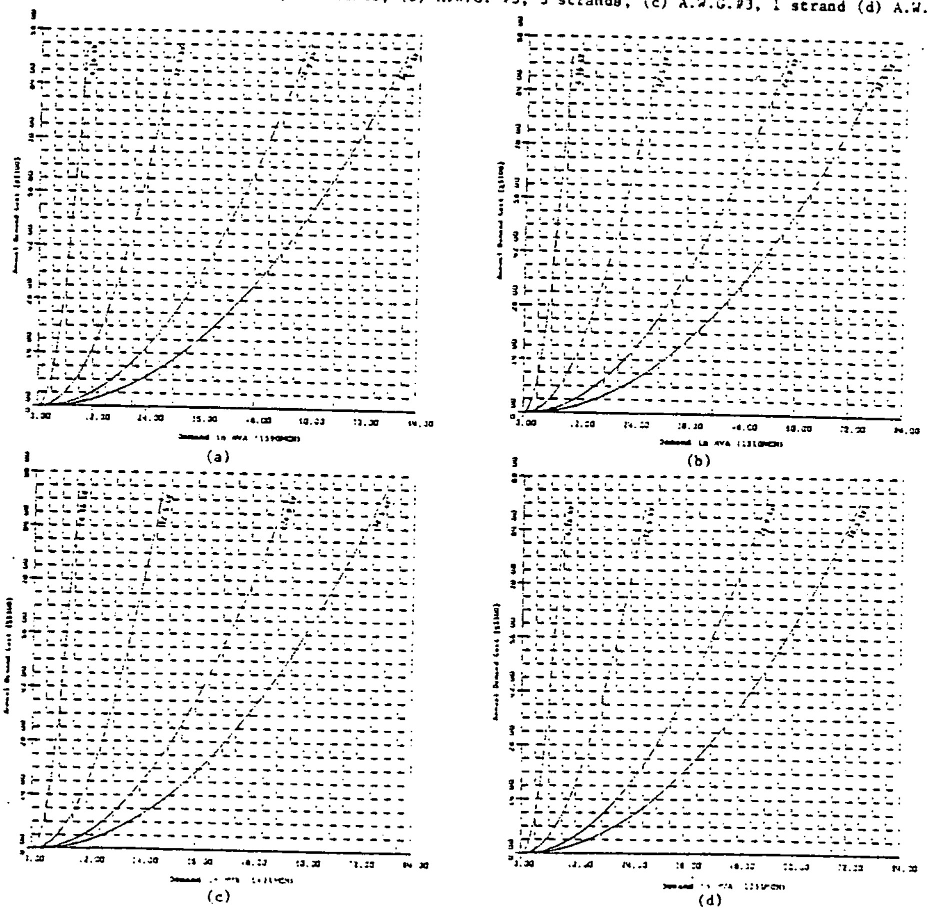


Figure D.29. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 1590 CM, (b) 1510 CM, (c) 1431 CM, (d) 1351 CM

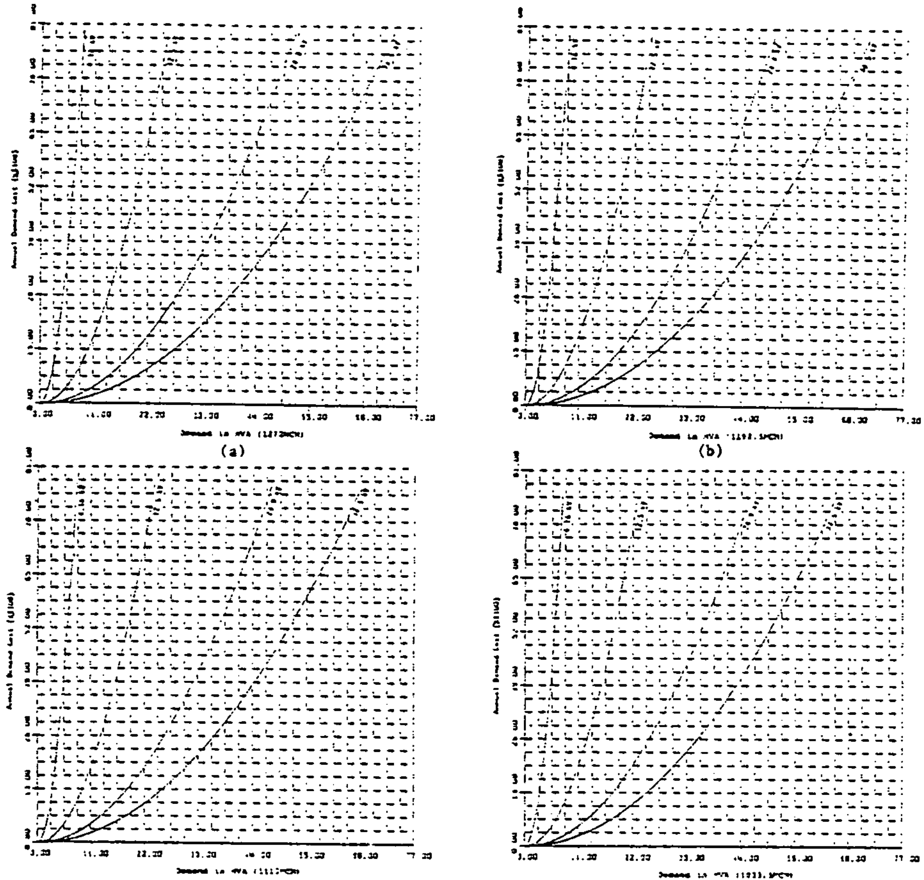


Figure D.30. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 1272 CM, (b) 1192.5 CM, (c) 1113 CM (d) 1033.5 CM

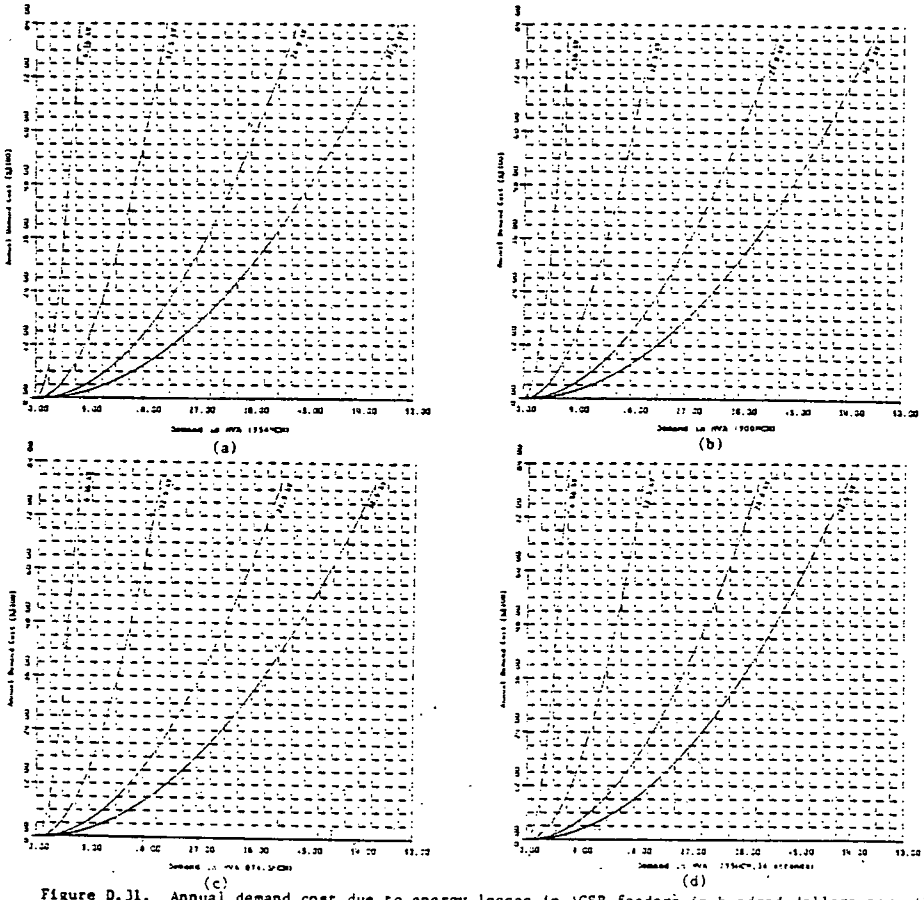
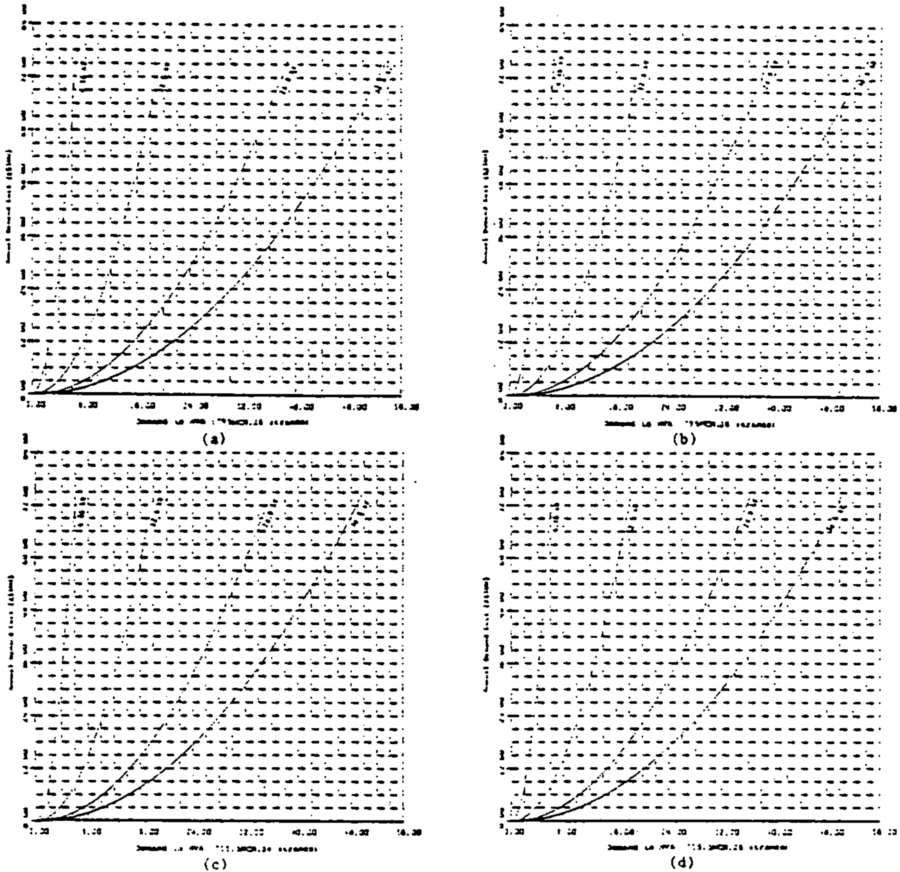
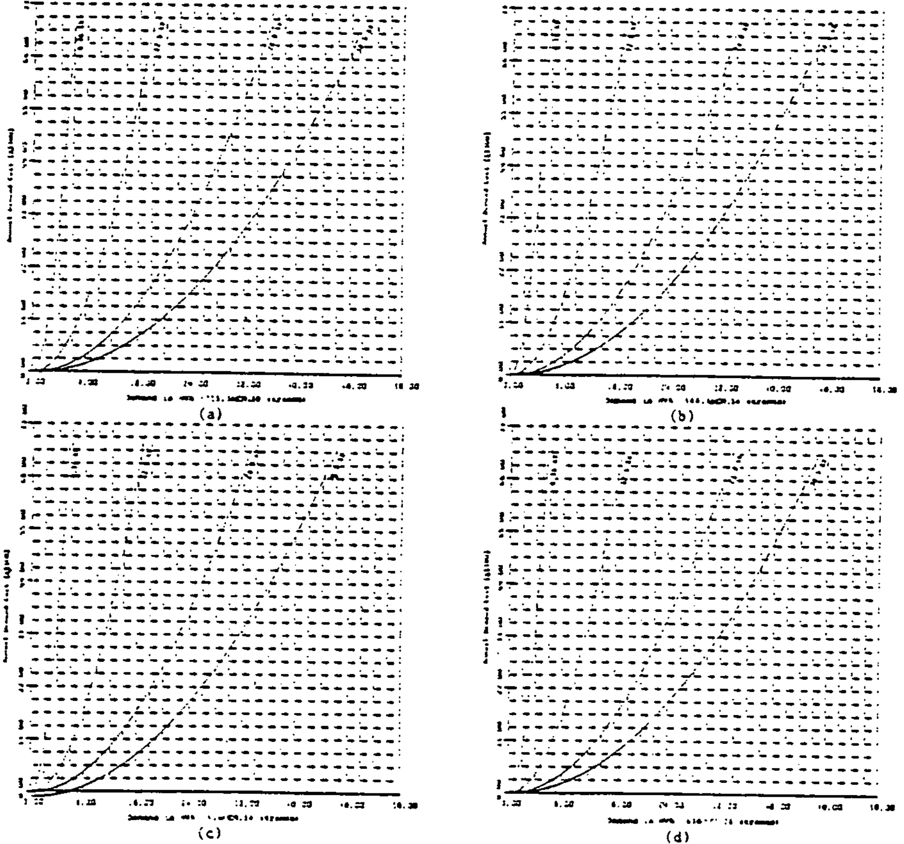


Figure D.31. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 954 CM, (b) 900 CM, (c) 874.5 CM, (d) 795 CM, 54 strands



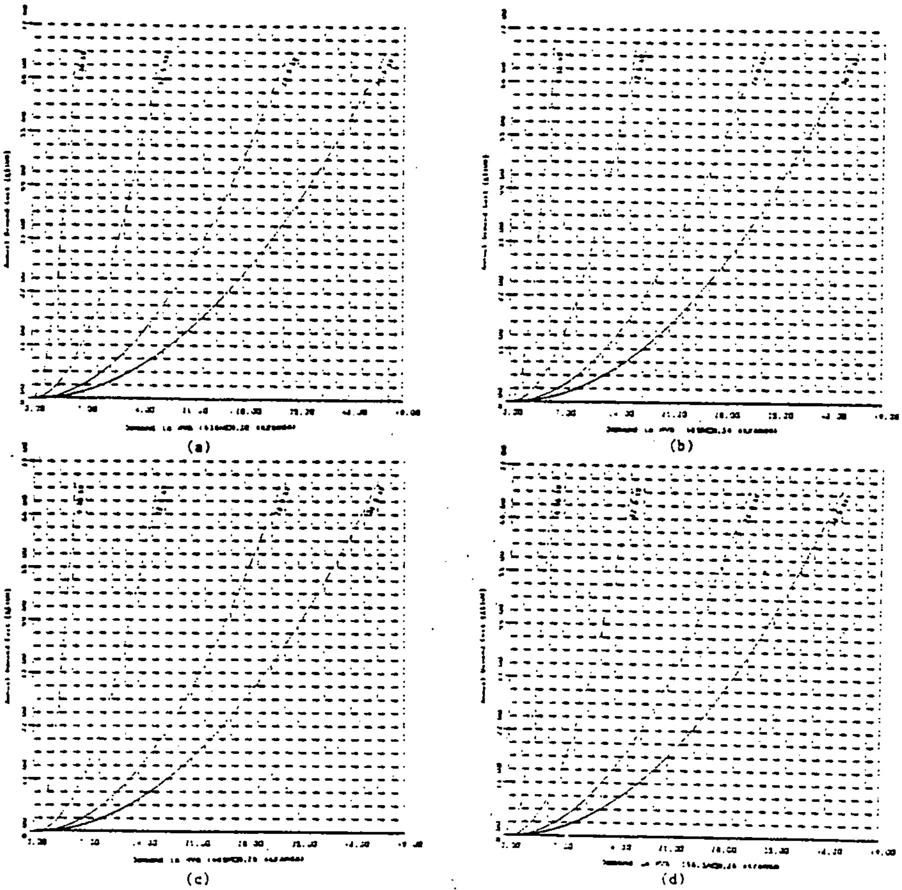
(a) (b) (c) (d)

Figure D.32. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 795 CM, 26 strands (b) 795 CM, 30 strands, (c) 715.5 CM, 54 strands, (d) 715.5 CM, 26 strands



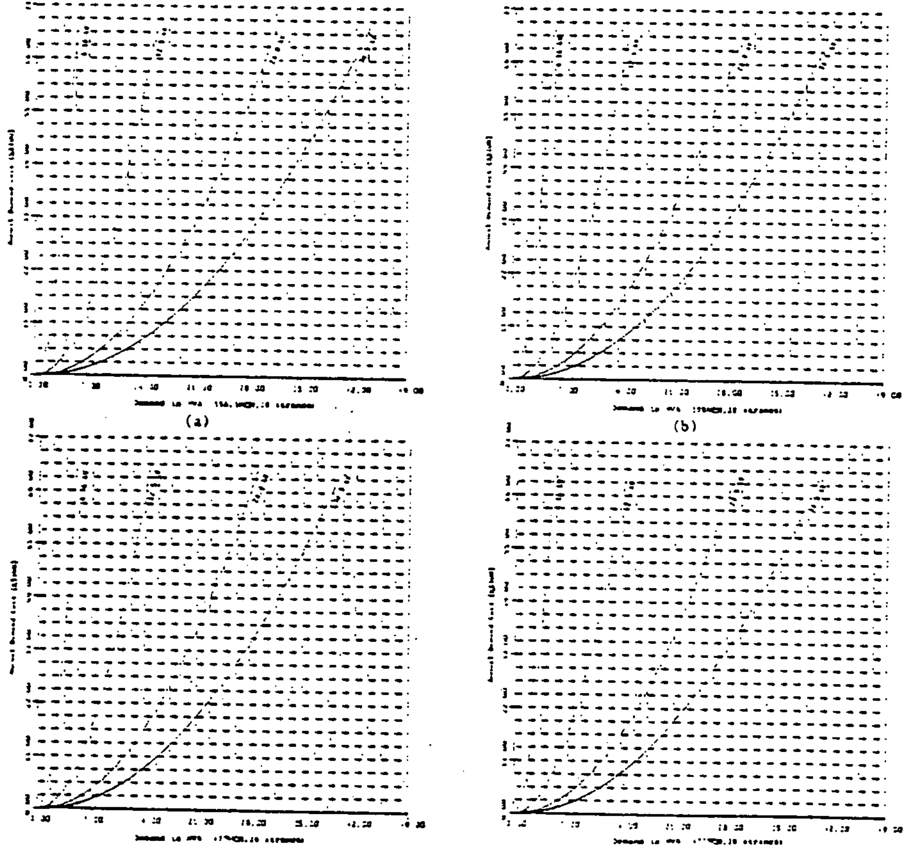
(a) (b) (c) (d)

Figure D.33. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 715.5 CM, 30 strands, (b) 666.6 CM, 54 strands, (c) 636 CM, 54 strands, (d) 636 CM, 26 strands



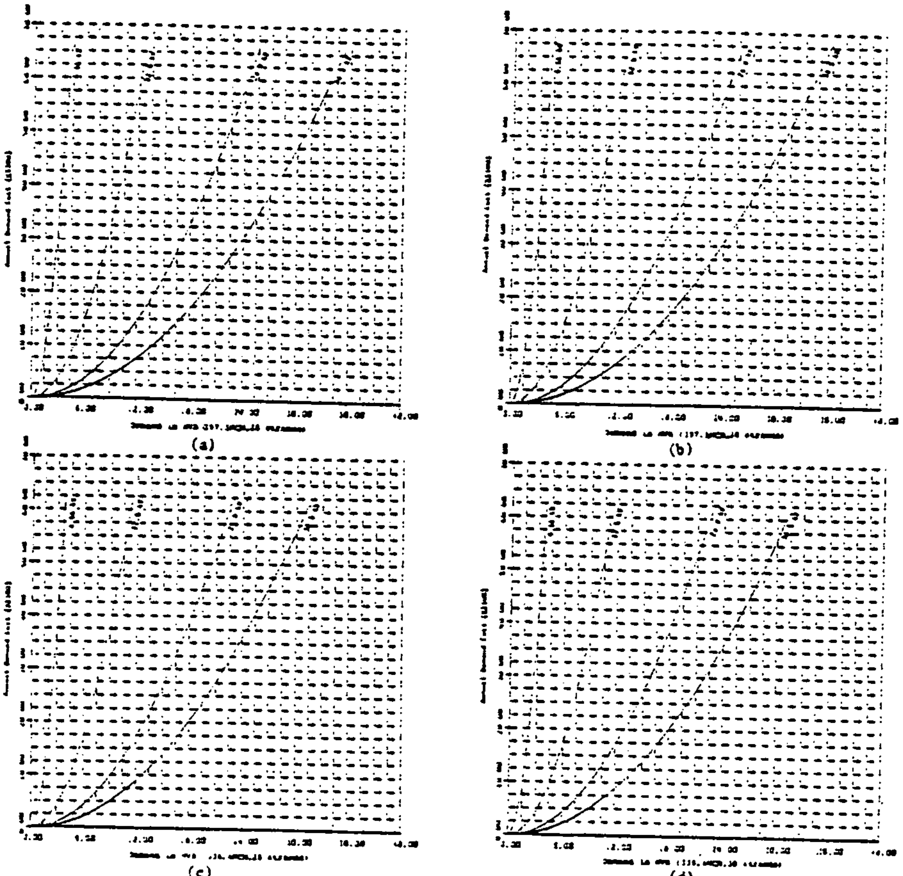
(a) (b) (c) (d)

Figure D.34. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 636 CM, 30 strands, (b) 605 CM, 54 strands, (c) 605 CM, 26 strands, (d) 556.5 CM, 26 strands



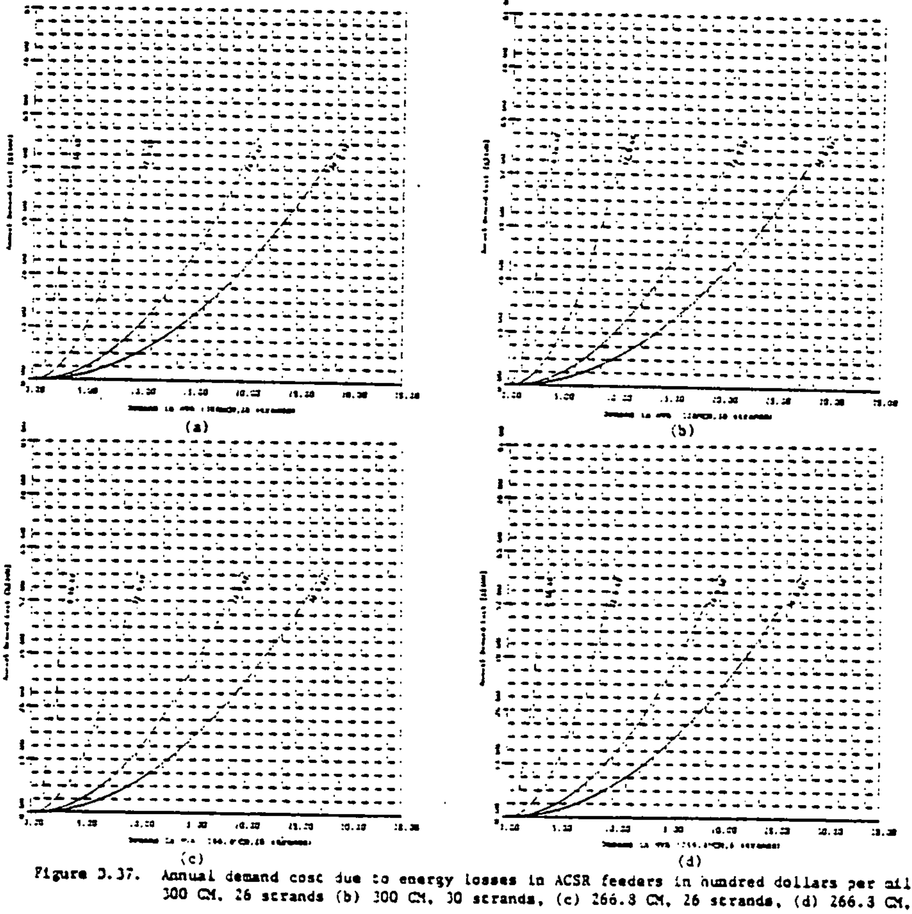
(a) (b) (c) (d)

Figure D.35. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 556.5 CM, 30 strands (b) 500 CM, 30 strands (c) 477 CM, 26 strands, (d) 477 CM, 30 strands



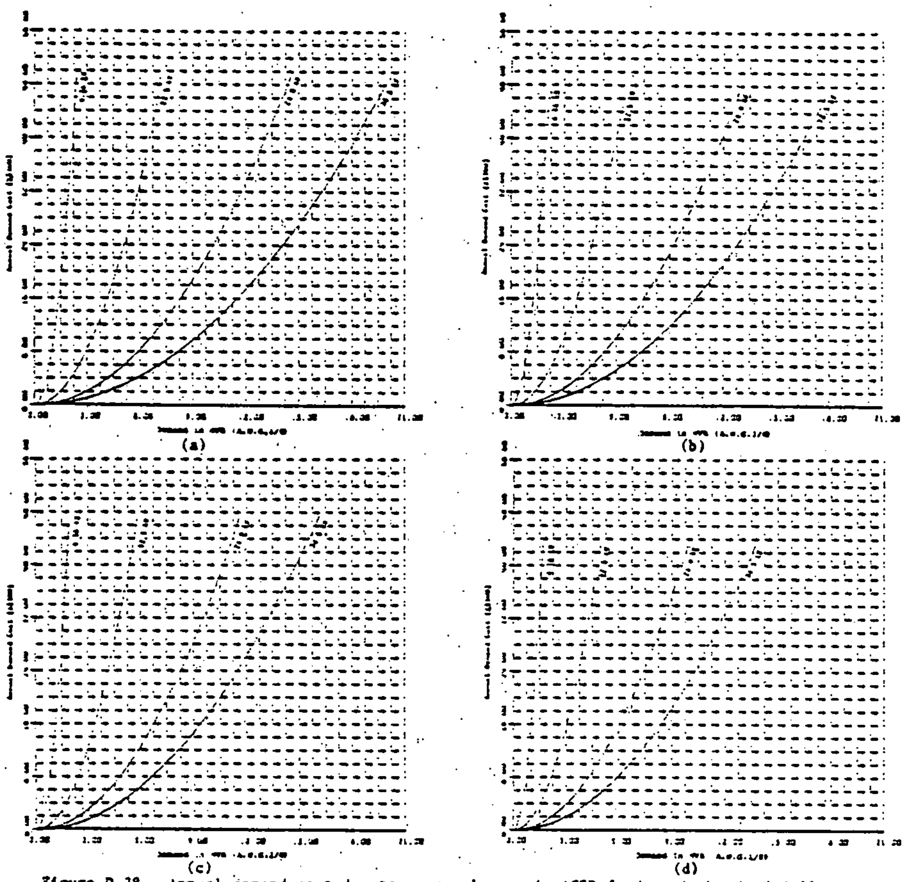
(a) (b) (c) (d)

Figure D.36. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 397.5 CM, 26 strands, (b) 397.5 CM, 30 strands (c) 336.4 CM, 26 strands, (d) 336.4 CM, 30 strands



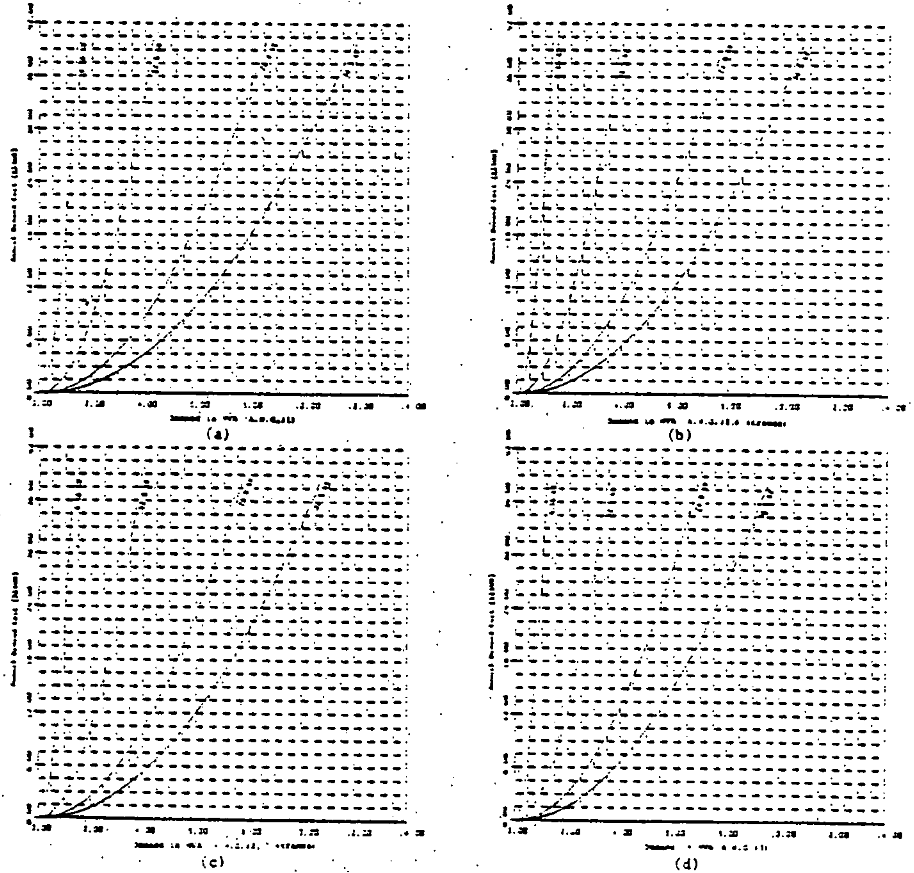
(a) (b) (c) (d)

Figure D.37. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) 300 CM, 26 strands (b) 300 CM, 30 strands, (c) 266.8 CM, 26 strands, (d) 266.3 CM, 6 strands



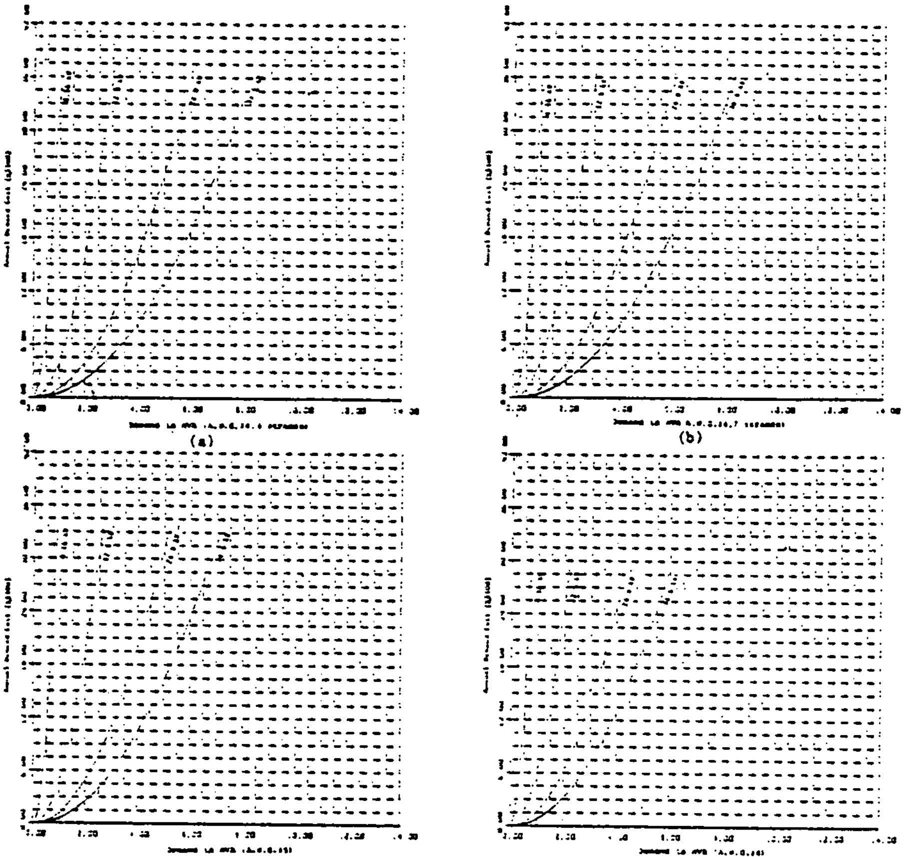
(a) (b) (c) (d)

Figure D.38. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.4/0, (b) A.W.G.3/0, (c) A.W.G.2/0, (d) A.W.G.1/0



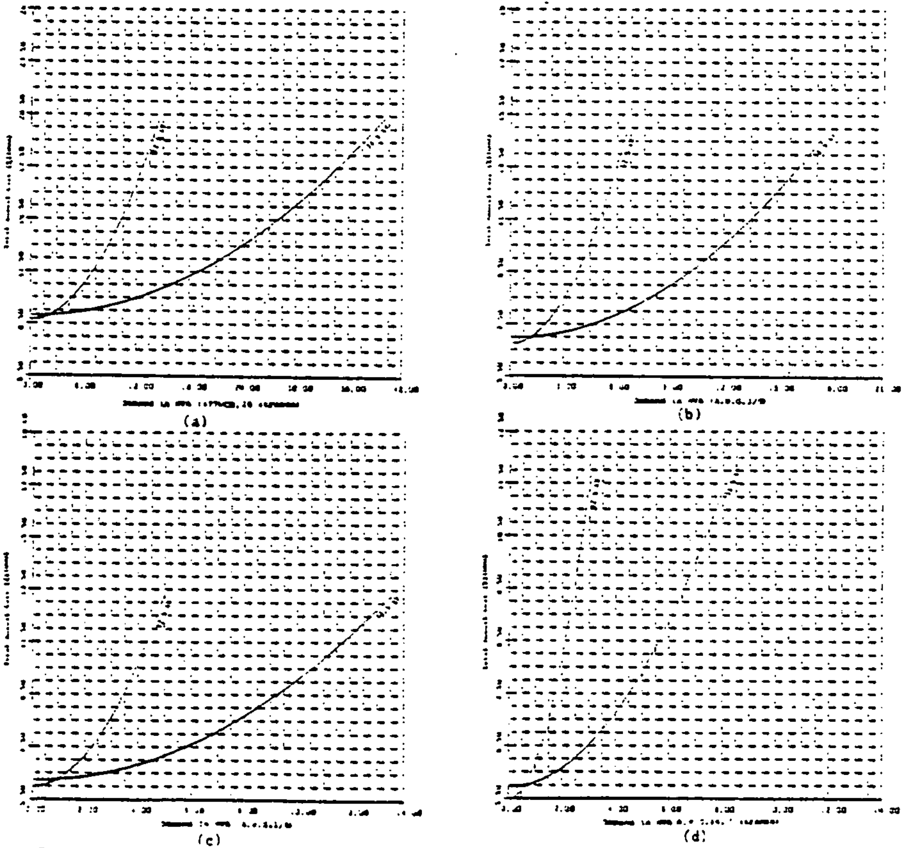
(a) (b) (c) (d)

Figure D.39. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.#1, (b) A.W.G.#2, 6 strands, (c) A.W.G.#2, 7 strands, (d) A.W.G.#3



(a) (b) (c) (d)

Figure D.40. Annual demand cost due to energy losses in ACSR feeders in hundred dollars per mile for: (a) A.W.G.#4, 6 strands (b) A.W.G.#4, 7 strands, (c) A.W.G.#5, (d) A.W.G.#6



(a) (b) (c) (d)

Figure D.41. Total annual equivalent cost of ACSR feeders for urban areas in thousand dollars per mile for: (a) 477 CM, 26 strands, (b) A.W.G. 3/0, (c) A.W.G.1/0, (d) A.W.G.#4, 7 strands

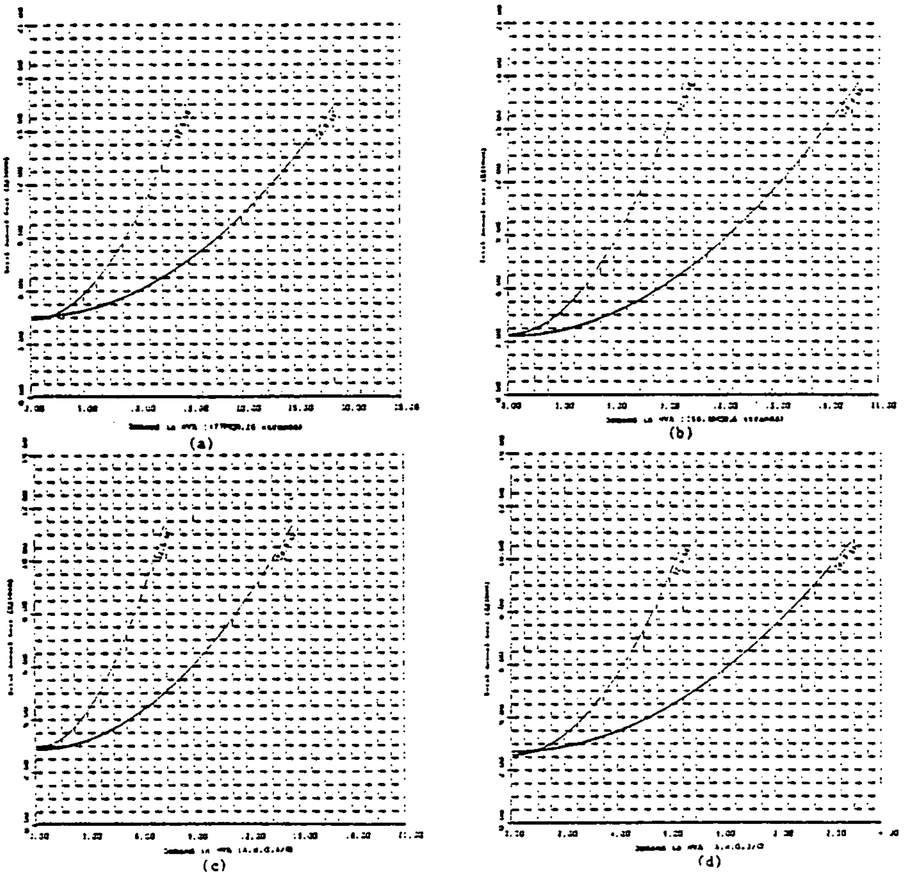


Figure D.42. Total annual equivalent cost of ACSR feeders for rural areas in thousand dollars per mile for: (a) 477 CM, 26 strands, (b) 266,3 CM, 6 strands, (c) A.W.G.4/0, (d) A.W.G.3/0

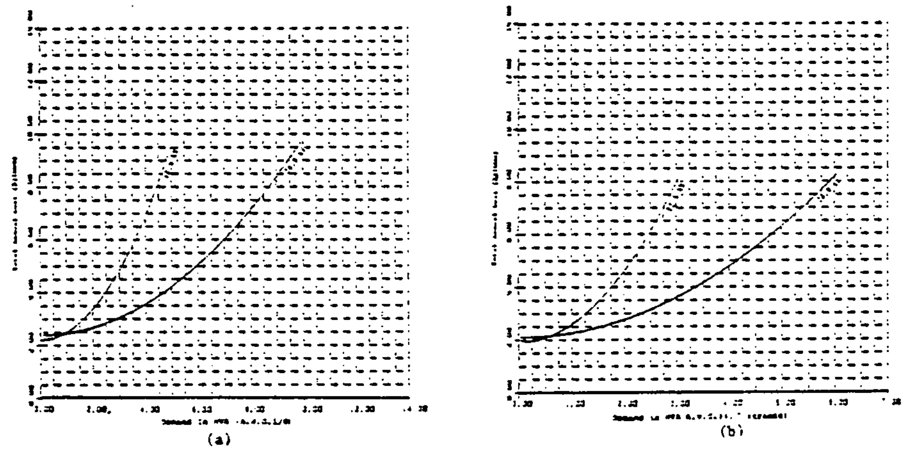


Figure D.43. Total annual equivalent cost of ACSR feeders for rural areas in thousand dollars per mile for: (a) A.W.G.1/0 (b) A.W.G.#4, 7 strands

APPENDIX E

MULTI-OBJECTIVE EVALUATION

A multi-criteria evaluation approach that has the advantage of simplicity in application is the Churchman - Akoff Approach.* The approach is simply the following:

1) Determine a set of independent, mutually exclusive objectives. For example, minimize costs and increasing reliability are two objectives which meet the criteria. Increasing reliability and increasing redundancy do not meet the criteria of independent and mutually exclusive because increasing redundancy will increase reliability. It may be possible, however, to reduce cost without affecting reliability by choice of demand centers served by substations, etc.
2) Determine weights for the objectives. In the Churchman - Akoff method this is done by a decision tree approach. The basic inputs are subjective quantifications by decision makers.
3) Determining the alternatives.
4) Find an appropriate measure of attainment of an objective.
5) Finding the efficiency of each alternative relative to each objective by using the measures defined in step 4.
6) Computing

$$E_j = \sum_{j=1}^{n} O_i \times e_{ij}$$

where

E_j = efficiency of the jth alternative,
O_i = weight of the ith objective,
e_{ij} = efficiency of the jth alternative relative to the ith objective.

As an example, consider a session with Mr. Lester Burris of OG & E, to demonstrate the approach. Mr. Burris named four objectives:

O_1 Minimize cost,
O_2 Meet voltage requirements for several years in the future,
O_3 Reliability,
O_4 Meet the esthetic demands of the environment.

The objectives were ranked in importance as follows:

voltage requirements,
cost,
esthetics,
reliability.

These objectives were then renumbered to correspond with the ranking. The measures for each objective were determined as follows:

Cost: Annual equivalent cost of plan at interest rate i, labor, materials, maintenance, taxes, interest, energy losses, etc.
Reliability: Projected load-hours lost per year.
Esthetics: An ordinal scale ranking with underground = 3, aluminum poles = 3, route change = 1 and regular poles = 0.
Voltage requirements: Number of years voltage requirement is met.

To measure efficiency, since none of these objectives have a well-defined efficiency as determined by the measure, efficiency curves were used.

For cost 100% efficiency would be zero, by definition. Zero efficiency would be an infinity cost. Two possible curves are concave and convex, as are in Figures E.1 and E.2, respectively.

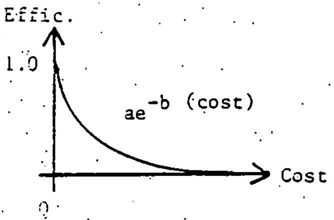


Figure E.1. Concave

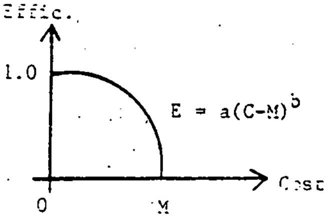


Figure E.2. Convex

In this case the convex curve was preferred. Parameters can be obtained by specifying a point (c,E) in addition to (0,1), and (M,0).

The curves shown in Figures E.3 and E.4 were proposed for reliability and esthetics, respectively. The voltage requirements efficiency curve was given by Figure E.5.

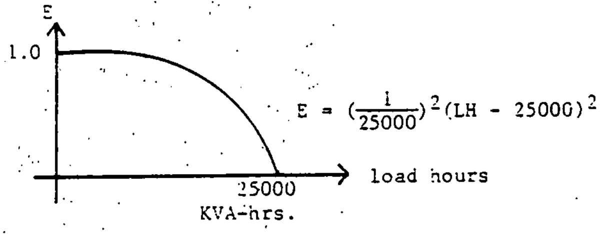


Figure E.3. Efficiency curve for reliability

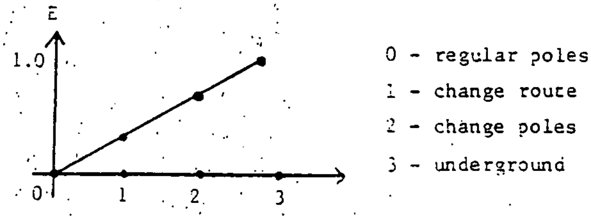


Figure E.4. Efficiency curve for esthetics

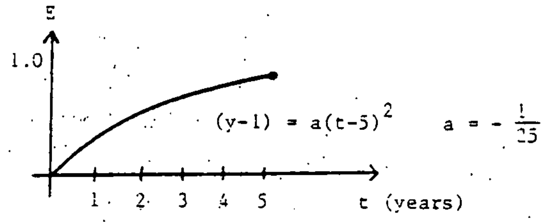


Figure E.5. Voltage requirements efficiency curve

Next weights W_i for each objective O_i were determined by posing the following type questions: Do you prefer (yes or no) O_1 to O_2, O_3, O_4 which means if O_1 is obtained at its maximum possible efficiency and O_2, O_3, O_4 were at their lowest acceptable levels of efficiency? The answer is yes then $W_1 > W_2 + W_3 + W_4$. Otherwise $W_1 < W_2 + W_3 + W_4$. If $W_1 > W_2 + W_3 + W_4$ then certainly $W_1 > W_2 + W_3$. If $W_1 < W_2 + W_3 + W_4$ then W_1 could be such that $W_1 > W_2 + W_3$ so O_1 vs. O_2, O_3 must be asked. The resultant decision tree is given by Figure E.6.

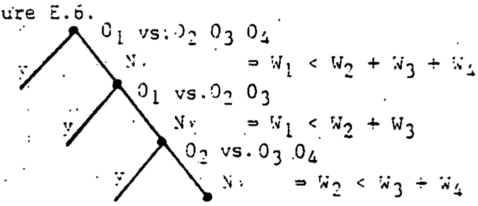


Figure E.6. The resultant decision tree

---

*Churchman, C.W., Akoff, R.L., and Arnoff, E.L., Introduction to Operations Research, John Wiley & Sons, Inc., New York, N.Y., 1957.

A first assignment of weight was:

voltage	100
cost	35
Esthetics	70
Reliability	69

These weights satisfied the constraints on the weights given by the decision tree questions. If the weights would not have satisfied the inequalities, then new weights would have been chosen by the decision maker in order to satisfy the inequailties violated. The weights were normalized to obtain more consistency and therefore:

$$\begin{aligned} W_1 &= 0.31 \\ W_2 &= 0.26 \\ W_3 &= 0.22 \\ W_4 &= \underline{0.21} \\ & \quad 1.0 \end{aligned}$$

The objective function is then

$E_j = .31e_{1j} + .26e_{2j} + .22e_{3j} + .21e_{4j}$

The e_{ij} are determined by evaluating the plan with the respective measures and then using the curves developed to find the efficiencies.

A further elaboration of the efficiency curve is now appropriate. As an example, consider a source with three load centers as shown in Figure E.7.

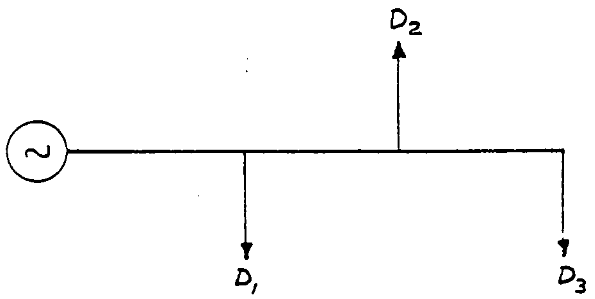


Figure E.7. A source with three load centers

It is desired that the voltage at each load center, i.e., D_i, be 121 volts plus or minus 6 volts. A measure of success in meeting this requirement might be

$$M = \infty \quad V_i \notin (115,127) \text{ for any } i$$

$$M = \sum_{i=1}^{3} (V_i - 121)^2$$

where

V_i = actual voltage at the load center D_i. Our voltage requirement goal is $V_i = 121$ Volt for each i. Thus if a circuit design is 100 percent <u>effective</u>,

$$M = (121 - 121)^2 + (121 - 121)^2 + (121 - 121)^2 = 0$$

If a voltage V_i is at the endpoint of (115,127) for each i, then further deviation will render $M = \infty$, thus giving the largest acceptable value of M that can be tolerated. That is

$$M = 6^2 + 6^2 + 6^2 = 108$$

Hence, for M = 108, the design is at its minimum acceptable efficiency level. Thus, two points, (100,0) and (0,108), are established.

What if a circuit design gives an M value of 34? What is the relative value of this plan, compared to a plan with M = 68? It is here the concept of an efficiency (effectiveness) curve comes in. The following can be postulated.

1) Effectiveness is measured by values from zero to one hundred.
2) E = F(M) is a concave function, as shown in Figure E.8.

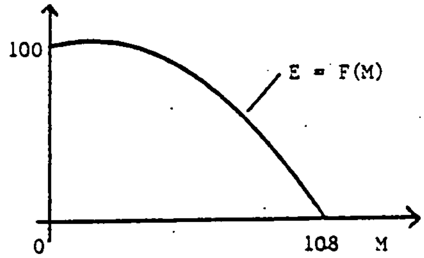


Figure E.8. E = F(M) concave function

This arises from the fact that subjectively values of M close to zero are really 100% effective also. Values of M close to 108 are very close to zero in effectiveness. Thus, we want a flat curve near M equal to zero and a vertical slope near M = 108. A parabolic curve fits these needs. This leads to a third postulate. That is

3) The relative effectiveness of plans is computed by

$$E - 1 = -\frac{1}{(108)^2}(M)^2$$

If the decision maker can define a third point, for example, the point $(50, M_{50})$ then $E = aM^2 + bM + c$ can be fitted to the data.

From this analysis the computation of effectiveness is based on axiomatic considerations. It is arbitrary, but so is the typical definition of efficiency

$$\text{efficiency} = \frac{\text{output}}{\text{input}} \times 100$$

In these terms, the effectiveness or efficiency diagram looks like the one shown in Figure E.9.

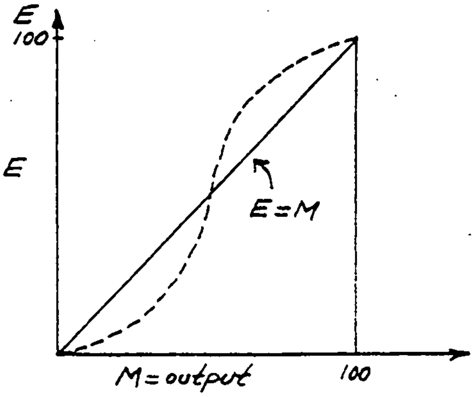


Figure E.9. Effectiveness or efficiency diagram

The measure M is the output measured in terms of percent of input. Of course, even in this case, it is possible to use a convex or concave curve. In fact, an S-shaped curve might be more appropriate for rotating machinary.

A second illustration is the result of an interview with another experienced distribution engineer, Mr. Tom Littleton, of Oklahoma Gas & Electric Company. The purpose of the session was to test the general evaluation procedure on a test distribution planning case. This was a large scale distribution planning case in the Arkansas River Basin area.

On this problem after lengthy discussion, protection and voltage requirements became the most important constraints. Further, it was concluded that any plan which clears a fault in three seconds is acceptable and also a plan which clears in one second is no better than a plan which clears in three second. Furthermore, all customer voltage must be within the given limits. A plan which has all customers at the center of the limit values has no real preference over a plan which has substantial variation but no limit is violated. The following three objectives were finally developed:

O_1: Minimize present worth of 30 years cost.
O_2: Minimize initial capital investment.
O_3: Maximize the reliability of the plan.

No formal reliability calculation or projection was made although it was formally recognized that one plan decreased reliability by reducing the number of substations and lengthening the line. Reliability is measured by

$$\frac{\text{total load hours demanded - load hours of outrage}}{\text{total load hours demanded}}$$

Since the study was not focused towards this method, some of the data was estimated. However, a large amount of summary data was available which put the estimated on a solid ground. The weights of the objectives were:

O_1: 0.39
O_2: 0.31
O_3: 0.29

The efficiency curves were developed, as shown in Figures E.10, E.11 and E.12.

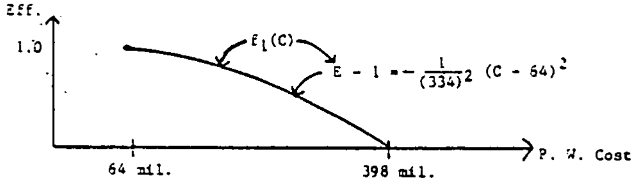


Figure E.10. The developed efficiency curve: efficiency vs. present worth of costs

Sixty-four million was the cost of the cheapest plan. 398 million was the present worth of revenue projected from the project and thus the maximum allowable expenditure.

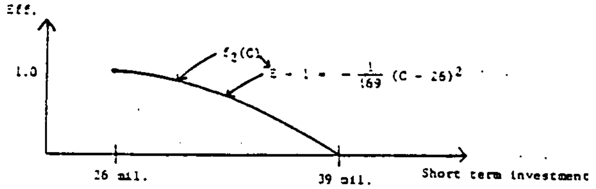


Figure E.11 The developed efficiency curve: efficiency vs. short term investment costs.

Twenty-six million was the smallest initial investment of the plans and 39 million was based on the importance of the project being enough to use up to 70 percent of the 55 million in capital which OG & E projected they could raise for expansion.

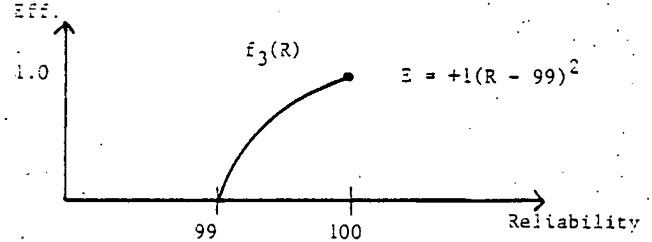


Figure E. 12. The developed efficiency curve: efficiency vs. short term investment costs.

Ninety-nine percent was a subjective estimate of the lowest ratio acceptable for reliability. Therefore, given the following data for the alternatives.

Plan	A	B
O_1	68	64
O_2	29	26
O_3	99.98	99.95

The results of the evaluation of the alternatives are

$$\begin{aligned} E_A &= .39f_1(68) + .31_2(29) + .29f_3(99.98) \\ &= .39(.9999) + .31(.95) + .29(.9996) \\ &\cong .9745 \end{aligned}$$

$$\begin{aligned} E_B &= .39f_1(64) + .31f_2(26) + .29f_3(99.95) \\ &= .39(1.0) + .31(1.0) + .29(9978) \\ &\cong .989 \end{aligned}$$

The results show that the plan B is better and was the chosen plan. However, the results of the evaluation are so close that a computation of reliability that was exact might reverse the results. It was, however, felt that the estimate of reliability was conservative in favor of Plan A. The approach then did predict the decision in an actual case.

REFERENCES FOR APPENDIX E

[1] Churchman, C.W., Ackoff, R.L., and Arnoff, E.L., Introduction to Operations Research, John Wiley and Sons, Inc., New York, N.Y., 1957.

[2] Saaty, L.L., "A Scaling Method for Priorities in Hierarchical Structures", Journal of Mathematical Psychology, Vol.15, June 1977.

[3] Bammi, Deepak and Bammi, Dalip, "Development of a Comprehensive Land Use Plan by Means of a Multiple Objective Mathematical Programming Model", Interfaces, Vol. 9. 2, part 2, Feb. 1979, pp. 50-64.

[4] Crawford, D.M., Huntzinger, B.C. and Kirkwood, C.W., "Multiobjective Decision Analysis for Transmission Conductor Selection", Management Science, Vol. 24, 16, Dec. 1978, pp. 1700-1710.

APPENDIX F: ELECTRIC ENERGY DISTRIBUTION PLANNING DATA BASE MANAGEMENT SYSTEM - PROBLEM SPECIFICATION

SECTION 1: SCOPE

This document is a functional specification of the prototype system being developed as part of the Electric Energy Distribution Planning Project. The specifications are not phrased in terms of specific programs. This will be done in the design phase of the Project (see the Project Plan, P01 for an overview of the project organization). The specifications are in terms of system capabilities and as such, define the system to which the design must lead.

This is a baseline document and its most recent edition is to be strictly adhered to by all Project personnel.

SECTION 2: APPLICABLE DOCUMENTS

(Reference codes specify documents as they are found in the Data Base Group library).

Ref. Code	Author	Short Title
R20	Date, C.J.	Intro. Database Systems
P01	--	Project Plan
T01	Thompson, J.C.	Advantages of Relational Model
A03	Gonen, T.	Task Summary
R19	Kernighan & Plouger	Software Tools
T03	Cubert, R.M.	"Network Editor"
R21	Fry & Sibley	Evolution of Database Management Systems

SECTION 3: REQUIREMENTS

The system to be designed and implemented consists of the following components:

- Database Model
- Prototype Database
- Database Management System
- Network Editor
- Shell Program
- Distribution Planning Programs
- Text Editor

The programs (the last five of the above) are to perform in an interactive environment in which the human engineer has direct control over the major subtasks in the planning process. The only program components of the system which are not constrained to have response times on the order of seconds are the Distribution Planning Programs. Their performance characteristics are not subject to further modification by the implementation effort associated with this project.

Because of their interactive behavior, the portions of the Problem Specification dealing with the programs will be expressed in terms of what are called protocols. These are expressions written in Backus-Naur form which specify the functions which each program is to possess. No specification below the protocol level will be presented. All further design will be found in the Design Specification. It should be noted that the protocols are not descriptions of the syntax of the interactive commands. They are the functional specifications for the commands. Syntax will be specified as part of the design in the design document.

3.1 Database

The database is the total collection of data that all of the operational programs require in order to execute. In principle, the database(s) also includes data that conceivably might be required at a later point in time by more sophisticated versions of the system.

The real issue however, is not so much what data constitute the database but whether the data model (data organization) is capable of supporting the system. It has been recognized for some time (cf. R21) that there are three distinct and well understood database organizations or data models. There are the hierarchical model, the network model and the relational model.

For reasons not dealt with here (see T01), the relational model has been selected as the type system to be implemented for this project. This choice relegates the problem of actual database implementation (as opposed to database management system implementation) to the status of a simple exercise. In fact, we may say that building the database only involves selecting the schema, i.e., dividing the relevant data into relations and inputing the actual data values.

In order to do this latter task, the required data must be known. Determining which data should be included is the responsibility of task members other than those in the Data Base Group (see A03). Once that is done and the data relations defined, the task of defining the conceptual database is completed.

3.1.1 Relational Database Model

A complete discussion of relational database models will be found in the literature cited in Section 2. Here a few terms will be defined for future reference. Let D_1, D_2, ..., D_n be a collection of (not necessarily distinct) sets. A relation R, on these sets is an ordered n-tuple (d_1, d_2, ..., d_n) where $d_i \in D_i$, for 1,2, ..., n. The sets D_i are the domains of R. The value of n is the degree of R. In a relation R, the ith attribute is associated with D_i and the attribute values for a given attribute are those values forming a subset of D_i in R. The domains, the attributes and their relative ordering in the tuples of a given relation are defined by the schema for the relation.

3.2 Database Management System

Since the early work of Codd, (see R20 for an extensive bibliography) there have been two well known descriptions or languages for manipulating elements of a relational database. One is known as the relational calculus and the other as the relational algebra. A full discussion of either of these is beyond the scope of the present discussion (see T01). Suffice it to say that the relational algebra is expressed primarily in terms familiar from set theory and is the more procedural of the two. It also appears the simpler of the two to implement. For these reasons, the relational algebra approach was adopted for the present system.

In the following sections performance parameters, functional specifications and human performance requirements are discussed.

3.2.1 Performance Parameters

The database management system will be designed to operate on databases of up to 100 million characters in size. Average maximum response time will be limited to five seconds. This number represents the time it takes to do a single operation on the database when running as a stand-alone system. Because response characteristics of different operating systems vary so much from situation to situation, this response time may not be realizable in all interactive environments. However, this is the value for which the system will be engineered.

3.2.2 Operational Requirements

The operational requirements for the database management system are listed below in the form of protocols. An index of terms appearing in the protocols will be found in Section 4.

3.2.2.1 ⟨select⟩ function

Functional specification.
⟨select⟩ ⟨input relation name⟩ ⟨qualification⟩
⟨output relation name⟩
⟨qualification⟩ ::= ⟨tuple relation⟩ ⟨qualification⟩
| ⟨tuple relation⟩
⟨tuple relation⟩ ::= ⟨attribute volume⟩⟨boolean operator⟩ ⟨Tuple funtion⟩*
Semantics:
The subrelation of the input relation specified by qualification is assigned to the output relation.

3.2.2.2 ⟨join⟩ function

Functional specifications.
⟨join⟩⟨input relation name$_1$⟩⟨input relation name$_2$⟩
⟨bi-tuple boolean relation⟩⟨output relation name⟩
⟨bi-tuple boolean relation⟩ ::= ⟨attribute value$_1$⟩
⟨boolean operator⟩
⟨attribute value$_2$⟩
Semantics:
The ⟨attribute value$_1$⟩'s are taken from tuples of the relation specified by ⟨input relation name$_1$⟩ while the ⟨attribute value$_2$⟩'s are taken from tuples of the relation specified by ⟨input relation name$_2$⟩. All such tuples which satisfy the ⟨bi-typle boolean relation⟩ are conccatenated together pairwise to form the relation referenced by ⟨output relation name⟩.

3.2.2.3 ⟨project⟩ function

Functional specifications.
⟨project⟩ ⟨input relation name⟩ ⟨attribute name list⟩
⟨output relation name⟩
⟨attribute name list⟩ ::= ⟨attribute name⟩
⟨attribute name list⟩
| ⟨attribute name⟩
Semantics:
Attributes of ⟨input relation⟩ included in ⟨attribute name list⟩ are used to form a new relation referenced by ⟨output relation name⟩. Only one instance of identical tuples is retained.

3.2.2.4 ⟨divide⟩ function

Functional specifications.
⟨divide⟩ ⟨input dividend relation name⟩
⟨input unary divisor relation name⟩
⟨output unary quotient relation name⟩
Semantics:
Indicate the ⟨divide⟩ operation with the following expression $Q = X \div Y$. Then $q \in Q$ implies $(q, y_i) \in X$ for all y_i Y.

3.2.2.5 ⟨minus⟩ function

Functional specifications.
⟨minus⟩ ⟨input subtrahend relation name⟩
⟨input minuend relation name⟩
⟨output result relation name⟩
Semantics:
The relations specified by ⟨input subtrahend relation name⟩ and ⟨input minuend relation name⟩ must be union- compatible, i.e., must be of the same degree n and the jth attribute of one must be drawn from the same domain as the other $1 \leq j \leq n$. The relation specified by ⟨output result relation name⟩ consists of those tuples of the ⟨subtrahend relation⟩ which are not in the ⟨minuend relation⟩.

3.2.2.6 ⟨union⟩ function

Functional specifications.
⟨union⟩ ⟨input relation name$_1$⟩ ⟨input relation name$_2$⟩
⟨output relation name⟩
Semantics:
The input relations must be union-compatible. The ⟨output relation⟩ is the relation containing tuples which appear in either ⟨input relation$_1$⟩or⟨input relation$_2$⟩ (or both).

3.2.2.7 ⟨intersect⟩ function

Functional specifications.
⟨intersect⟩⟨input relation name$_1$⟩ ⟨input relation name$_2$⟩
⟨output relation name⟩
Semantics:
The input relations must be union-compatible. The relation output relation consists only of those tuples contained in both the relations ⟨input relations$_1$⟩ and ⟨input relations$_2$⟩.

3.2.2.8 ⟨multiply⟩ function

Funcitional specifications.
⟨multiply⟩ ⟨input relation name$_1$⟩ ⟨input relation name$_2$⟩
⟨output relation name⟩
Semantics:
The relation ⟨output relation⟩ consists of all tuples which can be formed by concatenating tuples of ⟨input relation$_1$⟩ to those of ⟨input relation$_2$⟩ in that order.

3.2.2.9 ⟨invoke database management system⟩ function

Functional specifications.
⟨invoke database management system⟩ ⟨database name⟩
⟨user_id⟩
Semantics:
The database management system is invoked to perform transactions against the database specified by ⟨database name⟩. The ⟨user_id⟩ must be supplied. If no database by this name exists, one will be created with its database administrator (DBA) corresponding to the ⟨user_id⟩.

3.2.2.10 ⟨destroy database⟩ function

Functional specifications.
⟨destroy database⟩ ⟨database name⟩ ⟨user_id⟩
Semantics:
The database specified by database name is removed from the system, provided the ⟨user_id⟩ is identical to that of the database administrator. Otherwise, an error results.

3.2.2.11 ⟨create relation⟩ function

Functional specifications.
⟨create relation⟩ ⟨relation name⟩ ⟨schmea specification⟩
⟨access method⟩ ⟨retension status⟩
⟨permission⟩
⟨schema specification⟩ ::= ⟨attribute list⟩ ⟨key list⟩
⟨attribute list⟩ ::= ⟨attribute name⟩ ⟨domain specification⟩ ⟨attribute list⟩
| ⟨attribute name⟩ ⟨domain specification⟩
⟨key list⟩ ::= ⟨attribute name⟩⟨key list⟩
| ⟨attribute name⟩
⟨domain specification⟩ ::= ⟨int⟩ | ⟨float⟩ | ⟨char⟩

---

*It is understood that the tuple function depends on attribute values drawn from the same tuple as the attribute value being compared.

⟨field length⟩ ⟨retension status⟩ ::= ⟨temporary⟩|
⟨retension date⟩ ⟨permission⟩ ::= ⟨write⟩|⟨execute⟩
Semantics:

A relation with the name ⟨relation name⟩is created. The schema is specified by means of the ⟨schema specification⟩ . This command is illegal unless the user has ⟨write⟩ or ⟨execute⟩permission (or is the DBA).

3.2.2.12 ⟨destroy relation⟩function

Functional specifications
⟨destroy relation⟩ ⟨relation name⟩ ⟨permission⟩
⟨permission⟩ ::= ⟨execute⟩
Semantics:

The relation specified by ⟨relation name⟩ is removed from the database. For relations whose ⟨retension status⟩ is not equal to ⟨temporary⟩ , this command is illegal unless the user's ⟨permission⟩ equals ⟨execute⟩.

3.2.2.13 ⟨copy⟩ function

Functional specifications.
⟨copy⟩ ⟨external database name⟩ ⟨external relation name⟩
⟨local relation name⟩ ⟨access method⟩
⟨retension status⟩ ⟨external permission⟩
⟨external permission⟩ ::= ⟨execute⟩
Semantics:

A relation specified by ⟨external relation name⟩ residing in a database specified by ⟨external database name⟩ is copied into the database currently under the control of the user. ⟨execute⟩is the only legal value for the user's ⟨permission⟩ relative to the ⟨external database⟩.

3.2.2.14 ⟨modify⟩ .function

Functional specifications.
⟨modify⟩ ⟨input relation name⟩ ⟨new access method⟩
⟨output relation name⟩ ⟨permission⟩
⟨permission⟩ ::= ⟨execute⟩
Semantics:

The relation specified by ⟨input relation name⟩ is copied into a new relation organized according to the ⟨access method⟩ specified. The new relation is referenced by ⟨output relation name⟩.

3.2.2.15 ⟨output relation⟩ function

Functional specifications
⟨output relation⟩ ⟨relation name⟩ ⟨format specification⟩
Semantics:

This command coerces the relation referenced by ⟨relation name⟩ into a form suitable for output on the system output file. The ⟨format specification⟩ optionally controls conversion from one set of units to another.

3.2.2.16 ⟨output schema⟩ function

Functional specifications.
⟨output schema⟩ ⟨relation name⟩ ⟨user_id⟩
Semantics:

⟨output schema⟩converts the schema associated with the pair (⟨relation name⟩, ⟨user_id⟩) into a form suitable for output on the system output file.

3.2.2.17 ⟨input relation⟩ function

Functional specifications.
⟨input relation⟩ ⟨relation name⟩ ⟨format specification⟩
Semantics:

Data for a predefined relation is read from the standard input file under format control. The format specification optionally controls conversion from one set of units to another.

3.2.2.18 ⟨delegate⟩function

Functional specifications.
⟨delegate⟩ ⟨relation name list⟩ ⟨access list⟩ ⟨user⟩
⟨relation name list⟩ ::= ⟨relation name⟩ ⟨relation name list⟩
| ⟨relation name⟩

⟨access list⟩ ::= ⟨schema specification⟩ ⟨permission⟩
⟨user_id list⟩ ⟨access list⟩
|⟨schema specification⟩ ⟨permission⟩
⟨user_id list⟩
⟨permission⟩ ::= ⟨read⟩|⟨write⟩|⟨execute⟩
⟨user_id list⟩::= ⟨user_id⟩ ⟨user_id list⟩
| ⟨user group id⟩ ⟨user_id list⟩
| ⟨user_id⟩|⟨user_group_id⟩
⟨user⟩ ::=⟨DBA⟩
Semantics:

⟨delegate⟩ is used by the DBA to delegate certain priveledges to other users of the database. The permission ⟨read⟩ permits the user to form queries only. ⟨write⟩ permission allows a user to form queries and to ⟨define⟩ and⟨save⟩ new relations. ⟨execute⟩ permission, in addition to priveledges associated with the above, permits the user to ⟨update⟩relations. Note however, that even a user with ⟨execute⟩ permission is not the DBA since the DBA alone has the power to ⟨purge⟩ relations and ⟨destroy⟩ databases.

A ⟨user_group_id⟩ is a designation that applies to a group of users. ⟨user_group_id⟩'s are defined by the DBA in the system database.

3.2.2.19 ⟨save⟩ function

Functional specifications.
⟨save⟩ ⟨relation name⟩ ⟨retension status⟩ ⟨permission⟩
⟨retension status⟩ ::=⟨temporary⟩ ⟨retension date⟩
⟨permission⟩ ::= ⟨write⟩|⟨execute⟩
Semantics:

The relation specified by ⟨relation name⟩ is to be retained in the database for the period of time specified by ⟨retension status⟩ of⟨temporary⟩ at their time of creation unless this default value is overridden. Relations with ⟨temporary⟩ status are destroyed when the user's interactive session terminates.

3.2.2.20 ⟨purge⟩ function

Functional specifications.
⟨purge⟩ ⟨user_id⟩
⟨user_id⟩ ::= ⟨DBA⟩
Semantics:

Execution of ⟨purge⟩removes all relations from the database whose ⟨retension status⟩ has matured.

3.2.2.21 ⟨assign⟩function

Functional specifications.
⟨assign⟩ ⟨input relation name⟩ ⟨assignment list⟩
⟨output relation name⟩ ⟨permission⟩
⟨assignment list⟩ ::= ⟨attribute name⟩ ⟨tuple function⟩
⟨assignment list⟩
| ⟨attribute name⟩ ⟨tuple function⟩
⟨permission⟩ ::= ⟨write⟩|⟨execute⟩
Semantics:

The relation specified by ⟨output relation name⟩ is formed by assigning to attribute values, the results computed according to the ⟨tuple function⟩'s. The assignment is made for every tuple in the relation referenced by ⟨input relation name⟩.

For example, suppose a relation salary is defined by the schema

salary {employee#, pay} KEY (employee#)

and that each employee is to receive a 10% raise. Then ⟨assign⟩ could be used as follows:

⟨assign⟩ salary "pay := (1.1*pay)" new salary.

Attribute values not mentioned in the ⟨assign⟩ statement are copied unchanged to the output relation.

3.2.2.22 ⟨update⟩ function

Functional specifications.
⟨update⟩ ⟨input old relation name⟩ ⟨input modifying relation name⟩ ⟨output new relation name⟩ ⟨permission⟩
⟨permission⟩ ::= ⟨execute⟩
Semantics:

The relations referenced by ⟨input old relation name⟩ (the old relation), and ⟨input modifying relation name⟩ (the modifier), must be union-compatible. Tuples are compared on a 1-1 basis and those in the old relation with tuples whose keys match those of the modifier are replaced by the corresponding tuples in the modifier to form the new relation. The new relation has the same degree and cardinality as the old relation.

3.2.2.23 ⟨restore⟩function

Functional specifications.
⟨restore⟩ ⟨database name⟩ ⟨edition identifier⟩
⟨transaction file identifier⟩ ⟨user_id⟩
⟨user_id⟩ ::=⟨DBA⟩
Semantics:

This function is used to restore the database to current status after a system failure. The transaction file contains a list of all permanent changes to a specific edition of the database over some period of time.

3.2.2.24 ⟨dump⟩ function

Functional specifications.
⟨dump⟩ ⟨database name⟩ ⟨edition identifier⟩ ⟨output file name⟩ ⟨user_id⟩
⟨user_id⟩ ::=⟨DBA⟩
Semantics:

Execution of this command creates a backup copy of the database referenced by ⟨database name⟩.

3.2.2.25 ⟨help⟩ function

Functional specifications.
⟨help⟩ ⟨command name⟩ ⟨description qualifier⟩
Semantics:

⟨help⟩ supplies information about system capabilities upon user request. The description will be more or less detailed according to the ⟨description qualifier⟩.

3.2.3 Human Performance

The human support for the database management system is minimal. It consists primarily of human-directed management of the system files on which the dbms depends for such things as ⟨user_id⟩'s and ⟨permission⟩ assignments made to users.

Requirements on the user are not onerous either. If the basic ideas behind a relational database are well understood and if the user has a minimum familiarity with relational algebra, the system should be straightforward to use.

3.2.4 Language Specifications

The syntax for the relational algebra-base queries will not be specified in this document. The spirit which is to be infused into this language can be identified as the same spirit manifest in the command languages exhibited in the Software Tools (R19). It remains to be seen however, precisely what form this will impart to the query language.

3.3 Network Editor

The network editor is an interactive subsystem which will allow the planning engineer to create and modify network models for processing by other planning programs. From a human engineering point of view, it is much more satisfactory to allow the engineer to work directly with network concepts than to effect the same process by using the database management system. A supplementary discussion of the network editor will be found in T03.

3.3.1 Performance Parameters

Networks are stored in the system as relations in a network database. Generally, therefore, the network editor must meet the same performance goals that the database management system must meet.

3.3.2 Network Objects

Network objects are described here to the same level of detail that was supplied for the database management system.

3.3.2.1 ⟨network⟩

⟨network⟩ ::= ⟨object list⟩
⟨object list⟩ ::= ⟨object⟩ ⟨object list⟩| ⟨object⟩
⟨object⟩ ::= ⟨primitive object⟩|⟨composite object⟩

A network is a collection of objects or components which may be either primitive (atomic or simple) or composite, i.e., composed of a collection of simple or composite elements.

3.3.2.2 ⟨primitive object⟩

⟨primitive object⟩ ::= ⟨name⟩ ⟨object class⟩ ⟨connection list⟩
⟨object class⟩ ::= ⟨description⟩ ⟨input port list⟩ ⟨output port list⟩
⟨description⟩ ::= ⟨parameter list⟩ ⟨qualitative description⟩
⟨parameter list⟩ ::= ⟨parameter⟩ ⟨parameter list⟩ |⟨parameter⟩| ⟨NULL⟩
⟨parameter⟩ ::=⟨constant⟩|⟨variable⟩|⟨numerical state variable⟩ |⟨boolean state variable⟩
⟨input port list⟩ ::= ⟨port⟩ ⟨input port list⟩|⟨port⟩
⟨output port list⟩ ::= ⟨port⟩ ⟨output port list⟩ | ⟨port⟩
⟨connection list⟩ ::= ⟨list for each port⟩ ⟨connection list⟩|⟨list for each port⟩
⟨list for each port⟩ ::= ⟨list⟩ | ⟨OPEN⟩
⟨list⟩ ::= ⟨connection⟩ ⟨list⟩ | ⟨connection⟩

Primitive objects consist of a name, an object class and a list of connections to other objects. An object class consists of a parametric description e.g., values of capacitance for a capacitor, qualitative information such as the device class name, and a list of input ports and output ports.

The parametric description is general enough to permit the parameters to be variables as well as constants. The variables are of two types. One type, called state variables, are global to the entire network and may assume different values resulting in different network configurations. This is especially true for network components such as switches which may be opened or closed according to the value for the boolean state variable which defines their state.*

The second type of variable, found in both primitive and composite object, to be discussed below, is analogous in some respects to a subroutine in that the composite object may have formal variables, whose values are determined at instantiation time. Parameters of primitive objects denoted as variables assume the values assigned to the formal variables of the composite object in which they are embedded.

3.3.2.3 ⟨composite object⟩

---

*cf. the discussion of switches in T03.

⟨composite object⟩::=⟨name⟩ ⟨composite object class⟩
⟨connection list⟩
⟨composite object class⟩::=⟨composite description⟩ ⟨input port list⟩ ⟨output port list⟩
⟨composite description⟩::=⟨parameter list⟩ ⟨qualitative description⟩ ⟨network⟩

Composite objects are described in the same way that primitive objects are except that a description of the (sub-) network connecting their constituient parts must be included. Parts defined in object class 's of subnetwork components but for which connection lists are unspecified are identified in the ⟨composite object class⟩ and are connected via the ⟨connection list⟩ of the ⟨composite object⟩.

3.3.3 Operational Requirements

In this section, the funcitonal specifications for the network editor will be described.

3.3.3.1 ⟨invoke network editor⟩

Functional specifications.
⟨invoke network editor⟩ ⟨network name⟩ ⟨permission⟩
⟨permission⟩ ::=⟨write⟩ | ⟨execute⟩
Semantics:

This command activates the network editor. The user must have either ⟨write⟩or ⟨execute⟩ permission to invoke the editor on a given network. ⟨write⟩ permission allows the user to create and save networks, ⟨execute⟩ permission is required to modify a network already created.

3.3.3.2 ⟨create object class⟩

Functional specifications.

⟨create object class⟩ ⟨input object class description⟩ ⟨output object class⟩
Semantics:

This command establishes the paradigm for the network component. No specific object is created by this command (since no cursor is returned to point to a specific object). This function is used to create new primitive ⟨object class⟩'s not included in the system vocabulary and for building composite ⟨object class⟩'s.

3.3.3.3 ⟨create object⟩

Functional specifications.
⟨create object⟩ ⟨input object class⟩ ⟨input name⟩ ⟨input parameter specification⟩ ⟨output object cursor⟩
Semantics:

A network object is created. A cursor (pointer) is returned which can be used to connect the object to the network,

3.3.3.4 ⟨discard object⟩

Functional specifications.
⟨discard object⟩ ⟨input object cursor⟩
Semantics:

⟨discard⟩ eliminates an object from the system. An object which is contained in a network cannot be ⟨discard⟩ed. It must first be removed from the network (using ⟨remove⟩).

3.3.3.5 ⟨add object⟩

Functional specifications.
⟨add object⟩ ⟨input object cursor⟩ ⟨input network cursor⟩ ⟨input port-port descriptor⟩ ⟨output network cursor⟩
⟨port-port descriptor⟩::=⟨object port name⟩
⟨network object port name⟩
Semantics:

An object not initially attached to the network is connected to the network via a connection from the object to an object already included in the network. Objects already included in the network cannot be ⟨add⟩ed.

3.3.3.6 ⟨remove object⟩

Functional specifications.
⟨remove object⟩ ⟨input network cursor⟩ ⟨output object cursor⟩ ⟨output network cursor⟩
Semantics:

An object contained within the network is removed from the network. The object still exists (there is a pointer to it) but it is no longer included in the network. The operation returns an error if the object is not initially in the network.

3.3.3.7 ⟨connect network objects⟩

Functional specifications.
⟨connect network objects⟩ ⟨input network cursor$_1$⟩ ⟨input network cursor$_2$⟩ ⟨input port$_1$-port$_2$ descriptor⟩ ⟨output network cursor⟩
⟨port$_1$-port$_2$ descriptor⟩ ::=⟨network object$_1$ port name⟩
⟨network object$_2$ port name⟩
Semantics:

Two network objects are connected together. Both must already be included in the network and must have compatible port characteristics.

3.3.3.8 ⟨disconnect network objects⟩

Functional specifications.
⟨disconnect network objects⟩ ⟨input network cursor$_1$⟩ ⟨input network cursor$_2$⟩ ⟨input port$_1$-port$_2$ descriptor⟩ ⟨output netwrok cursor⟩
Semantics:

A connection between two network objects is broken. This command is illigal if it breaks the only connection between an object and the network.

3.3.3.9 ⟨set state variables⟩

Functional specifications.
⟨set state variables⟩ ⟨input parameter assignment list⟩
⟨parameter assignment list⟩::=⟨parameter assignment list⟩
| ⟨parameter assignment⟩
⟨parameter assignment⟩ ::=⟨parameter name⟩ ⟨parameter expression⟩
Semantics:

State variables are assigned values with this command. A state variable is assigned the value associated with the expression comprising parameter expression . All variables appearing in a parameter expression must have previously been defined.

3.3.3.10 ⟨find object⟩

Functional specifications.
⟨find object⟩ ⟨input network cursor⟩ ⟨input context description⟩ ⟨output network cursor⟩
Semantics:

This function is analogous to the context search command found in high performance text editors. The ⟨context description⟩ will probably be defined using regular expressions over the object class alphabet. The search will be done in a specified traversal order.

3.3.3.11 ⟨save ...⟩

Functional specifications.
⟨save network⟩ ⟨network name⟩ ⟨permission⟩
⟨permission⟩::= ⟨write⟩ | ⟨execute⟩
⟨save objects⟩ ⟨object cursor list⟩ ⟨permission⟩
⟨object cursor list⟩ ::=⟨object cursor⟩ ⟨object cursor list⟩ ⟨object cursor⟩
⟨save object class⟩ ⟨object class list⟩ ⟨permission⟩

⟨object class list⟩::=⟨object class⟩ ⟨object class list⟩
| ⟨object class⟩
Semantics:

Each of the entities mentioned is permanently retained within the system.

3.3.3.12 ⟨list ...⟩

Function specifications.
⟨list network⟩ ⟨input network cursor⟩ ⟨input list quantifier⟩ ⟨input format specification⟩ ⟨output network text⟩
⟨list network objects⟩ ⟨output object list⟩
⟨list network object classes⟩ ⟨output object classes⟩
Semantics:

The ⟨list network⟩ command may be used to list the topology of a network or a subnetwork or even to list the properties of a single component in the network. These different options are controlled by the ⟨list quantifier⟩. The form of the output is determined by the ⟨format specification⟩.
⟨list network objects⟩ and ⟨list network object classes⟩ permit the user to recall what parts and what object classes are currently available.

3.3.4 Human Performance

The same remarks made in Section 3.2.3 and 3.2.4 are appropriate for the network editor as well and will not be repeated here.

3.4 The Shell Program

The shell program is the overseer or executor program. It is, in essence, a miniature operating system and is to be designed so as to insulate the other major components of the system from a possible inhospitible environment which the host operating system would otherwise provide.

The general capabilities of the shell provide a file system, the ability to support and invoke the dbms and the network editor as well as the distribution planning programs, a sophisticated prompting facility and ⟨help⟩ features.

3.4.2 Shell Files

There are five standard shell files: ⟨INfile⟩, ⟨OUTfile⟩, ⟨ERRfile⟩ ⟨EXTinfile⟩ and ⟨EXToutfile⟩. Each is discussed below.

3.4.2.1 ⟨INfile⟩

This is the standard input file will be associated with the terminal input device.

3.4.2.2 ⟨OUTfile⟩

This is the standard output file and will be associated with the terminal output device.

3.4.2.3 ⟨ERRfile⟩

This is the standard error output file. It is associated with the terminal.

3.4.2.4 ⟨EXTinfile⟩

This is the standard input file for reading bulk data. Normally, it will be associated with a disk or tape file maintained by the host operating system.

3.4.2.5 ⟨EXToutfile⟩

This is the standard output file for writing bulk data. Normally, it will be associated with a disk or tape file maintained by the host operating system.

3.4.2.6 Internal files

The shell program maintains its own file system. This system is logically independent of the file system supported by the host operating system. These files are known via a catalogue internal to the shell. The permissible operations are discussed below, beginning with Section 3.4.3.6.

3.4.3 Operational Requirements

The shell commands are listed below.

3.4.3.1 ⟨invoke shell⟩

Functional specifications.
⟨invoke shell⟩ ⟨user_id⟩ ⟨database name⟩ ⟨EXTinfile⟩ ⟨EXToutfile⟩ ⟨audit file name⟩
Semantics:

The shell user must supply his ⟨user_id⟩, a database name to which the shell is to be connected and bulk I/O file names. The ⟨database name⟩ identifies the database against which all transactions are to be conducted for the duration of the shell session. All actions affecting the database are recorded on the file specified by ⟨audit file name⟩. This file can be postprocessed to produce ⟨transaction file⟩ for use by the ⟨restore⟩ function (see Section 3.2.2.23).

3.4.3.2 ⟨execution⟩

Functional specifications.
⟨execute⟩ ⟨program name⟩ ⟨from⟩ ⟨to⟩
⟨from⟩ ::= ⟨INfile⟩ | ⟨file name⟩
⟨to⟩::= ⟨OUTfile⟩ | ⟨file name⟩
Semantics:

The program identified by program name is invoked. The input file is identified by ⟨from⟩ and the output file is identified to ⟨to⟩.

3.4.3.3 ⟨pipeline⟩

Functional specifications.
⟨pipeline⟩ ⟨program name list⟩ ⟨from⟩ ⟨to⟩
⟨program name list⟩ ::= ⟨program name⟩ ⟨program name list⟩ ⟨program name⟩
⟨from⟩::=⟨INfile⟩ | ⟨file name⟩
⟨to⟩ ::=⟨OUTfile⟩ | ⟨file name⟩
Semantics:

⟨pipeline⟩ execution allows the user to execute several programs whose output is the input for the next program in the series. The programs identified in ⟨program name list⟩ will be executed in the order in which they appear in the list. Input to the first program is the file identified by ⟨from⟩ and the output of the list program is placed in the file identified by ⟨to⟩.

3.4.3.4 ⟨execute control file⟩

Function specifications.
⟨execute control file⟩ ⟨control file name⟩
Semantics:

The file identified by ⟨control file name⟩ is to be a control file, i.e., a file of shell commands and processor invocations. These are executed as they are encountered. When the file is exhausted, control returns to the file in which the ⟨execute control file⟩ was embedded (if the exhausted file is the terminal file ⟨INfile⟩, the shell waits for further input).

3.4.3.5 Subsystem invocations.

The invocations for the subsystems already discussed are listed here for completeness. One further system remains to be discussed, namely, the text editor. Its invocation is also included.

⟨invoke database management system⟩ (3.2.2.9)
⟨invoke network editor⟩ (3.3.3.1)
⟨invoke text editor⟩ (3.6)

3.4.3.6 ⟨create⟩

Functional specifications.
⟨create⟩ ⟨input file name⟩ ⟨input user_id⟩ ⟨input access list⟩ ⟨input retension status⟩ ⟨output file pointer⟩
⟨access list⟩ ::=⟨permission⟩ ⟨user_id list⟩ ⟨access list⟩ | ⟨permsission⟩ ⟨user_id list⟩
⟨permission⟩::= ⟨read⟩ | ⟨write⟩ | ⟨read-write⟩
⟨user_id list⟩::=⟨user_id⟩ ⟨user_id list⟩ | ⟨user_group_id⟩ ⟨user_id list⟩ ⟨user_id⟩ ⟨user_group_id⟩
⟨retension status⟩::=⟨temporary⟩ | ⟨retension date⟩
Semantics:

⟨create⟩ produces a mapping from a ⟨file name⟩ to a ⟨file pointer⟩ which is a descriptor used by the system to permit other system processes to access the file of interest. A file may be designated as ⟨read⟩-only, ⟨write⟩-only, or both ⟨read⟩ and ⟨write⟩. The result of ⟨create⟩ is always to produce a pointer to an empty file.

3.4.3.7 ⟨open⟩

Functional specifications.
⟨open⟩ ⟨input file name⟩ ⟨input user_id⟩ ⟨input mode⟩ ⟨output file pointer⟩
⟨mode⟩::= ⟨read⟩ | ⟨write⟩ | ⟨read-write⟩
Semantics:

⟨open⟩ returns the ⟨file pointer⟩ to the user pointing at the first block of the file. The mode specifies the use to which the file will be put and must be compatible with the ⟨permission⟩ definition made at the time the file was created. Attempting to ⟨open⟩ a file which is already open is an error.

3.4.3.8 ⟨close⟩

Functional specification.
⟨close⟩ ⟨input file pointer⟩
Semantics:

⟨close⟩ releases the file designated by ⟨file pointer⟩ from the user's control.

3.4.3.9 ⟨unlink⟩

Functional specifications.
⟨unlink⟩ ⟨input file name⟩
Semantics:

⟨unlink⟩removes the file specified by ⟨file name⟩ from the system. Files whose ⟨retension period⟩'s have exceeded, are automatically unlinked by the system.

3.4.3.10 ⟨save file⟩

Functional specifications.
⟨save file⟩ ⟨input file name⟩ ⟨input retension period⟩
Semantics.

The file designated by ⟨file name⟩ is to be retained by the system for the period specified by the retension period.

3.4.3.11 ⟨list files⟩

Functional specifications.
⟨list files⟩ ⟨input user_id⟩ ⟨output file list⟩
Semantics:

The list of files associated with ⟨user_id⟩is routed to the output file.

3.5 Distribution Planning Programs

Because of the necessity of performing text edition while using the system, a text editor will be included as part of the system. The particular editor will be the one defined in R19.