CO 370 Deterministic OR Models

UNIVERSITY OF WATERLOO

Winter 2010

Optimal Allocation of Postal Distribution Centers in the Waterloo County

Group Number: **2**

Project Members:

Sachin Mehta
Tiffany Cheuk
Rajesh Swaminathan
Thanusan Ehamparanathan

March 31, 2010

Table of Contents

1 Executive Summary

The county of Waterloo wants to build additional postal distribution centers in anticipation of a population rise due to the booming tech industry. The county is working on a tight budget for this mega project. The county has done some preliminary analysis taking into consideration land surveys and locations of present distribution centers. They have proposed a set of potential locations where new distribution centers can be built.

The county would like to determine which subset of these locations should be chosen to build new distribution centers on to satisfy the projected demand of the county's numerous postal code regions in an optimal manner keeping in mind the county's stringent budget limitations. Furthermore, the maximum distance traveled by any postman to and from these distribution centers must be minimized to reduce operational costs in the long run.

2 Problem Definition

There are J proposed locations for new distribution centers, and I postal code regions that need to be serviced by these new distribution centers. Each postal code region should be associated with exactly *one* distribution center since all mails from a given postbox in a given postal code region travel in the same sack to the same distribution center. Each distribution center, in contrast, can service multiple postal code regions.

The goal is to determine which subset of proposed locations must be chosen to build new distribution centers on, and to assign each postal code region to exactly one of these new distribution centers that is going to be built. The size of the chosen subset will be at most I since each postal code region can be associated with just one distribution center.

The anticipated population of each postal code region i in the next few years is p_i and will be provided by the county's town center. The cost of building each distribution center is a function f of its size which depends on the number of people the center services. f is defined piecewise as follows:

$$f_j(x) = \begin{cases} 3x, & \text{if } 0 \leq x \leq 4 \\ 12 + 2(x-4), & \text{if } 4 \leq x \leq 8 \\ 20 + 4(x-8), & \text{if } x \geq 8 \end{cases}$$

where x is the number of people (in hundreds) being serviced by distribution center j.

This cost function f is a piecewise linear function where the cost per person serviced initially decreases due to economies of scale, but starts to increase beyond a certain threshold (800 in this case) due to management and communication overhead involved. Furthermore, the county wants to discourage huge distribution centers from being built so as to alleviate the risk from infrastructural failures such as power and internet failures.

The objective is to minimize the operating costs once all new distribution centers have been built. The operating cost is primarily determined by the gas needed for the mail delivery truck to make its routes everyday, which is directly dependent on the distance traveled to and from the distribution centers. Other operating costs that are also distance-dependent are truck maintenance costs (paint, engine oil, etc.) and tire/brake replacement costs. Thus we want to minimize the aggregate maximum distance that postmen have to travel between post boxes within a postal code region and the single distribution center that is assigned to them. Minimizing travel distance will reduce gas costs and other maintenance costs significantly in the long run.

The total budget for this project provided by the county of Waterloo for the construction of new distribution centers is B dollars. The project is not allowed to exceed its budget.

The decision of which locations to build new distribution centers on must also adhere to certain additional constraints. These constraints are as follows:

1. To be economically feasible, a distribution center, if built, must service at least 200 people.
2. Certain postal code regions in the Downtown area receive significantly more business-related mail compared to other regions. Downtown is therefore a high-congestion region for mail pickup and delivery. Forecasts predict that locations 1, 5, and 8 or locations 2, 5, and 9 can service all of Downtown while still delivering adequate service to the nearby regions. Thus, one of these combinations of locations must be chosen.
3. The county of Waterloo does not want two distribution centers in close proximity to each other. The county wants to discourage distribution centers from being built too closely to one another so as to alleviate the risk of infrastructural failures such as power and network failures. This means if a distribution center at location j is built, then we are not allowed to build a distribution center at either sites $j-1$ or $j+1$, assuming the locations have been provided in an order sorted by proximity.
4. Because the size of a distribution center is limited by the lot size at its location, there is a cap on the number of people that distribution center can service. (The more people a distribution center services, the bigger it needs to be.) Thus the lot size available at location j limits the number of people the j^{th} distribution center can service.

3 Model Formulation

We model the above optimization problem as a mixed integer linear programming problem (MIP). The following sections specify the mathematical model in detail by describing the variables, objective function and constraints of the model.

3.1 Decision Variables

We introduce the following $\{0, 1\}$ variables:

1. Let y_j be J binary variables representing whether we build a distribution center at location j or not.
2. Let x_{ij} be $I \times J$ binary variables representing whether postal code region i is serviced by distribution center j or not.

For convenience, let K be the set containing the integers from 1 to I (inclusive), and L be the set containing the integers from 1 to J (inclusive).

3.2 Constraints

1. Each postal code region i is to be serviced by exactly one distribution center. Thus $\sum_{j \in L} x_{ij} = 1, \forall\ i \in K$.

2. A postal code region i is not allowed to be serviced by a location j that will not contain a distribution center. This means that if we choose to build a distribution center at location j, then we can have at most I postal code regions serviced by it. But if we choose to not build a distribution center at location j, then this location is not allowed to service any postal code regions. Thus the constraint can be modeled as a big-M constraint like so: $\sum_{i \in K} x_{ij} \leq y_j I, \forall\ j \in L$. This formulation introduces $|L| = J$ constraints.

   The value of M in this big-M constraint needs to be I (which is usually a big number) since we need to allow for the possibility of a single distribution center servicing *all* postal code regions. This makes the formulation weak due to the large feasible region permitted by a large I. However, if we limited the number of postal code regions serviced by a single distribution center to say 10, the value of M in this big-M constraint would be reduced yielding a stronger formulation closer to the convex hull of the MIP.

   An alternate formulation for this constraint is to add a constraint for each pair of postal code region and distribution center location: $x_{ij} \leq y_j$, for all $i \in K$ and for all $j \in L$. These $I \times J$ constraints enforce the implication that if a postal code region i is assigned to distribution center j, then distribution center j *must* be built. In other words, $x_{ij} = 1 \Rightarrow y_j = 1$.

   Given that both formulations are valid IP formulations, they both contains the same set of integer solutions. However, which one is stronger? We argue that the second formulation is stronger since the feasible region of its LP relaxation is a *strict* subset of the feasible region of the first formulation's LP relaxation. This can be seen by noticing that every feasible solution to the second formulation is also feasible for the first (the first formulation is simply the *aggregated* version of the second), but the solution ($y_1 = 0.5, x_{11} = 1, x_{i1} = 0\ \forall\ i > 1$) for $I \geq 2$ is feasible for the first formulation's LP relaxation but not for the second. Thus the optimal value to the relaxation of the second formulation will give us a closer approximation to the optimal solution of the original IP problem than that of the relaxation of the first formulation.

Nonetheless, we will continue to use the weaker first formulation in our model since the stronger second formulation introduces a considerably larger number of constraints than the first. Because we are going to be using a student solver that is limited in the number of constraints that can be used in the input, we want to use as few constraints as possible in our model formulation.

3. We cannot go over the stated budget of B dollars (in hundreds) allocated for this project. Therefore, the cost of building all new distribution centers must be less than or equal to the allocated budget. The cost of each distribution center will depend on its size which in turn depends on the number of people it will service. The number of people each distribution center services depends on the projected population of the postal code regions that are assigned to it.

   Let s_j be the number of people (in hundreds) serviced by distribution center j. Then $s_j = \sum_{i \in K} x_{ij} p_i,\ \forall\ j \in L$, where p_i is the projected population of postal code region i (in hundreds).

   Now, $f(s_j)$ is a piecewise-linear function that represents the cost of building a distribution center that services s_j people.

   Thus the budget constraint is specified as

   $$\sum_{j \in L} f(s_j) \leq B$$

   How do we model this piecewise function $f(s_j)$?

   Let us express s_j as $s_j = v_1(j) + v_2(j) + v_3(j)$, where

   $v_1(j)$ corresponds to the amount s_j exceeds 0, but is less than or equal to 4
   $v_2(j)$ corresponds to the amount s_j exceeds 4, but is less than or equal to 8
   $v_3(j)$ corresponds to the amount s_j exceeds 8.

   Clearly, $0 \leq v_1(j) \leq 4$, $0 \leq v_2(j) \leq 4$, and $0 \leq v_3(j) \leq \sum_{i \in K} p_i - 8$.

   So $f(s_j)$ is therefore $3v_1(j) + 2v_2(j) + 4v_3(j)$.

   To enforce the correct relationship between the v_1, v_2 and v_3 variables, we introduce binary variables $u_1(j)$ and $u_2(j)$. We now need to encode the fact that:

   - $v_2(j) > 0 \Rightarrow u_1(j) = 1 \Rightarrow v_1(j) \geq 4$ (and hence equal to 4 since $v_1(j) \leq 4$)

- $v_3(j) > 0 \Rightarrow u_2(j) = 1 \Rightarrow v_2(j) \geq 4$ (and hence equal to 4 since $v_2(j) \leq 4$)

To enforce the above two encodings, we use the following constraints:

(3a) $v_2(j) \leq 4u_1(j)$ forces $u_1(j)$ to be 1 if $v_2(j) > 0$

(3b) $v_1(j) \geq 4u_1(j)$ forces $v_1(j)$ to be ≥ 4 if $u_1(j) = 1$

(3c) $v_3(j) \leq 10u_2(j)$ forces $u_2(j)$ to be 1 if $v_3(j) > 0$

(3d) $v_2(j) \geq 4u_2(j)$ forces $v_2(j)$ to be ≥ 4 if $u_2(j) = 1$

Multiplying constraint (3a) above by -1 (causing the inequality to flip) and adding it to constraint (3d) gives us $u_1(j) \geq u_2(j)$. Thus $u_2(j)$ can be 1 only if $u_1(j)$ is 1.

4. In order to be economically feasible, a distribution center must service at least 200 people. Thus $s_j \geq 2y_j, \ \forall \ j \in L$.

5. Certain postal code regions in Downtown receive way more business-related mail compared to the other regions. Downtown is therefore a high-congestion region for mail pickup and delivery. Forecasts predict that locations 1, 5, and 8 or locations 2, 5, and 9 can service all of downtown while still delivering adequate service to the nearby regions. This can be modeled as an either-or constraint like so:

$$y_1 + y_5 + y_8 \text{ or } y_2 + y_5 + y_9$$

This may be modeled using an auxiliary binary variable k (where $k \in \{0, 1\}$):

$$y_1 + y_5 + y_8 \geq 3k$$

$$y_2 + y_5 + y_9 \geq 3(1 - k)$$

6. The county of Waterloo does not want two distribution centers in close proximity to each other. This means if distribution center at location j is built, then we are not allowed to build a distribution center at either locations $j - 1$ or $j + 1$. The proposed J locations have been provided to us sorted by proximity. Thus, location j is adjacent to locations $j - 1$ and $j + 1$ and so on. This constraint can be modeled as follows:

$$y_j + y_{j+1} \leq 1, \ \forall \ j \in \{1, \ldots, J - 1\}$$

7. Because the size of a distribution center is limited by the lot size at its location, there is a cap on the number of people that distribution center can service. (The more people a distribution center services, the bigger it needs to be.) Thus the lot size available at location j limits the number of people that the j^{th} distribution center can service.

   This constraint is $s_j \leq ls_j, \ \forall \ j \in L$, where ls_j is the maximum number of people (in hundreds) that distribution center j can service if it is built.

8. Non-negativity and binary constraints: $y_j \in \{0, 1\} \ \forall \ j \in L$, and $x_{ij} \in \{0, 1\} \ \forall \ i \in K, j \in L$.

3.3 Objective Function

Let dd_{ij} be the distance between the postbox in any postal code region i and any distribution center j. Let d_i be the distance between postal code region i and the single distribution center assigned to it. So $d_i = \sum_{j \in L} (dd_{ij})(x_{ij})$, for each i.

The maximum distance D traveled by any postman is then given by $\max_{i \in K} d_i$. Our objective is to minimize $T \times D$, where T is the cost of gas per kilometer traveled. The objective function minimize $\max(d_i)$ is equivalent to minimize(D) subject to $D \geq d_i, \ \forall \ i \in K$.

3.4 Full Model

The complete integer model with variables, constraints, and objective function for the postal distribution center allocation problem is stated in Appendix A.

3.5 Data Required

The following data will need to be collected/generated to solve the above IP formulation:

1. The number of postal code regions I
2. The number of proposed distribution center locations J
3. The projected population p_i of each postal code region i (in hundreds)
4. The distance dd_{ij} between postal code region i and distribution center j (in kms)
5. The cost function $f(x)$ (given) of building a distribution center that is capable of servicing at most x people (in hundreds)
6. The cost T of gas and maintenance and tire/brake replacement per kilometer traveled. Past data has shown this number is roughly 15 cents/km
7. The maximum number of people ls_j (in hundreds) that a distribution center at location j can service
8. The budget B of this entire undertaking. In this instance, the budget is 2.5 million dollars.

4 Model Assumptions

1. We assume that the demand for mail delivery and pick-up does not fluctuate seasonally. For example, during Christmas and New Year, the number of mails sent and received typically increases. We therefore ignore all stochastic uncertainty in mail traffic. This is a reasonable assumption to keep our model simple and allows us to assume fixed

demand throughout the year that is proportional to the population. This assumption is necessary to keep our model 100% deterministic.

For now, let p_i represent the population corresponding to the *maximum* demand experienced throughout the year. This will allow all demand to be satisfied throughout the year, but will result in occasional periods where the distribution centers are not working at maximum capacity.

2. We assume that the distance between the postal code regions and the proposed distribution center locations are fixed and do not change over time. This may not always be true since there will always be detours due to construction, bad weather conditions, and changes in street layouts with time.

   To model the uncertainty correctly, we could make the distances random variables with an associated probability distribution. The distribution can be specified as a collection of scenarios, and each scenario would have an associated probability with which it is likely to occur.

3. We assume that the population projections for the county in the next few years is accurate. We assume there is no variance or error in these projections. In reality, there almost always will be some error. This assumption is once again needed to keep our model simple and understandable. To model the uncertainty correctly, we would make p_i a random variable following a normal distribution with a mean of pm_i and a variance of pv_i.

5 Solution Methodology

We generated random data for this problem using a script in the PHP programming language. This script was able to generate data for arbitrary values of I and J and to produce an AMPL data file as output. Other than this, no special measures were taken to obtain a solution.

We used the Gurobi solver that comes part of the free AMPL student package to solve our mathematical model. The student edition of the solver limited us to 500 variables and 500 constraints. We therefore limited our problem to 495 variables and 224 constraints by setting $I = 22$ and $J = 15$.

6 Problem Solution

The problem was solved using the student edition of the Gurobi solver that comes part of the free AMPL package from `http://www.ampl.com`. The solution was obtained in less than a second on a commodity-grade Fujitsu laptop with a 1.6 GHz processor and 2 GB of

RAM. Sensitivity analysis was performed by turning on the `solnsens` parameter in the Gurobi solver.

The solution obtained is as follows:

$y_j = 0$, except $y_2 = y_5 = y_7 = y_9 = y_{11} = y_{13} = y_{15} = 1$. This implies we build 7 distribution centers at locations 2, 5, 7, 9, 11, 13, 15.

- Location 2 services regions 20 and 21
- Location 5 services regions 10, 18, and 22
- Location 7 services regions 5, 12, and 16
- Location 9 services regions 1, 2, 7, and 11
- Location 11 services regions 3, 4, and 19
- Location 13 services regions 13, 15, and 17
- Location 15 services regions 6, 8, 9, and 14.

The total cost of building these 7 distribution centers is $\sum_{j \in L} f(s_j) = 2.08$ million dollars.

The maximum distance traveled by any delivery truck is 14 kilometers per trip at a cost of $14 \times \$0.15 = \$2.1$ per trip.

7 Solution Analysis

7.1 Sensitivity Assumptions

In the following sections, we will assume that we will always have the necessary budget needed to fulfill all demand. We will also assume that we will have enough potential distribution center locations to fulfill all demand.

Since there is no formal sensitivity analysis that can be performed on IP formulations the way they can be performed on LP formulations, we resort to re-solving the entire model each time a perturbation is made to the model.

7.2 Sensitivity Analysis

1. Adding a New Neighbourhood With Population Less Than 200 People

Suppose a construction company bought a piece of land in Waterloo for the purpose of building residential homes in a new neighbourhood. Moreover, this new neighbourhood is not within an existing postal code region and hence its development introduces an addition of a new postal code region to our problem. However, due to size limitations of the purchased land, the new neighbourhood can only house a population of 100 residents.

With regards to our model, it was determined that in order to be economically feasible, a distribution center must service at least 200 people. Since our new neighbourhood has a population of only 100 residents, it is not feasible to build a new distribution center just for this new neighbourhood. Hence, we need to make use of excess capacity from existing distribution centers to service this neighbourhood. The alternative is to build a new distribution center and have its minimum capacity fulfilled from the excess capacity served by other distribution centers.

The solution implies that we still build the same 7 distribution centers, but the assignments change as follows:

- Location 2 services regions 20 and 21
- Location 5 services regions 14, 18, and 22
- Location 7 services regions 10 and 12
- Location 9 services regions 1, 4, 7, 19, and 23
- Location 11 services regions 3, 9, 11, and 15
- Location 13 services regions 2, 13, and 17
- Location 15 services regions 5, 6, 8, and 16.

The total cost of building these 7 distribution centers is $\sum_{j \in L} f(s_j) = 2.10$ million dollars, an increase of $\$20,000$ over the original solution.

The maximum distance traveled by any delivery truck is still the same at 14 kilometers per trip.

2. Adding a New Neighbourhood with Population More Than 200 People

We take the same situation as described above, but now we consider that the new neighbourhood can provide housing for a population of 300. With a larger population, it may be necessary that a new distribution center needs to be built or perhaps it may be more feasible to use existing centers with excess capacity to facilitate the needs of this new neighbourhood.

The solution implies that we still build the same 7 distribution centers, but the assignments change as follows:

- Location 2 services regions 20 and 21
- Location 5 services regions 10, 18, and 22
- Location 7 services regions 1, 2, 8, 12, and 16
- Location 9 services regions 7, 14, and 23
- Location 11 services regions 3, 4, 9, 11, and 19
- Location 13 services regions 13, 15 and 17
- Location 15 services regions 5 and 6.

The total cost of building these 7 distribution centers is $\sum_{j \in L} f(s_j) = 2.12$ million dollars, an increase of $\$40,000$ over the original solution.

The maximum distance traveled by any delivery truck is still the same at 14 kilometers per trip.

3. Adding a Far-Away Neighbourhood

Suppose a construction company bought a piece of land in the outskirts of Waterloo for the purpose of building condominiums. Since these condominiums are located at a more remote part of the county, their development introduces an additional postal code region to our problem. As well, due to its location, the neighbourhood puts itself significantly distant from the proposed distribution center locations. Furthermore, the condominiums are constructed to provide housing for a population of 1,200 residents.

The solution changes drastically as follows:

$y_j = 0$, except $y_1 = y_3 = y_5 = y_8 = y_{10} = y_{12} = y_{14} = 1$. This implies we still build 7 distribution centers but at these new locations: 1, 3, 5, 8, 10, 12, 14.

- Location 1 services regions 3, 4, 14, and 18
- Location 3 services regions 8, 17, 19, and 22
- Location 5 services regions 10, 11, and 15
- Location 8 services regions 13 and 23
- Location 10 services regions 6, 16, and 20
- Location 12 services regions 2, 5, 9, and 21
- Location 14 services regions 1, 7, and 12.

By adding the new condominiums along the outskirts of Waterloo, it drastically changes the locations at which the distribution centers are to be built. However, the total number of distribution centers built remains the same at 7.

The total cost of building these 7 distribution centers is $\sum_{j \in L} f(s_j) = 2.48$ million dollars, an increase of $\$400,000$ over the original solution. This shows how sensitive our original solution is to the addition of a distant neighbourhood with a relatively high population.

The maximum distance traveled by any delivery truck now increases to 23 from 14 kms per trip at a cost of $23 \times \$0.15 = \$3.45$ per trip.

4. Population Within a Postal Code Area Increases

The maximum distance traveled by the postmen indirectly depends on the populations in each of the existing postal code areas (which have been provided to us by the County of Waterloo) since the populations determine the locations of the distribution centers.

Now say for instance, we look at a particular postal code region located close to the University of Waterloo, which currently consists of several residential homes. Waterloo Living, a student housing company, has bought a large number of houses in this postal code region, and is demolishing them to build several apartment buildings in their place.

This change will drastically increase the overall population and mail flow through this postal code region. The following sensitivity analysis is performed to analyze the effects of a drastic population increase in a single postal code region. We have taken postal code region 8 for example, and increased its population from 100 to 500 people.

The solution implies that we still build the same 7 distribution centers, but the assignments change as follows:

- Location 2 services regions 20 and 21.
- Location 5 services regions 8, 10, and 18.
- Location 7 services regions 1, 2, 9, 12, and 16.
- Location 9 services regions 7, 17, and 19.
- Location 11 services regions 3, 4, 11, and 15.
- Location 13 services regions 6, 13, and 22.
- Location 15 services regions 5 and 14.

By increasing the population in the single postal code region 8 from 100 to 500, we have created the need to re-route several postal trucks to new regions in order to minimize the overall distance traveled.

The maximum distance D traveled by any delivery truck continues to be 14 kms per trip, which costs $14 \times \$0.15 = \$2.1$ per trip. Although the maximum distance traveled has not changed, the total cost of the above 7 distribution centers increases by $\$100,000$ to $\sum_{j \in L} f(s_j)$ $= 2.18$ million dollars.

5. Loss of Use of One Potential Distribution Center

After the county of Waterloo had determined that there were 15 potential lots to locate the mail distribution centers, there has been a major collapse in an adjacent landfill site, which made one of the lots inoperable. This reduces the potential number of lots to 14, and results in the following solution:

$y_j = 0$, except $y_1 = y_5 = y_8 = y_{10} = y_{12} = y_{14} = 1$. This implies we build only 6 distribution centers at locations 1, 5, 8, 10, 12, and 14, as opposed to the 7 originally proposed. The assignments are as follows:

- Location 1 services regions 1, 3, 6, 10, and 12.
- Location 5 services regions 14, 15, and 22.
- Location 8 services regions 11, 17, and 21.
- Location 10 services regions 4, 5, and 8.
- Location 12 services regions 2, 7, 9, and 16.
- Location 14 services regions 13, 18, 19, and 20.

By losing a potential location for a distribution center, we are not able to service postal code regions from that location. In particular, by losing the availability of potential lot #15, we are only utilizing the above 6 locations in order to maintain the maximum distance traveled. The re-routing of the service paths results in the maximum distance traveled by any delivery truck to be 15 km per trip, which costs $15 \times \$0.15 = \$2.25$, an increase of 15 cents per trip. The total cost of the above 6 distribution centers increases from 2.08 to $\sum_{j \in L} f(s_j) = 2.28$ million dollars, a change of $200,000 over the original solution.

7.3 Bottleneck Analysis

The bottleneck, which is the most critical situation in this solution, is the maximum distance D traveled by any postal worker in this solution. The bottleneck is significant in the solution as the goal is to minimize D in the objective function $T \times D$, with T being the cost of gasoline per kilometer traveled.

In the solution to our problem,

- d_i values are in the range between 6 to 14 kilometers
- This implies the maximum distance D traveled by any postal truck is 14 kilometers.

The five highest d_i values in the solution are:

- Regions 10 and 21's distance to their respective distribution centers is 14 km
- Region 20's distance to its assigned distribution center is 13 km
- Regions 5, 6, 12, 14, 17, and 19's distance to their distribution centers is 12 km
- Regions 4, 9, 11, and 18's distance to their respective distribution centers is 11 km
- Regions 13, and 22's distance to their respective distribution centers is 10 km

Apart from the two regions with the maximum distance of 14 kilometers from their respective distribution centers, most of the regions have relatively high d_i values, with one of them being 13 kilometers away from its assigned distribution center, and six others being 12 kilometers away. In addition, 15 out of the 22 total regions (68%) are at least 10 kilometers away from their respective distribution centers.

Thus, the maximum d_i has limited impact on the quality of this solution. Removing these bottlenecked regions (ie. regions 10 and 21) would only narrowly improve the solution, since most of the regions have a distance fairly close to D.

Appendices

A Full Mathematical Formulation

The final complete mixed integer linear programming (MIP) model for the postal distribution center allocation problem is as follows (K is the set of integers from 1 to I, and L is the set of integers from 1 to J):

$$\text{minimize} \quad T \cdot D$$

$$\begin{aligned}
\text{s.t.} \quad \sum_{j \in L} x_{ij} &= 1 && \forall\, i \in K \\
\sum_{i \in K} x_{ij} &\leq y_j I && \forall\, j \in L \\
s_j &= \sum_{i \in K} p_i x_{ij} && \forall\, j \in L \\
f(s_j) &= 3v_1(j) + 2v_2(j) + 4v_3(j) && \forall\, j \in L \\
v_2(j) &\leq 4u_1(j) && \forall\, j \in L \\
v_1(j) &\geq 4u_1(j) && \forall\, j \in L \\
v_3(j) &\leq 10u_2(j) && \forall\, j \in L \\
v_2(j) &\geq 4u_2(j) && \forall\, j \in L \\
s_j &= v_1(j) + v_2(j) + v_3(j) && \forall\, j \in L \\
\sum_{j \in L} f(s_j) &\leq B \\
2y_j \leq s_j &\leq l s_j && \forall\, j \in L \\
y_1 + y_5 + y_8 &\geq 3k \\
y_2 + y_5 + y_9 &\geq 3(1-k) \\
y_j + y(j+1) &\leq 1 && \forall\, j \in \{1, \ldots, J-1\} \\
d_i &= \sum_{j \in L} (dd_{ij})(x_{ij}) && \forall\, i \in K \\
D &\geq d_i && \forall\, i \in K \\
\\
y_j &\in \{0, 1\} && \forall\, j \in L \\
x_{ij} &\in \{0, 1\} && \forall\, i \in K, j \in L \\
k &\in \{0, 1\} \\
0 \leq v_1(j) &\leq 4 && \forall\, j \in L \\
0 \leq v_2(j) &\leq 4 && \forall\, j \in L
\end{aligned}$$

$$
\begin{aligned}
v_3(j) &\geq 0 && \forall\ j \in L \\
u_1(j) &\in \{0,1\} && \forall\ j \in L \\
u_2(j) &\in \{0,1\} && \forall\ j \in L
\end{aligned}
$$

The decision variables are y_j and x_{ij}.

The auxiliary decision variables are D, k, s_j, d_i, $f(s_j)$, $v_1(j)$, $v_2(j)$, $v_3(j)$, $u_1(j)$ and $u_2(j)$ $\forall\ i \in K, j \in L$.

B AMPL Model

```
param I; # of postal code areas
param J; # of proposed distribution center locations

# unit cost of fuel in dollars
param T;

# population (in hundreds) of each postal code area
param p{i in 1..I};

# total budget alloted (in thousands of dollars) for this project
param B;

# min. # of people serviced per dist. center
param m;

# distance between region i and location j
param dd{i in 1..I, j in 1..J};

# maximum pop. (in hundreds) serviceable by dist. center j
param ls{j in 1..J};

var y{i in 1..J} binary;
var x{i in 1..I, j in 1..J} binary;

# max distance travelled by any one postman
var D;

# population (in hundreds) serviced by each distribution center
var s{i in 1..J};

# distance between region i and the location j associated with it
var d{i in 1..I};

# for downtown either-or constraint
var k binary;

# piecewise cost function vars
var f{j in 1..J} >= 0;
var v1{j in 1..J} >= 0;
var v2{j in 1..J} >= 0;
var v3{j in 1..J} >= 0;
var u1{j in 1..J} binary;
var u2{j in 1..J} binary;
```

```

# objective function: cost of fuel * max distance traveled
minimize cost: T * D;

# 1) each region can be assigned to exactly one distribution center
subject to onecenter{i in 1..I}: sum {j in 1..J} x[i,j] = 1;

# 2) regions cannot be assigned to unbuilt distribution centers
subject to noemptycenter{j in 1..J}: sum {i in 1..I} x[i,j] <= y[j] * I;

# 3) define the population s[j] serviced by each distribution center j
subject to popofcenter{j in 1..J}: sum {i in 1..I} (x[i,j]*p[i]) = s[j];

# 4) define the cost of building each dist. center
subject to centercost{j in 1..J}: f[j] = 3*v1[j] + 2*v2[j] + 4*v3[j];

# 5) piece-wise constraints modeled as LP constraints
subject to c5{j in 1..J}: v2[j] <= 4*u1[j];
subject to c6{j in 1..J}: v1[j] >= 4*u1[j];
subject to c7{j in 1..J}: v3[j] <= 10*u2[j];
subject to c8{j in 1..J}: v2[j] >= 4*u2[j];
subject to c9{j in 1..J}: s[j] = v1[j] + v2[j] + v3[j];

# 6) cannot exceed max budget for this project
subject to maxbudget: sum {j in 1..J} f[j] <= B;

# 7) min. people serviced by a dist. center to be considered feasible
subject to minpop{j in 1..J}: s[j] >= m*y[j];

# 8) max. people that can be serviced by center j
subject to maxpop{j in 1..J}: s[j] <= ls[j];

# 9) downtown extra either-or congestion constraints
subject to c11: y[1] + y[5] + y[8] >= 3*k;
subject to c12: y[2] + y[5] + y[9] >= 3*(1-k);

# 10) no adjacent distribution centers
subject to noadjcenters{j in 1..J-1}: y[j] + y[j+1] <= 1;

# 11) define the distance travelled by each postman in regin i
subject to distdef{i in 1..I}: d[i] = sum {j in 1..J} (dd[i,j]*x[i,j]);

# 12) define D to be the maximum distance traveled by any postman
subject to maxdist{i in 1..I}: D >= d[i];

```

```
# 13) piecewise cost function f() constraints
subject to c19{j in 1..J}: v1[j] <= 4;
subject to c20{j in 1..J}: v2[j] <= 4;

#option solver cplex;
option solver gurobi;
#option gurobi_options 'solnsens=1';
data dist.dat;
solve;

# display full solution
display x, y, d, f, s, v1, v2, v3, u1, u2;
display D;
display sum {j in 1..J} f[j];
```

C AMPL Data

```
# of postal code areas
param I := 22;

# of proposed distribution center locations
param J := 15;

# unit cost of fuel in dollars
param T := 0.15; # 15 cents

# total budget alloted in dollars for this project
param B := 250; # 2.5 million

# min. # of people serviced per dist. center
param m := 2;

include regions.dat;
```

```
# max people (in hundreds) each dist. center can service
param ls :=
1 14
2 10
3 13
4 11
5 14
6 15
7 12
8 13
9 11
10 13
11 12
12 13
13 11
14 12
15 13;

# pop. (in hundred) of each region
param p :=
1 1
2 2
3 1
4 1
5 4
6 4
```

```
7 5
8 1
9 1
10 4
11 3
12 4
13 1
14 3
15 6
16 4
17 4
18 5
19 2
20 4
21 6
22 5;

# distance (in kms) between region i and location j
param dd :
    1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 =
1   7  8 26 14 20 21 12  9  7 17 17 19 22  9 20
2  24 15 18 14 19 18  9 21  8 26 27 11  9 24 14
3   8 26  9 16 26  6 12 19 22 24  6 24  8 28 22
4  19 10 21 10 24 18 21 25  6 13 11 17  9  6 28
5  16 22 25 10 19 20 12 19 12 15 17 13 13 28  7
6  11  7 25  7 15 13 28  8 21 18 23 12 10 16 12
7  10 21 17 27 27 14 15 22  7 28 20  8 21 23 28
8   7 15 14 23  6 14  7 25 28  7 18 19 15 21  6
9  25 12 23 12 26  7 14 16 20 22 10 12 15 21 11
10  6  6 28  9 14 25  7  6 18 23 26 15 12  7 27
11 15 23 19  9 21 22 21  6 11 14 11 18 22 16 18
12  8 10 10 16 24 11 12 13 28  6 27  6 21  8 26
13 25 12 18 13 21 26 27  6 25 26 16 10 10 11 19
14 17  6 16 17 11 11 16 23  6 10 23 16 10 21 12
15  9 14  7 18  6 24  9 17 20 11 14 19  6 26 17
16 25 25 15 26 20  9  8 15  6 19 21 15 15 10 12
17 21 24 20 22 20 21 26  9  6 28 21 13 12 10 28
18 16 10 12 17 11  9 22 11 26 28 16 24 10  6 28
19 14 12  7 11 16 27 18 22  9 14 12 14  7 12 19
20 21 13 26 12 20 17 26 24 21 21 18 26 23 15 19
21 28 14 26  9 25 17 20  7 26 19 23 16 26 19 26
22  8 11  6  6 10 23 25 26 16 24 16 16  8 10 27;
```

D AMPL Solution Output

```
ampl: model dist.mod
Gurobi 2.0.1: optimal solution; objective 2.1
1179 simplex iterations; 0 branch-and-cut nodes
plus 111 simplex iterations for intbasis
x [*,*]
:     1   2   3   4   5   6   7   8   9  10  11  12  13  14  15    :=
1     0   0   0   0   0   0   0   0   1   0   0   0   0   0   0
2     0   0   0   0   0   0   0   0   1   0   0   0   0   0   0
3     0   0   0   0   0   0   0   0   0   0   1   0   0   0   0
4     0   0   0   0   0   0   0   0   0   0   1   0   0   0   0
5     0   0   0   0   0   0   1   0   0   0   0   0   0   0   0
6     0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
7     0   0   0   0   0   0   0   0   1   0   0   0   0   0   0
8     0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
9     0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
10    0   0   0   0   1   0   0   0   0   0   0   0   0   0   0
11    0   0   0   0   0   0   0   0   1   0   0   0   0   0   0
12    0   0   0   0   0   0   1   0   0   0   0   0   0   0   0
13    0   0   0   0   0   0   0   0   0   0   0   0   1   0   0
14    0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
15    0   0   0   0   0   0   0   0   0   0   0   0   1   0   0
16    0   0   0   0   0   0   1   0   0   0   0   0   0   0   0
17    0   0   0   0   0   0   0   0   0   0   0   0   1   0   0
18    0   0   0   0   1   0   0   0   0   0   0   0   0   0   0
19    0   0   0   0   0   0   0   0   0   0   1   0   0   0   0
20    0   1   0   0   0   0   0   0   0   0   0   0   0   0   0
21    0   1   0   0   0   0   0   0   0   0   0   0   0   0   0
22    0   0   0   0   1   0   0   0   0   0   0   0   0   0   0
;

:     y   d    f    s    v1   v2    v3  u1  u2    :=
1     0    7    0    0    0    0     0   0   0
2     1    8   28   10    4    4     2   1   1
3     0    6    0    0    0    0     0   0   0
4     0   11    0    0    0    0     0   0   0
5     1   12   44   14    4    4     6   1   1
6     0   12    0    0    0    0     0   0   0
7     1    7   36   12    4    4     4   1   1
8     0    6    0    0    0    0     0   0   0
9     1   11   32   11    4    4     3   1   1
10    0   14    0    0    0    0     0   0   0
11    1   11   12    4    4    0     0   0   0
12    0   12    0    0    0    0     0   0   0
```

```
13   1   10   32    11    4    4     3   1   1
14   0   12    0     0    0    0     0   0   0
15   1    6   24     9    4    4     1   1   1
16   .    8    .     .    .    .     .   .   .
17   .   12    .     .    .    .     .   .   .
18   .   11    .     .    .    .     .   .   .
19   .   12    .     .    .    .     .   .   .
20   .   13    .     .    .    .     .   .   .
21   .   14    .     .    .    .     .   .   .
22   .   10    .     .    .    .     .   .   .
;

D = 14

sum{j in 1 .. J} f[j] = 208

d [*] :=
 1  7     4 11     7  7    10 14    13 10    16  8    19 12    22 10
 2  8     5 12     8  6    11 11    14 12    17 12    20 13
 3  6     6 12     9 11    12 12    15  6    18 11    21 14
```