NBSIR 86-3482

The Internal Revenue Service Post-Of-Duty Location Modeling System - Final Report

Paul D. Domich, Karla L. Hoffman, Richard H. F. Jackson
Marjorie A. McClain

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
National Engineering Laboratory
Center for Applied Mathematics
Gaithersburg, MD 20899

July 1986

Issued February 1987

Sponsored by
…Research Division
…al Revenue Service
…ington, DC 20224

NBSIR 86-3482

THE INTERNAL REVENUE SERVICE POST-OF-DUTY LOCATION MODELING SYSTEM - FINAL REPORT

Paul D. Domich, Karla L. Hoffman, Richard H. F. Jackson
Marjorie A. McClain

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards
National Engineering Laboratory
Center for Applied Mathematics
Gaithersburg, MD 20899

July 1986

Issued February 1987

Sponsored by
The Research Division
Internal Revenue Service
Washington, DC 20224

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, *Secretary*
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Director*

ABSTRACT

This report documents a project undertaken by the National Bureau of Standards to develop a mathematical model which identifies optimal locations of Internal Revenue Service Posts-of-Duty. The mathematical model used for this problem is the uncapacitated, fixed charge, location-allocation model which minimizes travel and facility costs, given a specified level of activity. The report includes a discussion of the location problem and the mathematical model developed. Data sources identified and used are also described. Brief descriptions of the mathematical techniques used and the interactive, user-friendly computer system built to solve the problem are also provided. The system is microcomputer-based and uses menus and graphically displayed maps of tax districts for interactive inputs and solution outputs.

Keywords: facility location, uncapacitated, location-allocation, plant location, warehouse location, fixed-charge, personal computer, microcomputer, interactive graphics, greedy algorithm, heuristic algorithm.

ACKNOWLEDGMENT

There have been many who have participated in the successful completion of this project to whom we owe a great debt of gratitude. From the National Bureau of Standards, there is Ralph Schofer, whose experience in transportation modeling was invaluable. Thanks are also due to Dmitri Krass, who worked out the details of the graph coloring algorithm to display the final solution, to Dave Tate for his work on the solver routines, and to Howard Hung who helped in the model formulation. At the Internal Revenue Service, a huge debt is owed to Ellie Convery, whose guidance at all levels of project design and implementation made it possible to complete it successfully. Jim Iocozzia's untiring work behind the scenes ensured, as much as humanly possible, the receipt of data and documents on time and accurately. John Hiniker, John Angle, and John Kam also provided assistance, guidance, and answers when sorely needed.

TABLE OF CONTENTS

1. INTRODUCTION

The Internal Revenue Service (IRS) maintains field offices, called Posts-of-Duty (POD's), within each tax district. Staff at these offices are drawn from the Taxpayer Service, Examination, Collection, and Criminal Investigation Divisions. These POD's are located within the tax district to minimize travel and facility costs while maintaining a high degree of accessibility to the taxpayers. In order to keep costs at a minimum, the IRS would like its Posts-of-Duty close to its centers of activity. Thus whenever economic and demographic shifts within a state cause corresponding shifts in the location of tax activities, the POD's responsible for such activities might also need to be moved. Although general guidelines are set forth in an Internal Revenue manual, each district develops its own specific guidelines and methodology for selecting new POD locations or eliminating existing ones.

In view of the potential benefits to be realized by both the taxpayers and the IRS, the National Office of the IRS contracted with the National Bureau of Standards (NBS) to produce a mathematical model which finds optimal locations of Posts-of-Duty based on minimizing costs to the IRS and to the public. The work statement for this project required that the following factors be considered in this modeling effort:

a) the volume and complexity of the workload expected or already known to be in effect at the Posts-of-Duty;
b) the accessibility of the office (bearing in mind terrain and distance factors);
c) demographic considerations such as density, growth and migration of the population; and
d) the costs including staffing, availability of staff and the availability of Federal space in buildings where Posts-of-Duty could be located.

The decision problem addressed here is that of determining the minimum number of Posts-of-Duty needed to satisfy the current workload in a district and to locate them physically in the district so as to minimize travel and facility costs. More specifically, the problem can be stated as follows.

<u>GIVEN</u>
- a set of existing and potential POD locations,
- a specific level of activity in the district,

<u>LOCATE</u>
- POD's so as to minimize facility costs and travel costs (to IRS staff and to taxpayers).

A mathematical model for the problem of determining the optimal <u>number</u> of POD's in a tax district is known as a "facility location problem." When formulated as such a problem, the model is easy to describe, but may have thousands of variables making it impossible to solve to proven optimality on a microcomputer. We used this basic model, but built a system around a solution technique that determines "near" optimal locations of a <u>prespecified</u> number of POD's in very reasonable times on a microcomputer. Using Lagrangian relaxation techniques, we then provide a measure of how far from optimality the solution can be. This quick response time allows the model to be solved repeatedly, altering the allowable number of POD sites each time, thereby determining an optimal number of POD's as originally desired.

From discussions with IRS District Office staff, we found that, in most cases, this system will be used to relocate a small number (less than five) of POD's in a district where there are also a relatively small number of potential locations for the new offices (usually not more than ten). Various configurations and their costs can thus be determined in the manner described above, by running the model under different input conditions. If one wished to consider reconfigurations where more of the locations would be altered, that too is possible. The computation time increases significantly, however.

This paper is one of a series of reports documenting the Internal Revenue Service Post-of-Duty Location Modeling System. The reports in the series are as follows.

1) The Internal Revenue Service Post-of-Duty Location Modeling System: Final Report

This report describes the Post-of-Duty location problem and its mathematical model. It discusses the types of data which are considered in calculating costs, describes the methods used to solve the location problem, and gives a brief introduction to the computer implementation of the model. NBS Contact: Richard H. F. Jackson

2) The Internal Revenue Service Post-of-Duty Location Modeling System: User's Manual

This report is a user's guide for the Post-of-Duty location computer system. It gives hardware and software requirements, instructions for installing the system, descriptions of data files, and detailed instructions for operating the system. NBS Contact: Marjorie A. McClain

3) The Internal Revenue Service Post-of-Duty Location Modeling System: Programmer's Manual for FORTRAN Driver

The Post-of-Duty location program is written in two sections of code, one in FORTRAN and the other in PASCAL. This report describes the FORTRAN driver which handles graphics displays and controls input and output for the solution procedure. The report includes an alphabetical list of the FORTRAN routines, describing the purpose, the calling sequence and the variables of each routine. NBS Contact: Marjorie A. McClain

4) The Internal Revenue Service Post-of-Duty Location Modeling System: Programmer's Manual for PASCAL Solver

This report describes the second part of the Post-of-Duty location program, the PASCAL solver. It discusses the algorithms and data structures used to solve a location problem. NBS Contact: Paul D. Domich

The remainder of this report consists of four sections and two appendices. Section 2 provides some background material about the problem under consideration, describes the process of converting from the physical problem to a mathematical representation, describes the precise mathematical representation subsequently developed, states the assumptions made, and ends with a graphical depiction of the flow of the solution procedure. Section 3 contains a full description of the data used in the POD location system, and Section 4 contains an overview of the mathematical procedures used at each stage of the solution procedure. Section 5 is an overview of the complete system developed to solve the problem. It includes brief discussions of the two principal parts of that system: the graphically based input/output procedures in the driver and the mathematically based solution procedures in the solver. Conclusions regarding this effort are given in Section 6. Appendix A describes a process for computing weights that could be used to modify straight-line distances used in the model to reflect more accurately local travel impedances. Lastly, Appendix B contains sample output.

2. THE PHYSICAL AND MATHEMATICAL MODELS

The purpose of this section is to describe the mathematical model developed for the IRS Post-of-Duty problem. In order to do this satisfactorily, it is first necessary to discuss very briefly some of the data used in the model. The intention here is only to provide sufficient detail regarding the data so as to achieve a full understanding of the model, not to describe the data completely. More complete descriptions of the data are given in Section 3.

The first step in providing an accurate mathematical model of the IRS Post-of-Duty location problem is to describe in more detail what is meant by "level of activity" in the simple problem statement given in the preceeding section. This means essentially that, of all the data bases maintained by, or available to, the IRS, the most appropriate for use in representing activity level in this modeling effort must be identified. The next problem to be resolved is to what level this workload data must be aggregated, since there is typically more detail in the data bases than is needed. Also, as prescribed by the statement of work, the model must account for geographic accessibility of the POD's. Lastly, distances between zip codes and corresponding traveling costs must also be computed. Each of these issues is discussed in this section.

2.1 Quantification of Levels of Activity

This POD location model is driven by existing workload within a District. Therefore, the choice of workload measure is critical to our modeling success. A variety of possible measures of workload were postulated, and it was determined that many of them could be used in this model. Almost all of them are available from the IRS Individual Master File (IMF) and the Business Master File (BMF). We examined these data bases for applicability to each of the four main IRS District Office functions: Examination, Collection, Taxpayer Service, and Criminal Investigation. For the Examination activity, possibilities for workload measure include:

a) the number of returns with Discriminant Function (DIF) scores above a certain specified number, within a specified taxpayer income class, or
b) the number of returns in the top X percent (specified by the user or National Office) of DIF scores, by income class, or
c) the number of audits by return class.

For the IRS Collection function, the possibilities include:

a) the number of taxpayer delinquent accounts by return type and amount, or
b) the number of taxpayer delinquency investigations by return type,
c) the data collected for DIIP/DIAP (Delinquent Investigation Inventory Profile/Delinquent Account Inventory Profile) reports.

Similarly for the Criminal Investigation Division, we could use

a) the number of open cases, by return class, or
b) the number of closed cases, by return class.

Finally, for Taxpayer Service, the model could use the number of taxpayer notices sent (first, second, third, and fourth), by return class, and the number of computer-generated error notices sent.

A more complete description of the decisions made regarding the data and the precise data used in this modeling effort is given in Section 3.

2.2 Data Aggregation

As mentioned above, another decision that must be made regarding the data is the level of aggregation. The IMF and BMF data are completely disaggregated to the individual taxpayer record. For the Jacksonville, Florida Tax District, for example, there are more than nine million such records. Both processing considerations and taxpayer security interests dictate aggregation of this data. It was determined that aggregation to the zip code level was both feasible and advisable. Each record on the IMF and BMF contains a zip code, but does not contain any of the other usual information by which an aggregation can be done; e.g., SMSA, census tract, BEA zones. Thus, the decision was made that the data would be aggregated to the five-digit zip code level. This means that the workload data described above for each of the IRS functions was summed for each category to provide totals for each zip code. (Note that zip code data cannot be aggregated to the four-digit zip code level and that three-digit aggregation was deemed too gross to produce meaningful travel and income-level differences.)

2.3 Distance Measurements

The decision to aggregate introduced another problem, that of how to measure distances between zip codes. Since the model will be minimizing travel costs and travel costs are based on distances traveled, it is necessary to have some convenient method for measuring distances between (aggregated) data points, or zip code regions. We needed to associate a specific and unique point with each zip code to be used in distance calculations.

For this, the location of the main Post Office in each zip code region was first considered, but that proved to be difficult information to obtain and update. (The U.S. Postal Service does not keep these coordinates in a machine-readable form.) Also considered was computing a point in the interior of each zip code, which represents, in some sense, the center. Since at least two commercial vendors maintain machine-readable coordinates of zip code regions, this data could be used to compute various measures of the center of a zip code.

After an evaluation of the alternatives, it was found that no mathematical technique can guarantee that the computed point will lie completely within the boundary of a zip code region. E.g., crescent-shaped and doughnut-shaped regions will result in exterior "centers." For these regions, the centers must be moved inside the boundaries by hand. This was not done in the prototype system developed for the Jacksonville District, since the vendor from whom the boundary data files will be purchased also supplies files of "modified" centroids which are guaranteed to lie within the correct boundaries. Having the coordinates of these modified centroids allows the model to compute Euclidean distances between zip code regions quickly. In the interests of brevity, modified centroids will simply be called centroids in the remainder of this report.

2.4 Geographical Difficulty

The cost of traveling between zip code regions is based on the Euclidean distance, but it is not correct to assume that costs are independent of the location of the zip code regions. For example, traveling in the mountains on back roads is more time consuming than in the plains on interstate highways, and this should be reflected in a higher traveling cost for such regions. This is incorporated into the mathematical model by providing weights used to multiply distances between zip code pairs which are more difficult to travel. These weighted distances can be used in regions where mountain ranges exist, bodies of water impede, downtown traffic slows progress, or where direct line distances do not accurately reflect true roadway distances. They could also be used to incorporate other penalties like parking costs that are not explicitly a part of the model.

We have not been able to identify any machine-readable data source that can provide this information to the model for every region in the continental United States. Therefore, if a user wishes to change these values from their default settings, he must calculate the weights outside the model. The actual resetting of the weights can be done interactively inside the system, however. Furthermore, in Appendix A, we discuss a process by which the calculation of the weights can be made more rigorous.

2.5 Costs of Traveling Between Zip Codes

There are basically three types of costs that are used in the model: travel cost, operating cost, and the cost of opening or closing a POD. Travel cost is determined as a dollar-per-mile cost, and is provided to the model in one of the basic input files. Operating cost is based on square feet of office space required by a POD and the rental cost per square foot of that space. Each potential and existing POD site has associated with it a dollar-per-square-foot cost. The cost of opening or closing a POD is determined by the District Office staff and is included as part of the input data. (Opening/closing costs might include items such as cost of relocating employees, cost of hiring new employees, and the cost of training new employees.) Each of these is used in the model to determine the cost of a possible POD configuration.

2.6 The Physical Model

Since we assume that aggregation of workload factors to five digit zip code levels is adequate, the model can be structured physically around the locations and boundaries of these five digit zip codes. The model then becomes what is commonly known as a <u>network model</u>, and the physical realization of the zip codes, shown in Figure 1, can be converted from zip code boundaries in the "real world" to zip code nodes in a network. These nodes are connected by arcs associated with distances between zip code centroids. In this initial, simplified mathematical representation of the zip code network, the distance from node i to node j is denoted d_{ij}.

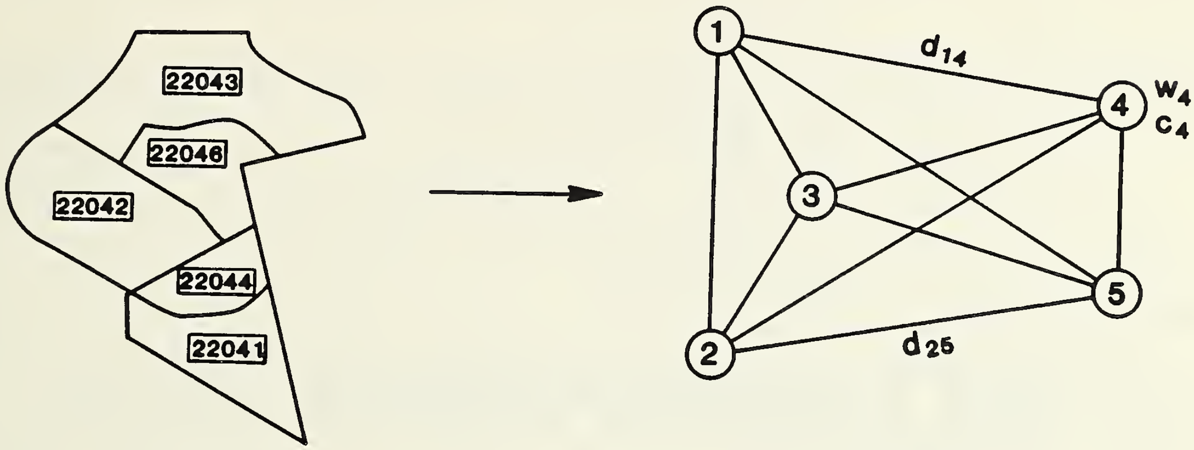

Figure 1: Conversion of Physical Problem to Mathematical Representation.

With such a mathematical representation, the problem of locating POD's in the network and allocating zip code activity to POD locations can be formulated in a mathematically succinct way. For small problems, this representation can aid in the hand calculation of a solution. For example, for the problem in Figure 1 with 5 nodes and 10 arcs, there would be 10 possible choices for locating two POD's. If one were also given the cost of opening/closing POD's, the cost of traveling between POD's, and the workloads at each node (which translate to a number of trips), one could solve this problem simply by calculating the cost of each possible choice and choosing that which is minimal. However, more realistic situations include districts having thousands of nodes and up to 50 potential or current Posts-of-Duty to be located/relocated. As the size of the problem increases, the complexity increases exponentially. Clearly, complete enumeration of such problems is impossible on a microcomputer.

Hence, a solution procedure must be developed that will capitalize on the special structure of the problem and solve it quickly and efficiently. There are many ways this can be done, but before these can be discussed, it is necessary to be more specific about the detailed structure of the proposed mathematical model. This mathematical structure is given next.

2.7 The Mathematical Model

The specific statement of the mathematical model of the IRS Post-of-Duty location problem is given on the next page:

$$\text{minimize} \quad \sum_{j \in J} u_j x_{jj} + \sum_{j \in J} v_j (1 - x_{jj}) + \sum_{j \in J} \sum_{i \in I} (T_{ij} + S_{ij}) x_{ij}$$

subject to

$$\sum_{j \in J} x_{ij} = 1, \; i \in I;$$

$$\sum_{j \in J} x_{jj} = k;$$

$$\sum_{i \in I} x_{ij} \leq M x_{jj}, \; j \in J;$$

where

I = set of all zip code numbers in the district;
J = subset of I containing zip code numbers of all current and potential POD sites;
x_{ij} = 1 if zip code $i \in I$ is assigned to POD $j \in J$, 0 otherwise;
u_j = cost of opening a POD at zip $j \in J$ (\$), if not already opened;
v_j = cost of closing a POD at zip $j \in J$ (\$), if already in existence;
T_{ij} = cost of round-trip travel from zip code $i \in I$ to POD $j \in J$, or

$$2 c g_{ij} d_{ij} f_i;$$

c = cost of travel (\$/mi.);
g_{ij} = a weight representing relative difficulty of travel between zip code $i \in I$ and POD $j \in J$;
d_{ij} = distance from zip code $i \in I$ to POD $j \in J$ (mi.);
f_i = weighted number of trips generated by zip code $i \in I$, or

$$\sum_{p=1}^{P} \sum_{q=1}^{2} \alpha_q \beta_p t_{pq} w_{ip};$$

P = number of IRS activities considered by model;
α_q = weight used in balancing IRS dollars and taxpayer dollars;
β_p = on/off switch for IRS workload class p, = 1 if use this class in problem, = 0 if not used;
t_{pq} = number of trips per workload unit for IRS workload class p made by traveller type q (trip factor);
w_{ip} = workload in zip code $i \in I$ for IRS workload class $p \in P$;
S_{ij} = office space cost for assigning zip code $i \in I$ to POD $j \in J$, or

$$h r_j s_j;$$

h = office space required for each IMF return (sq. ft.);
r_j = number of IMF returns filed in zip code $j \epsilon J$;
s_j = rental cost for office space in POD $j \epsilon J$ ($/sq. ft./yr.); and
k = total number of POD's desired in the solution; and
M = number of zip codes in the set I

This model is referred to in the literature of Operations Research as an uncapacitated, fixed-charged location-allocation model (see Francis and White (1974), and Wagner (1975)). The objective function is a measure of the cost of opening/closing Posts-of-Duty plus office space costs and the costs incurred by IRS personnel and taxpayers for traveling between POD's and taxpayer locations. The constraints are the feasibility conditions. The first constraint assures that each zip code region is assigned to one and only one Post-of-Duty for coverage. The second one requires that the number of Posts-of-Duty be some preset value, k. The third constraint assures that zip codes are assigned only to potential or current POD zip codes which have been chosen to be Posts-of-Duty in the solution. For the Jacksonville District, the broadest application (where each zip code can be a POD) of this model has at least one million variables and two thousand constraints, not counting the integer constraints. In that form it is too big to be solved by any existing integer programming code.

The goal, therefore, is to find ways to exploit the special structure of this problem to reduce its size. For example, whenever a Post-of-Duty site is fixed (i.e., must remain in existence), the number of variables in the problem decreases significantly. Similarly, whenever one specifies that a zip code site does not qualify for a POD location, the number of possible alternatives to evaluate is reduced. Finally, whenever one invokes the rule that no zip code can be assigned to a Post-of-Duty whose distance from it is greater than some prespecified limit (in the Jacksonville district this maximum distance is approximately 75 miles), the number of variables is also reduced. Each of these reasonable modeling specifications, and others, can be used to reduce both problem-size and problem-complexity.

One last note should be included here reqarding the calculation of S_{ij}, the office space cost for each zip code. Currently, this is computed by relating number of IMF returns to square feet of office space in use. A more appropriate measure would relate square footage to workload for each of the IRS functions. Only with such a measure can one determine how much square footage is necessary for each of the functions: Examination, Collection, Taxpayer Service and Criminal Investigation.

2.8 Assumptions

It is important to understand the following assumptions that are implicit in the use of this model.

a. Travel costs, operating costs, and costs of opening or closing Posts-of-Duty are the driving data for the model. Staffing costs are not to be considered in this model since all IRS functions will be treated according to current or projected tax activity within that functional area. (E.g., specification of DIF cutoff scores, which determine examination workload, is outside this modeling effort.) Thus the model is not intended to be used to compute the optimal allocation of IRS staffing funds among the IRS functions, but rather to determine the location of offices to minimize travel costs and facility operating costs given a specified amount of activity within each of the major IRS functions, and to assign specific zip codes to POD's.

b. Aggregation of workload factors to the five digit zip code level is adequate for the purposes of this model.

c. Travel costs to the taxpayer within this model are not necessarily considered equal to the travel costs to the IRS.

d. Taxpayer service walk-ins will be treated in the model. Taxpayer service phone activity will not.

e. There will be no upper or lower limits set on the staff-size of Posts-of-Duty. The model will determine both the location and the size of a facility.

f. Future demand projections will be supplied to the model and not generated internally. The model will produce near-optimal locations based on the static demand data available to it. When future projections are required, the data will be projected externally. The model will then solve this new static problem.

g. The model user (District Director, National Office analyst, or District Office analyst) will be asked to define reasonable ranges on the number of POD's to be located, geographical difficulty factors, maximum distance per trip that an agent or a taxpayer will be asked to travel, cost per square foot of office space for existing and potential POD sites, weights for IRS/taxpayer costs, and categories of workload to include.

2.9 Flow of Solution Procedure Use

In Figure 2, the method by which this solution procedure is to be used is indicated. In order to understand the use of this procedure, consider a situation where there are 25 current, fixed POD's, 10 potential sites, and the user wants up to 5 new POD's opened. In this case, five successive runs of the Location Modeling system would be made, with five different values of k, the number of desired sites in the final solution. The first step in the procedure is to perform the model simplifications noted previously that reduce the large model to a smaller, solvable one. With this reduced static data (any required data projection is performed outside of the model), a near-optimal POD configuration with k open POD's is computed, along with the allocation of the remaining zip codes to these POD's. In addition, the cost of that configuration is computed, and the solution is displayed to the user graphically. Hard-copy output is also available, as is a bound on the maximum difference between this near-optimal configuration and the optimal one. Empirical evidence reported in the literature (see Morris (1978)) indicates

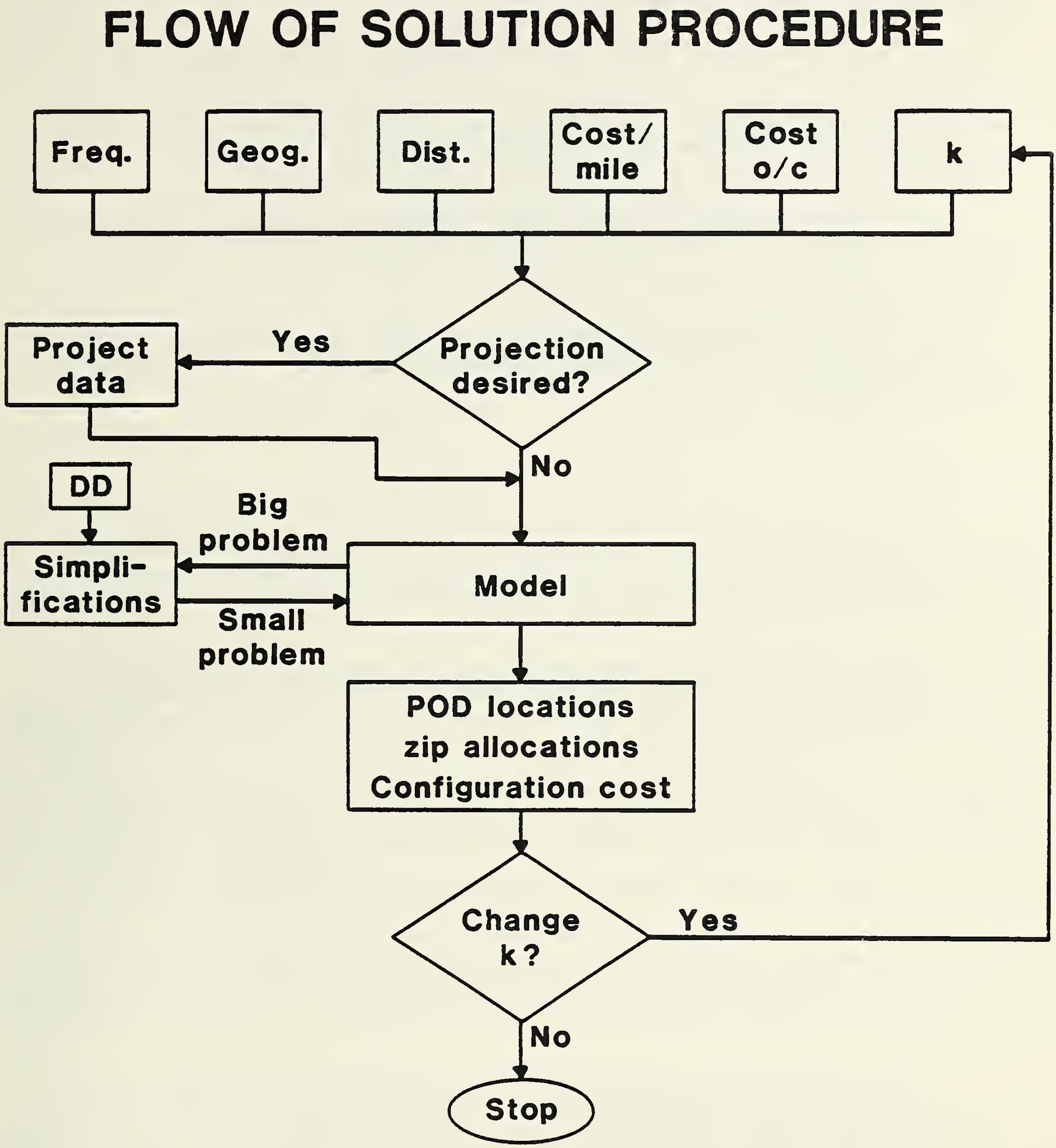

Figure 2: Flow of Solution Procedure Use.

that very often the solution obtained by this procedure is optimal. The results of Morris indicate that if the bounding procedure is used and allowed to run to termination, 96% of the time it will terminate with the optimal solution (Morris, (1978)). In any case, the solution is guaranteed to be no worse than the bounds provided.

This solution process stops once the user has obtained runs for all values of the input data of interest. A major assumption governing the use of this procedure is that each of the model runs can be performed quickly. This has turned out to be true. For example, one of the test runs with 23 current, 16 potential, and three new sites to be chosen, required less than 5 minutes on an IBM PC-XT to find a solution, which in fact was optimal.

The computer programs could be easily modified to produce solutions for a range of values of k automatically, thereby simplifying this iterative process. However, this would be done at the expense of increased solution times for single values of k.

Now that the model of the POD location problem has been given, we are ready to describe in complete detail the data used by it. This is given in the next section.

3. DATA DESCRIPTION

Critical to the success of any quantitative analysis effort is both the ability to model adequately the system under question, and the ability to acquire the data which this representation requires. This section will discuss the data acquisition issues related to modeling the IRS Post-of Duty Study as a "facility location problem".

When we first began this effort, we were asked to consider the following factors which affect the location of offices: 1) the workload currently existing or projected to exist in various regions, 2) the accessibility of the office (considering terrain and transportation patterns), 3) the likely shifts in population and tax activity, and 4) the costs, including opening and/or closing costs, availability of staff, and variable space costs.

We had a variety of meetings with IRS National Office and Field Personnel to determine which of this information was quantifiable and what might be the best ways of acquiring data. Not surprisingly, some data were much easier to acquire than others. For simplicity of exposition, we will categorize the data by whether they are: a) descriptive of activity within a zip code, b) descriptive of how zip code pairs relate, or c) general cost data which are not unique to specific zip codes.

3.1 Data on Activity Within a Zip Code

Data which pertain to activities within a zip code are workload data. It was determined that much of the information concerning the workload of a specific zip code could be ascertained from the IRS Individual and Business Master Files.

For <u>examination workload</u>, the Individual and Business Master Files contain information on the returns which had been audited and, for certain types of returns, the DIF score of every return. From this information, one could define workload as:

1) number of taxpayers with DIF score above national cutoff levels in zip code i and tax class j,
2) number of taxpayers with DIF scores above some inventory level in zip code i and tax class j,
3) number of taxpayers audited within tax class j and zip code i.

Each of these data items were collected for each zip code and tax class within a district. When discussing the use of this data as representative of "examination workload", both IRS field-office and national-office personnel agreed that DIF score information adequately represented workload for return classes with scores.

For the <u>collection workload</u>, the problem is more complicated. Field-personnel believe that the zip code DIIP/DIAP (Delinquent Investigation Inventory Profile/Delinquent Account Inventory Profile) reports would provide a good measure of workload. However, these data are not yet collected into machine-readable summary reports by zip code. (There are only District summaries at this time.) For this reason, the study group chose to use both the number of Individual Master File Taxpayer Delinquency Investigations

(TDI's) and Taxpayer Delinquent Accounts (TDA's) issued within a zip code. Field-office collection personnel believe that these data serve only as a weak substitute for the true measure of collection workload. We understand that DIIP/DIAP monthly workload summary reports are currently being instituted on a trial basis. When these data become more widely available, we recommend that they be substituted for the TDI and TDA data currently in use.

For the taxpayer-service function, surrogate data must be used, since currently counts do not exist of the number of people arriving at IRS offices for taxpayer assistance. What is needed is not the number of people who arrive, but the number of people residing within each zip code in a district who come to an IRS office for taxpayer assistance. Although this would be the "perfect" measure of taxpayer service workload, we do not believe it worthwhile to create a data-collection effort to acquire these data. Instead, we chose to use as a substitute the number of TDA and TDI first notices and the number of math error notices which originated in each zip code. (Second, third and fourth notices were ignored for fear of too much double counting.) These data can easily be retrieved from counts of the Individual and Business Master Files. Although these data are only a surrogate for the true data, they are believed to be a relatively good measure.

For criminal investigation, the number of criminal investigation cases were tallied by zip code. The total number of criminal investigation cases, in even the largest of Districts (e.g. Jacksonville), is very small relative to the amount of workload in other IRS functions. Due to this small workload count, we suspect that eliminating this data from the model will not affect the final locations of zip codes. The Criminal Investigation Function will not be adversely affected by this omission since many of these criminal investigations are performed jointly with local, state and federal law enforcement officials, and Criminal Investigation personnel can use offices provided by these agencies. Currently, the Criminal Investigation data is being collected and used in the modeling effort. We will investigate -- using sensitivity analysis -- the degree to which this data affects the overall results of the model.

Once these workload counts have been collected, we need to be able to relate the amount of workload within a zip code to the number of trips required to handle each of these instances of workload. For example, suppose we have determined that there are 25 examination returns with high DIF score in zip code 11111 for class 1. We must then be able to translate these 25 returns into the number of trips the taxpayer and/or an IRS employee takes to complete these audits. We have received data from Jacksonville indicating the conversion factors necessary to convert workload counts into trips for each audit class and each IRS function. Table 1 provides this data.

TABLE 1
TRIP FACTORS FOR JACKSONVILLE DISTRICT

Examination*

	% handled by IRS	% handled by Taxpayer	Trips per 100 cases-IRS	Trips per 100 cases-Taxpayer
Class 1	10.9	89.1	95	105
Class 2	18.4	81.6	88	165
Class 3	6.7	93.3	47	140
Class 4	12.0	88.0	72	185
Class 5	10.0	90.0	69	195
Class 6	51.1	48.9	100	230
Class 7	34.4	65.6	142	120
Class 8	40.2	59.8	176	148
Class 9	66.3	33.7	203	190
Class 10	26.7	73.3	113	123
Class 11	33.0	67.0	156	123
Class 12	61.3	38.7	212	138

* These data are used in the following way. If there are 100 cases in Class 1, they will generate a total of 104 trips because:
1) 100 x .109 x .95 = 10.4 trips,
2) 100 x .891 x 1.05= 93.6 trips, and
3) 10.4 + 93.6 = 104.

Collection

For each 100 cases, the following trips will be generated for IRS and taxpayers:

Trips/100 cases

Case type	IRS	Taxpayers	Total
TDI	300	75	375
TDA	200	30	230

Taxpayer Service

Trips per 100 notices = .0247.

Criminal Investigation

Trips per C.I. Case = 120.

3.2 Zip Code Pair Data

For each zip code-POD pair, a measure of the geographic difficulty of traveling between the zip code regions, and the distance between regions, is needed. As described above, a modified centroid is used as the specific location to be used with each zip code in distance calculations. The Euclidean distance between two centroids is then used to represent the average distance a taxpayer/IRS employee traveled.

We have incorporated geographic difficulty into the mathematical model by allowing the user to specify weights with which to multiply distances between zip code pairs which are more difficult to travel. For more on this see Section 2.4 and Appendix A.

Other data which must be described is data relating to each current or potential POD site.

3.3 Post-of-Duty Site Data

For each POD zip code region, we need to know the cost of opening a Post-of-Duty (if it is does not currently exist), the cost of closing an existing Post-of-Duty, and the operating costs associated with maintaining a Post-of-Duty in that zip code. Operating costs are defined to be the costs of leasing space in a building for the purpose of housing the personnel assigned to this POD site. We believe that this cost should be relative to the number of people assigned to that zip code. Therefore, square footage cost of space is used in the model.

The opening/closing costs for the Jacksonville district are provided in Table 2 below, as well as the cost of square footage for each potential and current POD zip code region. Note that the data on square footage costs cannot be uniform throughout a zip code. The Facilities Management personnel within IRS were asked to provide costs based on the most likely locations for an office within that zip code region. In discussions with field office personnel, they indicated that this type of information is relatively well-known and easy to acquire. Note also that since travel costs and rental costs are calculated on an annual basis, the one-time costs of opening/closing a POD must be amortized across some fixed number of years.

TABLE 2
OPENING, CLOSING, AND RENTAL COSTS FOR JACKSONVILLE DISTRICT

POD	TYPE	OPENING COST*	CLOSING COST*	OFFICE RENTAL COST/SQ.FT
32018	CURRENT	0.00	1.00	9.54
32202	CURRENT	0.00	1.00	8.93
32301	CURRENT	0.00	1.00	13.25
32401	CURRENT	0.00	1.00	8.99
32501	CURRENT	0.00	1.00	10.35
32601	CURRENT	0.00	1.00	8.53
32670	CURRENT	0.00	1.00	8.30
32771	POTENTIAL	1.00	0.00	10.00
32801	CURRENT	0.00	1.00	13.84
32901	CURRENT	0.00	1.00	8.23
33130	CURRENT	0.00	1.00	11.36
33169	CURRENT	0.00	1.00	13.25
33173	CURRENT	0.00	1.00	0.00
33174	POTENTIAL	1.00	0.00	13.00
33301	CURRENT	0.00	1.00	11.19
33401	CURRENT	0.00	1.00	0.00
33432	POTENTIAL	1.00	0.00	16.00
33450	CURRENT	0.00	1.00	10.13
33583	CURRENT	0.00	1.00	10.68
33602	CURRENT	0.00	1.00	11.95
33616	POTENTIAL	5.00	0.00	0.00
33702	FIXED	0.00	1.00	10.06
33801	CURRENT	0.00	1.00	6.10
33907	CURRENT	0.00	1.00	10.89

* These are fictitious, since actual costs of opening and closing were unavailable at the time these runs were made.

Finally, data regarding situations unique to specific districts might also be collected. Examples include offices which must remain open, areas which cannot be considered for POD locations, any unusual transportation costs (e.g., high parking fees, severe traffic problems), and any rules specific to the district (e.g., Revenue Officers never travel more than 60 miles, taxpayers are never asked to travel more than 45 miles for an audit).

The only other data which this modeling effort requires is data which is global in nature; that is, data not unique to a specific zip code region.

3.3 General Cost Data

The facility location model requires the specification of a functional relationship between workload and personnel. This function together with a square footage per person requirement determines the staff assigned to a specific zip code and the accompanying square footage required for the POD site. This data is not only needed by the model to represent adequately the costs of operating a POD site, but is also useful for presenting the results. It is far more informative to provide information regarding staffing needs within a POD site than to state only where each of the POD's is to be located.

Finally, one needs the cost per mile to be associated with traveling to or from a POD site for both an IRS agent and a taxpayer. This data should be provided in terms of a cost-per-mile figure. The tables provided as part of the sample output in Appendix B contain the existing realization of these data for the Jacksonville District Office. The next section gives an overview of the algorithms developed to solve the POD location problem using the data described above.

4. SOLUTION ALGORITHM OVERVIEW

Two approaches to the computer solution of the mathematical POD location problem can be categorized by whether the problem is to be solved on a mainframe computer or on a microcomputer. In the former case, using good reduction and reformulation techniques yielding less than 50 zero-one variables, the problem can be solved to optimality. If, on the other hand, the problem size cannot be reduced sufficiently, "heuristic" approaches can be used together with other integer programming techniques to yield "good" solutions with estimates of how far from optimality they could be. A major advantage of using the heuristic approach is that facility location problems such as those considered by IRS can typically be solved in seconds on a mainframe computer, which correspondingly means that they can be solved in a few minutes on a microcomputer.

The latter is an option which has associated with it a number of other advantages, including ease-of-use, portability, and interactive, graphically based input of problem parameters and output of solution configuration. It is for these reasons that the decision was made to use a heuristic approach to solving the problem and capitalize on the use of a microcomputer, thereby avoiding the communications problems of building a mainframe procedure that would be resident on a central computer and available on a time-sharing basis by dial-up for each District Office.

The method for finding a "good" solution to the facility location problem involves two well-known and dependable heuristic procedures. The first is a Greedy heuristic (see, for example, Cornuejols, et al., (1977)) and the other is the Interchange heuristic (see, for example, Teitz and Bart (1968)). To display the solution found by this method, a coloring of the zip code areas on the map is required. The graph coloring algorithm used is the Sequential Least-first Interchange Algorithm (see Matula et al., (1972)), which determines a coloring of the map so that no two adjacent zip code areas share the same color. These three procedures are briefly outlined below.

Assume that the desired number of POD's in the final solution is different from the number of currently existing POD locations. The Greedy heuristic in its simplest form will add/delete POD sites from the current POD configuration sequentially, so to minimize the total cost given the set of POD sites already selected. This procedure is called "greedy" since it will only examine the cost of a single addition/deletion at each step, and does not consider the addition, deletion, or exchange of two or more POD sites jointly. For the facility location problem, a Greedy procedure will terminate with a good, but not necessarily best, set of POD locations of the prescribed number.

Once the target number of facilities has been allocated by the greedy procedure, the interchange procedure tries to determine a better solution with the same number of open POD sites. The procedure iteratively locates pairs of POD sites, one which is currently selected and one not, such that if the two are interchanged in the current configuration, the overall cost will be reduced. When no such pair exists, the routine terminates with the last configuration.

The Greedy heuristic and the interchange heuristic described above are well-known (see Cornuejols, et al. (1977) and Erlenkotter (1978)) to produce good solutions to the facility location problem. A drawback of these procedures is that it is difficult to determine when the generated solution is in fact the optimal integral solution to the described problem. However, it is possible to investigate the optimality of a solution by generating lower bounds on the optimal integer objective function value. One such bound can be obtained by solving the linear programming (LP) relaxation of the original problem, i.e., the original problem without the integrality constraints.

In general, the LP formulation of the facility location problem has a large number of constraints in the problem description and it, too, can be difficult to solve. Lagrangian relaxation techniques can be used to produce the optimal LP objective function value in an efficient iterative manner, and provide at each step a lower bound on the optimal solution to the original facility location problem. The relaxation we used is one which removes the restriction that a zip code is serviced by only one POD and adds a penalty to the objective function for violations of these constraints. Further, by rounding the possibly fractional real-valued solution produced by this method, an improved integral solution may be found as a by-product. The interested reader may refer to the many articles in this subject (e.g., Cornuejols, et al. (1977)).

To display in color the final assignment of zip code areas to POD locations determined above, it is necessary to ensure that no two adjacent POD service areas, i.e., two areas sharing a common border, are colored with the same color. This is a map coloring problem, where the regions involved are groups of customers aggregated by their assigned POD facility. The problem is to choose colors C_j for the regions V_j of a graph G, such that C_i is not equal to C_j if V_i and V_j are adjacent regions, and in such a fashion that a "small" number of colors are used. Since all of the zip code maps can be represented as planar graphs (i.e., graphs that can be drawn on a sheet of paper so that no two edges cross), theoretically all can be colored using only four colors. In practice, to find a four-coloring is a very difficult problem, so a five- or six-coloring is used. For a description of the coloring algorithm, see Matula, et al., "Graph Coloring Algorithms," (1972). The procedure used is called the Sequential Least-first Interchange heuristic (SLI).

5. OVERVIEW OF THE LOCATION MODELING SYSTEM

The model and solution techniques described earlier in this report have been implemented in a microcomputer-based software system. It is a highly interactive, menu-driven system, that uses function keys to allow the user to move through the solution process. It was developed on an IBM PC-XT with a math coprocessor, a ten-megabyte hard disk and a color monitor, but has been run on a similarly configured compatible. The graphical displays are accomplished by way of zip code maps drawn on the screen from latitude-longitude data points that describe each zip code in the tax district. Zip code data files and IRS data files containing workload and other information are required by the system, and must be obtained by the user. The package performs the following functions.

1) Displaying Workload

The user may choose the type of workload to be displayed, such as the number of returns examined or the number of criminal investigation cases. A state zip code map is then drawn, with each zip code shaded according to the workload it generates.

2) Displaying Initial POD Sites

A state zip code map is drawn showing where POD's are currently located and also where new POD's could potentially be located. The user may make modifications interactively to this information to specify POD's which may not be or must be in the set of POD's determined by the solution procedure.

3) Solving for Optimal POD Locations

Using the initial POD information specified by the user, the solution procedure calculates the cost of assigning each zip code to each current or potential POD. The cost includes travel costs associated with the workload of each zip code, office rental costs, and costs of opening new POD's or closing old POD's. The user may set parameters indicating what types of workload (such as taxpayer service or criminal investigation) should be included in the travel cost calculations. Also, weights may be assigned to pairs of zip codes and POD sites to scale travel costs. The user must specify the number of POD's desired in the solution. The solution procedure then determines the set of POD's which will result in the least total cost for the district.

4) Displaying Optimal POD Locations

A state zip code map is drawn showing where the new POD's determined by the solution procedure are located. Also, a report file is generated which summarizes the problem specification and the solution. The report includes a list of which zip codes are to be assigned to which POD's.

5) Controlling Display of Maps

On any of the state zip code maps mentioned above, the user may zoom in on a small region, back up to a larger region, or find the five-digit zip code number of an area on the map.

There are four stages to solving a POD location-allocation problem with the IRS POD Location Modeling system. In the first stage, problem data are input by the user. In many cases, the data files will have been previously created and the user need only update them if necessary. In the second stage the user will interact with the POD Location System to provide additional local data, or modify problem parameters for a particular configuration desired. During this preprocessing and data manipulation stage, one can also obtain information about the data or the physical situation by using the display properties of the system. The third stage is the one in which the problem is actually solved using the heuristic approach mentioned above. Finally, in the last stage, the solution is displayed and a summary report of the solution is written. Each of these stages is discussed in more detail in (Domich, et al., (1986a)).

6. CONCLUSIONS AND FUTURE WORK

The model and system described here have been built and will be undergoing extensive field tests in the Jacksonville District Office of the Internal Revenue Service. Preliminary response to its use has been gratifyingly good. We will be responding to suggestions for improvements and modifications and will test it using data for the Pittsburgh District in the near future.

In this concluding section, we wish to point out that the IRS POD Location Modeling System has potential uses other than those described in this report. For example, it could easily be modified to solve other facility location problems in different application areas. In fact, the solution algorithm routines could be separated from the graphical display routines to become part of a general facility location solution routine. In addition, the graphical display routines could be used as a stand-alone data display package.

Another possible use is as a device for determining the optimal size of Posts-of-Duty. If there were good data available on the workload that can be handled by employees in each of the four IRS functional areas, then, once locations have been determined, optimal size could also be output.

Some of these ideas will be pursued in our future work.

BIBLIOGRAPHY

Cornuejols, G., Fischer, M.L., and Nemhauser, G.L., "Location of Bank Accounts to Optimize Float: An Analytical Study of Exact and Approximate Algorithms," Management Science, Vol. 23, 789-810, 1977.

Domich, P.D., Jackson, R.H.F., and McClain, M.A., "The Internal Revenue Service Post-of Duty Location Modeling System: User's Manual," National Bureau of Standards Technical Report NBSIR 86-3471, Gaithersburg, MD, July, 1986a.

Domich, P.D., Jackson, R.H.F., and McClain, M.A., "The Internal Revenue Service Post-of Duty Location Modeling System: Programmer's Manual for the FORTRAN Driver," National Bureau of Standards Technical Report NBSIR 86-3473, Gaithersburg, MD, July, 1986b.

Domich, P.D., Jackson, R.H.F., McClain, M.A., and Tate, D.M., "The Internal Revenue Service Post-of Duty Location Modeling System: Programmer's Manual for the PASCAL Solver," National Bureau of Standards Technical Report NBSIR 86-3472, Gaithersburg, MD, July, 1986c.

Erlenkotter, D., "A Dual-Based Procedure for Uncapacitated Facility Location," Operations Research, Vol. 26, no. 6, 992-1009, 1978.

Francis, R.L., and White, J.A., Facility Layout and Location, Prentice-Hall, Englewood Cliffs, NJ, 1974.

Garfinkel, R.S. and Nemhauser, G.L., Integer Programming, John Wiley & Sons, New York, NY, 1972.

Morris, J.G., "On the Extent to which Certain Fixed-Charge Depot Location Problems Can Be Solved by LP," Journal of the Operational Research Society, Vol. 29, no. 1, 71-6, 1978.

Matula, D.W., Marble, G., and Isaacson, J.D., "Graph Coloring Algorithms," in Ronald C. Read, Graph Theory and Computing, Academic Press, New York, NY, 1972.

Teitz, M.B. and Bart, P., "Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph," Operations Research, 16, 955-61, 1968.

Wagner, H.M., Principles of Managment Science, Prentice-Hall, Englewood Cliffs, NJ, 1975.

APPENDIX A: TRANSPORTATION VARIABLE DEVELOPMENT

In this appendix we describe, for the record, a process for modifying straight-line distances used in the model so as to reflect more accurately travel times between the points in question. The technique for incorporating these modifying weights is described in Section 2.4 of this report and also in (Domich, et al. (1986a)). However, what is not described in either place is how one might determine the weights required by the system. That is the subject of this appendix.

The purpose here is to develop a measure of travel impedance between pairs of zip code centroids for use in selecting efficient locations for POD's. Since a limited number of POD's usually will be provided within each state boundary, each zip code centroid in a state need not have an impedance value between itself and every other such centroid in the state: it is necessary only to develop impedance values among pairs of centroids which cluster around the established or projected POD's. The maximum allowable travel distance mentioned in section 2.7 can be used to determine these clusters.

We recommend that a pilot test be conducted using real data from the Jacksonville District. A prime objective of the test is to provide cost-effective impedances among zip code centroids for subseqeunt use in location studies. Many measures of impedence have been tested and explored by land-use and transportation modelers over the years. The exercise described here considers two of the most popular ones, and considers only off-peak (mid-day) travel times by automobile.

The simplest is to develop travel impedances among zip code centroids by computing the Euclidian distance between selected centroids, and multiplying this crow-flight distance by a distict-wide desired circuitry factor to obtain a better estimate of true "road" distance. To obtain time of travel, this distance would then be divided by an assumed speed, the value of which would be area-dependent: rural, suburban or urban. Terminal times (unparking, parking and walk time) and intrazonal times would also be developed and applied for the same three area categories. Some hand modification might be necessary to reflect impediments to travel such as bridges, toll booths and swamps.

A slightly more sophisticated alternative should also be tested. In this alternative, travel impedances in major urban areas would be manually developed and combined with the aforementioned procedures for the rest of the state. The procedure would be first to identify the larger urbanized areas. (In Florida, these are likely to be the Jacksonville area, the Tampa/St. Petersburg/ Clearwater/Saratoga area, and the West Palm Beach/Greater Miami combination.) Then, zip code boundary overlays of the 1/2" to 1 mile county maps containing these urban areas would be prepared. The best routing and probable mid-day, off-peak speed along these routings would be estimated and converted into interzonal travel times. Nominal terminal times would be added and a matrix of inter-zip travel times for these urban areas would be manually constructed. These values would be merged with other inter-zip zonal travel times developed in the first alternative.

Using the location model, IRS analysts could thus evaluate the solutions obtained from three alternatives:

1) mechanically derived impedances for the whole state;

2) alternative (1) with manually derived travel times among zip zones in large urbanized areas, and mechanically derived impedances in other areas of the state; and

3) Euclidean distances provided originally.

APPENDIX B: SAMPLE OF PRINTED OUTPUT

For the sake of completeness and instruction, output from a run of the system on Jacksonville, Florida data is provided in this appendix. Many of the input data values were fictitious, so the results are not realistic.

```
**********************************************************************
***                                                                ***
***                IRS POST-OF-DUTY LOCATION SYSTEM                ***
***                                                                ***
***                  REPORT OF SESSION BEGINNING:                  ***
***                       4/25/86    9: 8                          ***
***                                                                ***
**********************************************************************


**********************************************************************
***                                                                ***
***                     PROBLEM INITIALIZATION                     ***
***                                                                ***
**********************************************************************


********************** INITIAL POD INFORMATION ***********************
```

POD	TYPE	OPENING COST	CLOSING COST	OFFICE RENTAL COST
32018	CURRENT	0.00	1.00	9.54
32202	CURRENT	0.00	1.00	8.93
32301	CURRENT	0.00	1.00	13.25
32401	CURRENT	0.00	1.00	8.99
32501	CURRENT	0.00	1.00	10.35
32601	CURRENT	0.00	1.00	8.53
32670	CURRENT	0.00	1.00	8.30
32771	POTENTIAL	1.00	0.00	10.00
32801	CURRENT	0.00	1.00	13.84
32901	CURRENT	0.00	1.00	8.23
33130	CURRENT	0.00	1.00	11.36
33169	CURRENT	0.00	1.00	13.25
33173	CURRENT	0.00	1.00	0.00
33174	POTENTIAL	1.00	0.00	13.00
33301	CURRENT	0.00	1.00	11.19
33401	CURRENT	0.00	1.00	0.00
33432	POTENTIAL	1.00	0.00	16.00
33450	CURRENT	0.00	1.00	10.13
33583	CURRENT	0.00	1.00	10.68
33602	CURRENT	0.00	1.00	11.95
33616	POTENTIAL	5.00	0.00	0.00
33702	FIXED	0.00	1.00	10.06
33801	CURRENT	0.00	1.00	6.10
33907	CURRENT	0.00	1.00	10.89

```
************************** USER OPTIONS ***************************

#        ITEM                        POSSIBLE             ACTUAL
```

1	IMF EXAMINATION	0,1	1
2	IMF COLLECTION	0,1	1
3	IMF TAXPAYER SERVICE	0,1	1
4	IMF CRIMINAL INVESTIGATION	0,1	1
5	BMF EXAMINATION	0,1	0
6	BMF COLLECTION	0,1	0
7	BMF TAXPAYER SERVICE	0,1	0
8	BMF CRIMINAL INVESTIGATION	0,1	0
9	TAXPAYER TRAVEL COST WEIGHT	[0.0-1.0]	1.00000000
10	IRS TRAVEL COST WEIGHT	[0.0-1.0]	1.00000000
11	MAXIMUM TRAVEL DISTANCE IN MILES		80.00000000

1 = INCLUDED IN COST CALCULATION, 0 = NOT INCLUDED

********************** TRAVEL DIFFICULTY FACTORS *********************

ZIP CODE	POD ZIP CODE	FACTOR
33557	33702	3.0
33615	33702	5.0

**************************** TRIP FACTORS ****************************

IRS TRIP FACTORS	TAXPAYER TRIP FACTORS
0.1040	0.9360
0.1620	1.3460
0.0310	1.3060
0.0860	1.6280
0.0690	1.7550
0.5110	1.1250
0.4880	0.7870
0.7080	0.8850
1.3460	0.6400
0.3020	0.9020
0.5150	0.8240
1.3000	0.5340
3.0000	0.7500
2.0000	0.3000
0.0000	0.0002
1.2000	0.0000

************************** OTHER PARAMETERS **************************

OFFICE SPACE REQUIREMENT = 0.062 SQUARE FEET PER IMF RETURN

TRAVEL COST = $ 0.205 PER MILE

```
************************************************************************
***                                                                  ***
***                            SOLUTION                              ***
***                                                                  ***
************************************************************************
```

COST = $ 156166508.72 (Note: Fictitious result due to sample problem data)

ZIP CODES ASSIGNED TO POD 32018:

32005	32010	32012	32014	32015	32016	32017	32019	32020	32021
32022	32023	32028	32030	32032	32033	32036	32037	32045	32057
32069	32074	32080	32081	32088	32090	32230	32617	32624	32631
32632	32633	32634	32654	32660	32662	32663	32664	32681	32682
32690	32706	32709	32710	32720	32722	32732	32733	32734	32740
32744	32747	32759	32763	32764	32768	32777	32790	32793	32798
32814	32816	32853	32854	32855	32856	32858	32861	32959	33848
33858									

ZIP CODES ASSIGNED TO POD 32202:

32007	32009	32011	32034	32040	32043	32046	32058	32063	32068
32072	32073	32079	32082	32084	32087	32097	32201	32203	32204
32205	32206	32207	32208	32209	32210	32211	32212	32215	32216
32217	32218	32219	32220	32221	32222	32223	32224	32225	32226
32227	32229	32233	32234	32244	32250	32265	32267	32602	32616
32658									

ZIP CODES ASSIGNED TO POD 32301:

32013	32053	32059	32302	32303	32304	32305	32306	32307	32308
32309	32311	32312	32321	32322	32323	32324	32327	32330	32331
32332	32333	32334	32336	32337	32340	32343	32344	32346	32347
32350	32351	32352	32355	32356	32357	32358	32360	32361	32362
32423	32442	32460	32466						

ZIP CODES ASSIGNED TO POD 32401:

32320	32328	32335	32403	32405	32407	32409	32410	32420	32421
32422	32424	32425	32426	32427	32428	32430	32431	32432	32433
32437	32438	32439	32440	32443	32444	32445	32446	32448	32449
32452	32453	32454	32455	32456	32459	32461	32462	32463	32464
32465	32537	32538	32544	32564					

ZIP CODES ASSIGNED TO POD 32501:

32434	32503	32504	32505	32506	32509	32510	32511	32512	32522
32530	32531	32533	32535	32536	32541	32542	32560	32561	32563
32565	32567	32568	32569	32570	32577	32578	32579	32580	

ZIP CODES ASSIGNED TO POD 32601:

32008	32031	32038	32042	32044	32047	32048	32049	32052	32054
32055	32060	32061	32062	32066	32071	32077	32083	32089	32091
32094	32096	32359	32615	32618	32619	32621	32622	32626	32628
32635	32638	32640	32643	32648	32656	32666	32667	32669	32680
32683	32685	32692	32693	32694	32696	32697	32630		

ZIP CODES ASSIGNED TO POD 32670:

32039	32093	32620	32625	32627	32629	32636	32637	32639	32642
32645	32646	32647	32649	32650	32659	32661	32665	32668	32672
32673	32679	32684	32686	32691	32695	32698	32702	32731	32735
32748	32762	32785	33502	33513	33514	33521	33524	33526	33531
33536	33538	33554	33556	33585	33593	33604	33802	33849	

ZIP CODES ASSIGNED TO POD 32801:

32701	32703	32705	32707	32708	32711	32713	32725	32726	32729

32730	32737	32741	32745	32746	32750	32751	32753	32754	32755
32756	32757	32760	32761	32765	32766	32767	32769	32771	32775
32776	32778	32780	32784	32786	32787	32789	32792	32797	32803
32804	32805	32806	32807	32808	32809	32810	32811	32812	32813
32817	32820	33503	33530	33550	33587	33601	33623	33685	33835
33863									

ZIP CODES ASSIGNED TO POD 32901:

32739	32903	32905	32922	32925	32935	32937	32950	32951	32955
32957	32958	33438							

ZIP CODES ASSIGNED TO POD 33169:

33009	33010	33012	33013	33014	33015	33016	33019	33021	33023
33024	33025	33026	33027	33028	33029	33054	33055	33056	33126
33127	33132	33137	33138	33139	33140	33141	33147	33150	33154
33160	33161	33162	33167	33168	33179	33180	33181	33331	33332

ZIP CODES ASSIGNED TO POD 33173:

33001	33030	33031	33032	33033	33034	33035	33036	33037	33039
33051	33070	33122	33125	33128	33129	33130	33133	33134	33135
33136	33142	33143	33144	33145	33146	33155	33156	33157	33158
33165	33166	33170	33172	33174	33175	33176	33177	33178	33182
33183	33184	33185	33186	33187	33189	33190	33192	33193	33196
33925	33943								

ZIP CODES ASSIGNED TO POD 33301:

33004	33020	33060	33062	33063	33067	33068	33304	33305	33306
33308	33309	33311	33312	33313	33314	33315	33316	33317	33319
33322	33323	33324	33325	33326	33328	33330	33334		

ZIP CODES ASSIGNED TO POD 33401:

33403	33404	33405	33406	33407	33408	33409	33410	33411	33430
33440	33455	33458	33460	33461	33462	33463	33470	33480	

ZIP CODES ASSIGNED TO POD 33432:

33011	33022	33061	33064	33065	33066	33116	33124	33151	33153
33163	33164	33194	33197	33302	33303	33307	33318	33320	33327
33335	33338	33339	33431	33433	33434	33435	33436	33437	33441
33444	33445	33446	33493						

ZIP CODES ASSIGNED TO POD 33450:

32948	32949	32960	32970	32971	33402	33439	33456	33457	33472
33476	33490	33491	33492	33494	33852	33857			

ZIP CODES ASSIGNED TO POD 33583:

33504	33505	33506	33507	33508	33509	33517	33522	33523	33529
33532	33533	33545	33548	33551	33555	33558	33559	33560	33561
33564	33577	33578	33580	33581	33582	33586	33588	33595	33596
33730	33734	33736	33737	33738	33739	33740	33821	33842	33864
33865	33890	33902	33920	33927	33938	33945	33946	33952	33953

ZIP CODES ASSIGNED TO POD 33602:

33511	33512	33527	33534	33539	33547	33549	33552	33553	33557
33568	33569	33570	33574	33576	33584	33592	33594	33598	33603
33605	33606	33607	33608	33609	33610	33611	33612	33614	33615

33616 33617 33618 33619 33620 33621

ZIP CODES ASSIGNED TO POD 33702:
33515 33516 33519 33520 33528 33535 33540 33541 33542 33543
33563 33565 33572 33589 33590 33591 33701 33703 33704 33705
33707 33709 33710 33711 33712 33713 33714

ZIP CODES ASSIGNED TO POD 33801:
32736 33525 33537 33566 33597 33599 33803 33805 33820 33823
33825 33827 33830 33834 33837 33838 33839 33840 33841 33843
33844 33846 33847 33850 33851 33853 33854 33855 33856 33860
33866 33868 33870 33873 33877 33880

ZIP CODES ASSIGNED TO POD 33907:
33459 33471 33901 33903 33904 33905 33908 33922 33923 33926
33928 33929 33930 33931 33933 33934 33935 33936 33937 33939
33940 33944 33950 33954 33955 33956 33957 33960

NBS-114A (REV. 2-80)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR 86-3482	2. Performing Organ. Report No.	3. Publication Date FEBRUARY 1987

4. TITLE AND SUBTITLE

The Internal Revenue Service Post-of-Duty Location Modelling System - Final Report

5. AUTHOR(S)

Paul D. Domich, Karla L. Hoffman, Richard H. F. Jackson, Marjorie A. McClain

6. PERFORMING ORGANIZATION (*If joint or other than NBS, see instructions*)

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (*Street, City, State, ZIP*)

The Research Division
U. S. Internal Revenue Service
1201 E Street
Washington, DC 20224

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (*A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here*)

This report documents a project undertaken by the National Bureau of Standards to develop a mathematical model which identifies optimal locations of Internal Revenue Service Posts-of-Duty. The mathematical model used for this problem is the uncapacitated, fixed charge, location-allocation model which minimizes travel and facility costs, given a specified level of activity. The report includes a discussion of the location problem and the mathematical model developed. Data sources identified and used are also described. Brief descriptions of the mathematical techniques used and the interactive, user-friendly computer system built to solve the problem are also provided. The system is microcomputer-based and uses menus and graphically displayed maps of tax districts for interactive inputs and solution outputs.

12. KEY WORDS (*Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons*)

facility location, uncapacitated, location-allocation, plant location, warehouse location, fixed-charge, personal computer, microcomputer, interactive graphics, greedy algorithm, heuristic algorithm.

13. AVAILABILITY

[X] Unlimited
☐ For Official Distribution. Do Not Release to NTIS
☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
[X] Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES

37

15. Price

$11.95