FINAL REPORT

CLAYTON COUNTY SCHOOLS

Analysis of Attendance and Facilities Utilization

David J. Burkitt
Robert J. Graves

December 15, 1977

School of Industrial and Systems Engineering
Georgia Institute of Technology
E-24-658
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Problem Statement</td>
<td>1</td>
</tr>
<tr>
<td>1.2 State-of-Practice</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Literature</td>
<td>4</td>
</tr>
<tr>
<td>2.0 MATHEMATICAL MODELS</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Elementary School Model</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Junior High School Model</td>
<td>7</td>
</tr>
<tr>
<td>2.3 High School Model</td>
<td>8</td>
</tr>
<tr>
<td>3.0 SYSTEMS MODEL</td>
<td>8</td>
</tr>
<tr>
<td>3.1 Data Requirements</td>
<td>10</td>
</tr>
<tr>
<td>3.2 Operations</td>
<td>10</td>
</tr>
<tr>
<td>4.0 RECOMMENDATIONS</td>
<td>11</td>
</tr>
<tr>
<td>4.1 Phase I: Interactive Model</td>
<td>12</td>
</tr>
<tr>
<td>4.2 Phase II: Interactive-Optimization Model</td>
<td>13</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>15</td>
</tr>
</tbody>
</table>
1.0 INTRODUCTION

This report is the summary of activities performed on the internship project entitled "Analysis of Attendance and Facilities Utilization" and jointly sponsored by Clayton County Schools and Georgia Tech. The tasks performed during this project include the analysis of current procedures for enrollment projection and school assignment decisions at Clayton County Schools, at other local school systems and, where feasible, similar issues reported by the Department of Health, Education and Welfare and the National Institute of Education. The result of this analysis has led to the preliminary development of recommendations for re-design of the system for enrollment projections and school boundary determination in Clayton County Schools.

The general outline of this report begins with a brief statement of the problem facing school system decision-makers which is then followed by a presentation of the state-of-practice in this problem area and the related literature. Mathematical models for solution of the problem at the elementary school level, junior high level and the high school level are presented and discussed along with a systems model demonstrating information flows. Finally, recommendations for implementation are developed with concentration upon a two-phase approach for the computer model.

1.1 Problem Statement

The problem of school boundary design is a recurring one for educational decision-makers due to changing populations and population locations. Shifts in boundaries must be weighed carefully against the educational and social issues involved and the facilities utilization pressures. Thus, several major criteria appear as pertinent to the problem:
i) minimize the total distance that students travel to school;
ii) minimize the number of students assigned to a school different from their present school;
iii) maximize the number of students assigned to the closest school;
iv) minimize the number of students crossing interstate highways.

At the same time, several constraining conditions exist:

i) no school is to exceed its maximum enrollment figure;
ii) no school should fail to exceed its minimum enrollment figure;
iii) students within walking distance of a school should be assigned to that school;
iv) high school students should continue to be assigned to their present school.

These concerns about the problem reflect the difficulties of balancing the needs for efficiency (facilities utilization, transportation costs), for safety (walking routes) and for effectiveness (educational impact of switching current student assignments). Other school systems might add to or delete from such a list of concerns based on their own particular needs.

A number of knowledge bases are required for solution to the problem. Not all of this knowledge can be explicitly modeled however. Understanding of the physical characteristics of Clayton County is needed to gain insight into growth patterns and natural barriers. Familiarity with street and school locations would be needed when addressing the problem as well as the locations of residential, industrial and commercial areas. Knowledge of traffic flow patterns would assist in the determination of those elements related to transportation and pedestrian safety. Specific data on students' residences, either from historical or projected perspectives, is necessary for the final estimation of student flows into schools.
The following report section discusses the current procedures for resolving the problem at Clayton County Schools and presents results of examination of several other school districts' procedures.

1.2 State-of-Practice

In the last year, the process of setting school boundaries has progressed from an entirely manual process to a combination of machine and manual methods. Using the master student file and a file of land lots of the county, computer programs now give a partial student population of each land lot. The students living on streets crossing multiple land lots are pinpointed by manual methods. From these land lot population figures, the implications of school boundaries conforming to land lot boundaries can be easily assessed. However, since most school boundaries divide land lots, manual methods are used to determine effects on the individual school populations. The process is one of using the expert knowledge of school officials to set tentative boundaries, subsequently checking school populations using both the computer and manual methods, revise school boundaries, and continue the process until satisfactory results are obtained. The major drawback to this method is the time and effort needed to check the effects of tentative boundaries. Furthermore, the computer can not be utilized to give insight into other complicated boundary shifts.

Other school districts use similar methods for setting boundaries depending on the available data base and computer facilities. Those systems having a small data base use the knowledge of school principals pertaining to the individual school population distribution to set boundaries by hand tabulating methods. This is a time-consuming process which produces populations whose accuracy depends on the principals' familiarity with their areas.
Large school districts with extensive computer facilities usually have the data base available to use optimization programs to devise the best boundaries with the objective of minimizing student travel. Furthermore, with sophisticated graphic software, machine methods are used to produce maps to aid in the decision process. Even at this level of computer sophistication, the final boundaries are developed using both objective and subjective information.

1.3 Literature

The current technical literature on the school boundary problem uses operations research optimization techniques to minimize student distance traveled with the constraints of school capacities and, in many articles, racial balance. The problem is easily related to a class of allocation problems in operations research. The models are constructed either as linear programming models or as network models. The models are similar while using different techniques in the solution process. In both cases, methods are used to insure that a student location is not split with part being assigned to one school while the other part is assigned to a different school. The linear programming formulation has been further investigated in an attempt to use its dual formulation for sensitivity analysis. In addition to providing the optimal boundaries, some analysis of "what if" questions could be performed. Some working papers have dealt with the interactive aspect of the model using optimization techniques. The interactive models tend to use highly sophisticated software. The objective of explicitly minimizing transfers from present school assignments has been ignored in the literature. So many other aspects of education are difficult to quantify and therefore are also omitted from the models. The
assumption is that the school administrators will subjectively insert these aspects into the final solution to obtain the "best" boundaries.

2.0 MATHEMATICAL MODELS

Mathematical models have been developed for the elementary, junior high and high school levels at Clayton County. These different models reflect the different objectives at work in school boundary design for these school levels. The effort to develop mathematical models of the problem is a basic step in the operations research technique. Through careful analysis and communication, the operations research analyst and the decision-maker can examine, in an explicit way, the types of objectives to be pursued in solving the problem. Explicit development and statement of constraints to the problem serves to further reflect the realities of a decision-maker's problem solving. Thus, the following models represent the culmination of this modeling effort to date, although further refinements may occur as implementation takes place.

The basic physical unit examined by the model is that of a land parcel. Each parcel is a geographic unit; a collection of students identified by their home location. The parcel boundaries may be natural features such as streets, rivers and utility rights-of-way or artificial features such as present school boundaries and lot lines. These parcels are to be determined using both manual means and automated methods with data such as that of the DIME file from the Atlanta Regional Commission. Thus, a location variable for the elementary school model is defined as $x_{ij}$ which assumes a value of zero if parcel $i$ is not assigned to school $j$ and a value of one if parcel $i$ is assigned to school $j$. This variable is redefined for the junior high and high school model to $x_{ijk}$ which is zero if grade $k$ in parcel
i is not assigned to school j and becomes one if grade k in parcel i is assigned to school j.

2.1 Elementary School Model

Define the following:

A: a constant which assumes a value from 0 to 1 as specified by the user

d_{ij}: the distance from the center of parcel i to school j

t_{ij} = \begin{cases} 
0 & \text{if parcel i is presently assigned to school j} \\
M & \text{(some large positive number)} \quad \text{otherwise}
\end{cases}

N: total number of parcels in the county

6

P_i = \sum_{k=1}^{6} p_{ik} \quad \text{where} \quad p_{ik} = \text{the number of students presently in grade k residing in parcel i}

l_j = \text{the least number of students to be assigned to school j}

u_j = \text{the maximum number of students to be assigned to school j}

then the mathematical model assumes the following form

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i=1}^{N} P_i \left[ \sum_{j=1}^{25} \left( A d_{ij} + (1 - A) t_{ij} \right) x_{ij} \right] \\
\text{subject to:} & \quad l_j \leq \sum_{i=1}^{N} p_i x_{ij} \leq u_j \quad \text{for} \quad j = 1, 2, \ldots, 25 \\
& \quad \sum_{j=1}^{25} x_{ij} = 1 \quad \text{for} \quad i = 1, 2, \ldots, N \\
& \quad x_{ij} = 0 \text{ or } 1 \quad \text{for all} \quad i, j
\end{align*}
\]

The objective, as specified in (1) is to minimize a weighted distance
function with additional penalties for changing a parcel's current assignment. The decision-maker specifies the relative importance between distance and parcel reassignment through the parameter A and the penalty value M.

The constraint set serves to limit the attainment of the objective. Constraints specified in (2) keep the population assigned to each school between the specified lower and upper bounds. Constraint set (3) causes each parcel to be assigned to exactly one school and the constraints in (4) cause the assignment of entire parcels to one school thus preventing the model solution from splitting parcels.

2.2 Junior High School Model

The model for junior high schools is similar to that for elementary schools except for increased attention to the grade levels as evidenced by the introduction of subscript k.

Define the following:

- \( p_{ik} \): number of students in grade k residing in parcel i
- \( \lambda_{jk} \): least number of students in grade k to be assigned to school j
- \( u_{jk} \): maximum number of students in grade k to be assigned to school j

then the mathematical model is specified as

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i=1}^{N} \sum_{j=1}^{10} \left\{ \sum_{k=7}^{9} p_{ik} x_{ijk} [A d_{ij} + (1 - A) t_{ij}] \right\} \\
\text{subject to:} & \quad \lambda_{jk} \leq \sum_{i=1}^{N} p_{ik} x_{ijk} \leq u_{jk} \quad \text{for } j = 1, 2, \ldots, 10 \\
& \quad \text{and } k = 7, 8, 9
\end{align*}
\]
\[ \frac{x_{ijk}}{p_{ik}} \leq \frac{N}{9} \sum_{i=1}^{10} \sum_{k=7}^{9} p_{ik} x_{ijk} \leq u_{j} \quad \text{for } j = 1, 2, \ldots, 10 \quad (7) \]

\[ \sum_{j=1}^{10} x_{ijk} = 1 \quad \text{for } i = 1, 2, \ldots, N \quad (8) \]

\[ \text{and } k = 7, 8, 9 \]

\[ x_{ijk} = 0 \text{ or } 1 \quad \text{for all } i, j, k \quad (9) \]

Interpretation of the objective and the constraints is similar to that described earlier except now the concerns must hold for each grade level 7, 8 and 9 as distinguished from an aggregated grade level represented by \( P_i \) in the earlier model. Constraint set (6) places restrictions on each grade level's size in each school and constraint set (7) restricts the overall size of the school.

2.3 High School Model

The high school model is identical to the junior high model in all respects except for two indices. For \( j \), the number of schools involved, is limited to five reflecting the current number of high schools. Index \( k \), the grade levels, varies from ten through twelve rather than seven through nine.

3.0 SYSTEMS MODEL

A tentative systems model is presented in Figure 1 to display the types of data files which will be necessary and the operations to be performed on the data. Computer codes must be planned and developed to perform the data transfer and manipulation necessary in the implementation of the model.
Figure 1 Systems Model

- DIME
  - REDUCE
    - PENNY
      - COMBINE
        - SCHOOL
          - PATH
            - DISTANCE
          - PARCEL
            - GROUP
            - SDF
            - SMF
        - ASSIGN
          - USER
3.1 Data Requirements

The data required is the student population characterization of each parcel, the location and capacities of the schools, and a distance measure between the parcels and the schools. The data base required for this information will consist of the master student file, the school facility file, and the DIME file with various programs to actually construct the data.

The master student file contains information about each student and his/her location in the county. The information consists of student name, number, race, and sex. The student's present address, grade, and school assignment are also on the file. This file is in use at the present time by the data processing department for attendance purposes and is updated continually to keep it as accurate as possible.

The school facility file contains the location and capability of each school in the county. It also contains types of utilization of each school. This data file does not presently exist in computer form and therefore would have to be constructed.

The DIME file contains a network representation of the streets in the county. It contains information about the nodes and edges of the network including state plane coordinates useful in distance computation. This file is being developed by the Atlanta Regional Commission for the U.S. Census Bureau. The northern, most populous, two-thirds of the county will be available at the end of 1977 with the remainder of the county being completed during 1978.

3.2 Operations

Using the data base, computer programs will construct intermediate data files for use in the model. These include:
Penny - contains a reduced DIME file for efficient operation
Parcel - contains the characterization of each parcel
Distance - contains the distance from each parcel to each school
Student Distribution File - contains the student population characteriza-
tion of each parcel

The programs constructing the intermediate data files include:
REDUCE - a program to delete unnecessary information from DIME file
COMBINE - an interactive program to partition network into parcels
PATH - shortest path algorithm to construct distances from parcels to schools
GROUP - a program to match a student's present address to a unique parcel
ASSIGN - an interactive program to assign parcels to schools.

The interaction of the user with the model not only includes his involvement in ASSIGN but also in his capability to easily change the intermediate data files to answer "what if" questions.

4.0 RECOMMENDATIONS

The completion of the study of this problem, including experience from the experimental computer model developed during the first half of the current project, leads to recommendations for further effort toward the overall problem solution. The research and development is far from complete at this stage, but the efforts of this project have resulted in a clearer view of the path toward completion.

The implementation of the model is recommended to be a natural stepwise procedure building toward the comprehensive model. This approach will
yield valuable knowledge at intermediate points such as data integrity, system design and user interface problems while also yielding useful decision-making tools at the milestone points. The first phase is to lead to development of an interactive model for decision-makers. Such a model would be reflective in the sense that "what if" questions might be asked by the decision-maker and the model response would be that of computing the implications of the question and reporting such results back to the decision-maker. The second phase model also includes an optimization component which means that optimum (in the sense of explicitly modeled criteria) school boundaries could be determined for the decision-maker in response to hypothetical questions such as different fixed boundary sets at initiation.

4.1 Phase I: Interactive Model

The main effort of Phase I is to concentrate upon the development of an interactive model. A significant task will be that of data base design, construction and management. This task must be performed with concern for the flexibility of its use by the Phase II model as well. A second major task involves the design of a computerized distance algorithm which will closely approximate the pedestrian distance traveled by each student to his/her assigned school. Determination of the land parcels is required and a capacity for projection or forecasting of student populations and locations should be included.

An important element of the code development is the user interface design. Interaction at a remote terminal is seen as the proper approach at this stage with the decision-maker "what if" questions posed in the form of changing parcel-to-school assignments on an existing boundary solution. Responses to such questions could include such data as:
i) Population characteristics at each school such as grade, sex, and race or the incremental change in such characteristics from the current physical situation or from the last computer iteration;

ii) The average and/or maximum distance traveled by students;

iii) The number of students assigned to their closest school, second closest, ...;

iv) The number of students traveling further than some specified distance or distances, e.g. further than 1.5 miles;

v) The number of students within some specified distance of each school.

The decision-maker would examine the changes in statistics such as these as questions are posed to the model and ultimately arrive at a school boundary decision. Thus the computer model essentially performs the service of a computational aid.

Implementing Phase I will serve to establish important information pointing toward the comprehensive model development in Phase II. Data base verification and maintenance procedures should be completed and in place before Phase II begins. System design and the user interface should not only be in place but actual use experience available for potential modification to the system.

4.2 Phase II: Interactive-Optimization Model

The major additional development to be performed in Phase II is the inclusion of an intelligent solution procedure within the computer code so that the decision-maker can use the model to efficiently assign parcels to schools according to the explicit criteria and constraints of the model. This more comprehensive development places a powerful tool at the
administrator's hand to aid in this pupil assignment problem solution without reducing the model's capacity to accept implicit planner judgement through the "what if" questions in an interactive mode.

The model is envisioned as one providing the administrator with the following:

i) Optimal boundaries constructed from the explicit criteria as a benchmark solution;

ii) Optimal boundaries constructed with user specified student transfer parameters;

iii) Optimal boundaries for situations of new facilities expansion or facilities contraction decisions;

iv) Optimal assignment shifts in response to changed boundary specifications by the user,

v) Optimal assignment shifts due to user specified changes in the model constraints such as new barriers to student travel or new parcel developments.

It is intended that the Phase II model would provide new insights to administrators to aid in the decision-making about complicated planning decisions, while easing the manual workload required.
REFERENCES


