Garbage Collection in Chicago: A Dynamic Scheduling Model

Donald D. Eisenstein • Ananth. V. Iyer
Graduate School of Business, The University of Chicago, Chicago, Illinois 60637
Krannert Graduate School of Management, Purdue University, West Lafayette, Indiana 47907

We investigate the scheduling of garbage trucks in the city of Chicago. Analysis of data collected from the system shows that city blocks differ in the rate at which garbage is collected. However, in the current system, each truck visits the dumpsite two times each day. Our approach is to devise a flexible routing scheme in which some routes visit the dumpsite only once per day, while others visit the dumpsite twice per day depending on the blocks assigned to the route. We use a Markov decision process to model the impact on capacity of using flexible routes. This provides a dynamic scheduling algorithm that adjusts the number of dumpsite visits throughout the week to maximize service level. Results of the model suggest a potential reduction in truck capacity of 12–16% for a set of five pilot wards. This paper shows that flexible schedules can significantly reduce the capacity required to operate a system in the presence of variability.
(*Garbage Collection; Markov Decision Process; Public Sector Application; Capacitated Truck Routing*)

1. Introduction

The Department of Streets and Sanitation for the city of Chicago has an annual budget of over $70 million and deploys over 330 trucks daily for garbage collection. Increasing budgetary pressure has resulted in a steady decrease in the truck capacity available to the city. We study the city's deployment of truck capacity, and base our analysis on data gathered from five (out of 50) wards.

At the start of each day, each truck is provided a "time stamp" schedule, that is, each truck crew is directed to collect garbage for a *fixed* amount of time before going to the dumpsite. In the current schedule each truck collects garbage for 2.5 hours each morning, empties its load at the dumpsite, collects for another 2.5 hours in the afternoon, and then makes a final trip to the dumpsite. With only 2.5 hours of picking time before each trip to the dumpsite, the trucks are never weight constrained. Such time stamp schedules are easy for the crew to follow. Management thus feels that all implementable schedules should have this time stamp feature. Alternatively, schedules that require crews to make adaptive decisions are difficult to implement in the current environment.

We model the weight and time required to collect garbage from a block as a random variable whose parameters are estimated from data. This model suggests that blocks differ in their weight and time requirements. We introduce a dynamic scheduling algorithm that uses *flexible* truck routes—some of which visit the dumpsite twice a day and others that visit the dumpsite once a day. Each of these flexible route types are specified as a time stamp schedule. At the start of each day, we choose the route type that maximizes the service level (probability that the truck completes its assigned sequence of blocks for the week).

We model the choice of a truck route as a Markov decision process. The output of our model is an implementable schedule that provides daily work assignments for each truck. We compare the capacity required by the model to the capacity used in the current operation. We demonstrate that our approach decreases

0025-1909/97/4307/0922$05.00

truck capacity by 12–16%. This corresponds to a potential cost reduction of over $9 million.

The next section discusses related literature. Section 3 provides details of the garbage collection process and our sources of data. In §4 we provide a statistical model of the weight and time required to collect a sequence of blocks. Section 5 examines the use of flexible routes. In §6 we develop a dynamic scheduling algorithm for collecting a sequence of blocks over a week based on a Markov decision process. Section 7 provides results from application of the model to a data set collected from five pilot wards and §8 our conclusions.

2. Related Work

There is a stream of research that concentrates on forming routes along edges in a network so that travel is minimized. Such work usually concentrates on some variant of the Chinese Postman Problem such as capacities on the edges, a mix of directed and undirected edges, or multiple truck routes. (See for example, Eglese 1994, for routing winter gritting vehicles; Golden and Assad 1988; Golden and Wong 1983; Male and Liebman 1978.) Since these variants are computationally hard, these papers have reported on branch and bound approaches or various heuristic methods. Beltrami and Bodin (1974) consider the collection of large commercial bins in New York. Since these collection points are more sparse, they model such collection points as the nodes on a network, and use a heuristic procedure to solve this Vehicle Routing Problem. Since in these cases the network of streets is complicated, a travel time focus is appropriate. However, for our problem the grid structure of the blocks admits easy solutions to minimize travel time. Therefore, for our problem, there is little to be gained from a routing focus.

We do not account for interblock travel time explicitly, but instead restrict ourselves to routes where the blocks are contiguous. This is similar to other work in which a geometric area is partitioned according to some criteria. For example, Christofides and Whitlock (1977) and Biro and Boros (1984) report on cutting stock problems restricted to a guillotine cut structure. This is also related to work that first identifies feasible subregions (corresponding to, for instance, a truck's weekly schedule of blocks) and then tries to cover the whole region with as few subregions as possible. For instance, if each subregion is a rectangle then this covering problem can be posed as one of intersecting rectangles in the plane (see Roberts 1969, Maehara 1984, Imai and Asano 1983).

New York is another major city whose Department of Streets and Sanitation has solicited assistance from Management Science models (see Riccio 1984, Riccio and Litke 1986, Riccio, Miller, and Litke 1986).

Our approach differs from other work in this area in a number of ways. First, we consider the use of flexible routes that can visit the dumpsite either once or twice a day. Second, we explicitly model the uncertainty in the system by using a statistical model of the weight and time required for each block. And finally, we introduce a dynamic scheduling algorithm that will adjust the routing based on the progress of each truck through the week.

3. Garbage Collection in Chicago

The city does not pick up garbage from large apartment or commercial buildings, which leaves about 50% of the living units (separate households) for collection. The city is divided into 50 wards, and the garbage collection for each ward is managed independently by a ward superintendent. We chose a set of five adjoining pilot wards and built a database of information from the following sources: (i) interviews with the planners in city hall and ward superintendents, (ii) field observations from three different wards that provided the time spent driving and collecting garbage across the day, (iii) driver logs of all trucks collecting garbage from pilot wards for a set of six weeks (these logs provide the blocks generating each truck load), (iv) maps of each of the five pilot wards giving the location of each block, (v) a database of the living units in each block that the city is responsible for, and (vi) a database of the pounds of garbage dumped at the dumpsite for each truck visit. We use this information to describe the collection problem faced by the ward superintendent.

Each ward superintendent is allocated a fixed number of trucks that he deploys to collect all the garbage in the ward by the end of the week. His problem is twofold. First, the mean time to collect garbage and the weight collected varies from block to block. One reason for this variation is the different number of living units the city collects in each block.[1] His second concern is that each week the time required to collect garbage from a block and the associated weight picked up is uncertain—this uncertainty implies that the workload for a truck has to be rescheduled daily. He has to manage this variation while satisfying constraints with respect to the truck's weight capacity and the hours available in a work day for the truck to travel, collect, and dump garbage.

The ward superintendent manages this variability in a number of ways. First, he takes advantage of the grid structure of the blocks in a ward by routing trucks to move across alleys of contiguous blocks with minimal wasted travel time between blocks. Second, he has each truck follow the same time schedule, each making two trips to the dumpsite each day. This ensures that trucks are never weight constrained. And finally, each truck's weekly workload has slack capacity on Friday. While such simplifications make his problem easier, they do impact the capacity required to collect garbage.

3.1. Daily Collection

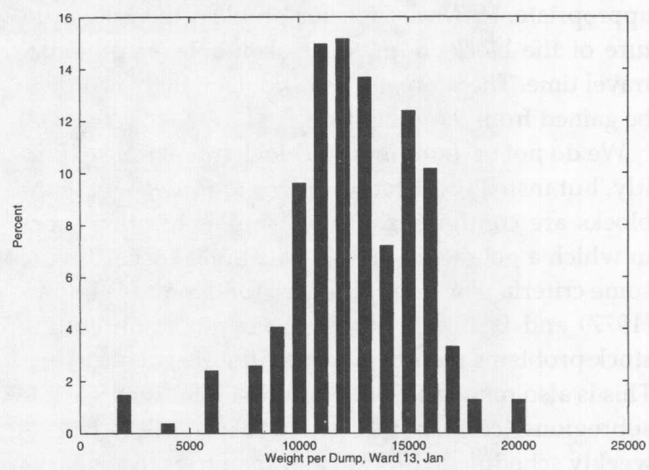
Residents of each living unit store their garbage in city provided carts which are usually placed in the alley for collection. Each truck is assigned a three-person crew; one driver and two laborers. The two laborers ride on the back of the truck and empty the garbage from the carts into the truck bay. The laborers are also responsible for collecting any garbage outside the carts and keeping the alleyways clear. The crew begins its daily collecting at 8:00 am and at about 10:30 am the truck leaves for the dumpsite to empty the morning load. At about 11:30 am the dump is completed and the truck is back in the ward. Collection resumes after lunch at 12:00 pm and continues until 2:30 pm, followed by a final trip to the dumpsite to empty the afternoon load. At about 3:30 pm the driver returns to the ward and completes his paperwork for the day. The roundtrip time to the dumpsite is about one hour. Thus each truck currently spends two hours out of seven traveling to the dumpsite, dumping its garbage, and returning to the ward. City regulations do not permit the garbage to be stored in any location other than the dumpsite, so all trucks must end their work day with a visit to the dumpsite.

Each dumpsite stores detailed data regarding the arrival time of the truck to the dumpsite, the pounds of garbage deposited by the truck, and the time at which the truck departs from the dumpsite. Figure 1 provides a histogram of the pounds of garbage dumped at each truck visit to the dumpsite for ward 13 (histograms for the other wards look very similar). The average garbage load per truckload across all wards is just over 11,500 pounds. However, the truck capacity that can be planned is 22,000 pounds, indicating an average capacity utilization of 52%. The total garbage picked up by a truck across the two dumps each day averages over 23,000 pounds—exceeding the capacity of one truck load. Thus, on average, a truck cannot easily forgo the midday dumpsite visit. The downside, however, is that truck capacity utilization is low.

3.2. Weekly Collection

Each truck crew is assigned a sequence of blocks to complete by the end of the week. Determining the sequence to route each truck is relatively easy—with but few exceptions the city blocks are laid out on a grid. Thus a truck follows a long snake-like route traveling from the

Figure 1 Histogram of Weight per Truck Visit to the Dumpsite for Ward 13 in January 1992



[1] For example in ward 12, the city collects 17 living units from block 322 and 158 living units from block 182. This variation reflects the fact that ward 12 has some blocks consisting primarily of single family dwellings and others consisting primarily of two and three flat buildings.

alley of one block right into the alley of an adjacent block, beginning one day where they left off the previous day. This grid structure of the blocks makes it easy for the ward superintendent to form routes so that wasted travel time between blocks is kept small.

The dumpsite data shows that the average load picked up on Fridays is less than 50% of the average loads during the rest of the week. The Streets and Sanitation personnel explained this as a "work and hope" collection system—you work Monday through Thursday and hope that enough slack is reserved for Friday to absorb any variance during the week. Most trucks complete their sequence by midday on Friday and are available in the afternoon to assist a crew that may need help for that week. Because of variance faced by the collection process, a resident is not promised garbage collection on a particular day of the week.

4. A Model of Block Load

We model the time to collect the garbage from block k as a normally distributed random variable, X_k, with mean collection time of t_k and a variance in collection time of $\sigma^2 t_k$. A constant block dependent parameter r_k represents the rate of garbage collection in pounds per unit time for block k. Thus the mean pounds of garbage picked up from block k is $r_k t_k$ and the variance in the pounds of garbage picked up is $r_k^2 \sigma^2 t_k$. In the appendix we detail a procedure that uses the data to estimate σ^2 and the parameters t_k and r_k for each block k. The appendix also details the use of a seasonality factor that adjusts these estimates for a block k based on the month of the year.

Let $S = \{i, i+1, \ldots, j-1\}$ be a set of adjoining blocks to be collected by a truck. Thus for this set of blocks, S, our model of the weight is normally distributed with mean $\Sigma_{k \in S} r_k t_k$ and variance $\Sigma_{k \in S} r_k^2 t_k \sigma^2$. Our model of the required time is normally distributed with mean $\Sigma_{k \in S} t_k$ and variance $\Sigma_{k \in S} t_k \sigma^2$.

We let T denote the total time available for a truck route and W be the total weight capacity of the truck. We define

$$\Phi_{ij} = P\left\{ \sum_{k=i}^{j-1} X_k \le T \text{ and } \sum_{k=i}^{j-1} r_k X_k \le W \right\}$$

to be the probability that a truck completes collection of blocks $i, i+1, \ldots, j-1$.

The time and weight required to collect a sequence of blocks is correlated (because of the constant rate factors r_k). The joint distribution of two linear combinations of normal random variables (i.e., the time (x) and weight (y) for a sequence of blocks) follows a bivariate normal distribution (see Rohatgi 1976). Thus

$$\Phi_{ij} = \int_{x=0}^{T} \int_{y=0}^{W} \phi_{ij}(x, y)\,dx\,dy,$$

where $\phi_{ij}(x, y)$ is the bivariate normal density function. The marginal densities of X (the time random variable) and Y (the weight random variable) are normally distributed with $\mu_x = \Sigma_{k=i}^{j-1} t_k$, $\sigma_x^2 = \sigma^2 \Sigma_{k=i}^{j-1} t_k$ and $\mu_y = \Sigma_{k=i}^{j-1} r_k t_k$, $\sigma_y^2 = \sigma^2 \Sigma_{k=i}^{j-1} r_k^2 t_k$. By definition, $\Phi_{ij} = 1$ for $j \le i$. We define ρ_{ij}, the correlation between X and Y for blocks i to j, as

$$\rho_{ij} = \frac{\Sigma_{k=i}^{j-1} r_k t_k}{\sqrt{\Sigma_{k=i}^{j-1} t_k \, \Sigma_{k=i}^{j-1} r_k^2 t_k}}\,.$$

4.1. Model Justification

In the appendix we show how we used our data set to obtain six deseasonalized data values for time and weight for each block. With such a small data set for each block, we cannot test for normality of the travel time at the block level. We instead standardized the times for each block (i.e., $(X_k - t_k)/\sigma\sqrt{t_k})$) and pooled these standardized times across all blocks—this pooled set of data is well approximated by a normal distribution. While this does not necessarily imply that the underlying data is normal, it does provide some justification for the assumption that the time to pick each block follows a normal distribution. Normality of the time to pick a block also means that the sum of the times from successive blocks also follow a normal distribution and that larger blocks have a lower coefficient of variation than smaller blocks—a model feature that is intuitively appealing.

For each block we then examined the deseasonalized rates across months. The coefficient of variation of these rates is small (less than 0.2 in almost all cases). We therefore chose to model the rate as a constant term that is block dependent. We remark that this modeling decision implies that we will have a correlation between the time spent in a block and the corresponding weight generated by that block. Furthermore, the joint distribution of time and weight follows a bivariate normal distribution. It is intuitively appealing that more time spent collecting a block would imply more weight generated by that block.

The predictive power of the model was validated in two ways. First, for each of the routes for which we had data in 1992, we used the information regarding the blocks picked up to generate a mean truckload weight and associated standard deviation of the truckload weight. We observe that in over 85% of the 2,200 routes examined in 1992, the actual truckload weight observed was within three standard deviations of the mean weight estimated by our model (based on the actual blocks collected by the truck). Second, we selected a random sample of 5% of the truck routes from the first week in June 1994. Again, we used information regarding the blocks picked up to estimate the mean truckload weight and associated standard deviation of the truckload weight. The actual truckload weight observed was within three standard deviations of the mean weight estimated by our model in over 90% of the selected routes. We thus conclude that we have an acceptable statistical model of the loads generated by blocks and can thus evaluate the effect of changes in the scheduling of trucks.

5. Flexible Truck Routes

We have seen that on average a truck collects over half its weight capacity on each of two daily trips to the dumpsite. Thus for a truck to reliably forego the midday dumpsite visit and thus gain about an hour of picking time, it would need to visit a number of blocks whose load was small relative to the time it takes to collect. That is, such a route would need to visit blocks of a sufficiently low collection rate (weight per unit time) of garbage so that it could collect garbage an extra hour while not reaching its weight capacity. We call such a hypothetical route that visits the dumpsite only once at the end of the day a 1-*dump* route. On the other hand, blocks of high collection rate can continue to be collected by routes that will go to the dumpsite twice a day—a 2-*dump* route. If enough 1-dump routes could be formed that don't exceed weight capacity, then the trucks would be better utilized since each 1-dump route gains an extra 20% picking time. We note that 1-dump routes are a viable strategy for the city to implement. The driver would be working the same number of hours, yet instead of a trip to the dumpsite, he would continue collecting. The laborers are not needed for a trip to the dumpsite, so they are currently allowed an extra hour off during the midday dumpsite visit; however, they are paid during this time and the city has the option to better utilize their time.

For ward 12, the block rates estimated by our model range from 55 pounds per minute for block 90 to 122 pounds per minute for block 121. Blocks have different rates for a number of reasons. First, the city does not pick up commercial properties or large apartment buildings—these must get private service. Thus a block interspersed with such buildings will tend to gather garbage at a slower rate because the truck must still travel the alley, while bypassing a number of buildings without collecting. Blocks also differ in the amount of garbage that collects outside of carts. In such cases this garbage must be swept up by the laborers, slowing the rate of collection. We observed this phenomenon when collecting data from two different wards; for one ward the time to collect a cart averaged 20 seconds, while another ward averaged 50 seconds. In addition, some alleys are narrower and harder to navigate than others.

In order for a 1-dump route to not be weight constrained, the collection rate must average no more than $22{,}000/360 = 61.1$ lbs/min. About 45% of the blocks have rates below 61.1 lbs/min. Furthermore, if one aggregates the weight and time of 90% of the blocks with the lowest rates, then this aggregated rate is below 61.1 lbs/min. This suggests that 1-dump routes could use a mix of blocks, with rates above and below 61.1 lbs/min. Thus the data indicates a potential for forming flexible routes.

For a truck starting the collection day, in a *flexible* routing system, we choose between one of two actions we denote by $a \in \{1, 2\}$. The action $a = 2$ implies that the truck runs a 2-dump route—in the morning it goes to the dumpsite after 2.5 hours and again collects for 2.5 hours in the afternoon before going to the dumpsite. The action $a = 1$ implies that the truck runs a 1-dump route and thus travels throughout the day and goes to the dumpsite only once after 6 hours of collecting.

We expand the definition of Φ_{ij}, T, and W to include the decision parameter a so that

$$\Phi_{ij}(a) = P\left\{\sum_{k=i}^{j-1} X_k \le T(a) \text{ and } \sum_{k=i}^{j-1} r_k X_k \le W(a)\right\}$$

$$= \int_{x=0}^{T(a)} \int_{y=0}^{W(a)} \phi_{ij}(x, y)dxdy,$$

is the probability that a truck beginning at the start of block i completes *at least* to the *start* of block j in a single day using action a.[2] For our system, the 1-dump routes are characterized by $W(1) = 22{,}000$ and $T(1) = 360$, and the 2-dump routes by $W(2) = 2W(1) = 44{,}000$ and $T(2) = 300$.

5.1. The Impact of Variability on Flexible Routes: An Example

We provide a simple example consisting of three blocks to demonstrate the impact of variability on the probability of completing these blocks under a 1-dump or a 2-dump route. Consider three blocks with mean garbage collection time of {50, 30, 23} and constant (block dependent) collection rates of {60, 45, 70}. Thus the mean weight and time across these three blocks are 5,960 pounds and 103 minutes, respectively. With $\sigma = 1.3$, the standard deviation of the weight is 773 pounds and the standard deviation of the time required is 13.2 minutes. Also, the correlation $\rho_{14} = 0.98$.

Let the weight capacity for a 1-dump route be 6,000 pounds and the time capacity for collection be 150 minutes. Let the corresponding values for a 2-dump route be 12,000 pounds and 130 minutes. In the absence of variability, both 1-dump and 2-dump strategies can complete all three blocks in one day. However, in the presence of variability, the service level under a 1-dump route is $\Phi_{1,4}(1) = 0.528$ and for a 2-dump route is $\Phi_{1,4}(2) = 0.979$. We therefore need an algorithm to choose the appropriate route type depending on the sequence of blocks to be collected.

6. A Markov Decision Process

We defined $\Phi_{ij}(a)$ as the probability of collecting a sequence of blocks *in one day* given a 1-dump or 2-dump route decision ($a = 1, 2$). We now consider how to best schedule a single truck over a longer, week long, sequence of blocks. We build a Markov decision model that optimally selects the action a truck should take depending on the day of the week, and the starting block for the day. Thus the truck follows a *Dynamic Scheduling Algorithm*, changing its choice of a 1-dump or 2-dump route each day, depending on its current situation in order to maximize its service level.

We then consider a sequence of blocks for the entire ward and use our Dynamic Scheduling Algorithm to divide it into block sequences, each designed for a single truck to complete in the course of a week within a specified service level—our *Weekly Workload Assignment Algorithm*.

6.1. A Dynamic Scheduling Algorithm

We consider a sequence of blocks $1, 2, 3, \ldots, n$ to be collected by a single truck over a five-day work week. For convenience, we append a dummy block $n + 1$ that has zero weight and time requirement. We let this sequence of blocks $\{1, 2, \ldots, n + 1\}$ define our states of a Markov decision process. Our stochastic process (garbage truck) is said to be in state i at day t if at the *start* of day t the truck is at the *start* of block i. Each morning a decision $a \in A = \{1, 2\}$ is made to run either a 1 or 2-dump route. We let $p(j \mid i, a)$ be the probability that the truck transitions to state j given that the truck made decision a from state i the previous day (this notation follows Puterman 1994). So for example, $p(n + 1 \mid 1, a)$ is the probability that the truck completes all n blocks in one day under action a.

We can specify the $p(j \mid i, a)$ for $i, j = 1, 2, \ldots, n + 1$ by

$$p(j \mid i, a) = \begin{cases} 0 & \text{if } j < i, \\ \Phi_{ij}(a) - \Phi_{i,j+1}(a) & \text{if } i \le j \le n, \\ \Phi_{i,n+1}(a) & \text{if } j = n + 1. \end{cases}$$

The $p(j \mid i, a)$ form a stochastic matrix of one-step transition probabilities. We define a value function, $v_t(i)$, to be the maximum probability that a truck will complete all n blocks by day 5 if it starts in state i on day t. We are interested in determining the set of actions that define $v_1(1)$, since the system always begins in state 1 on day 1.

Our deterministic Markovian decision rule $d_t(i)$ specifies which action should be chosen in state i at day t.

[2] For our data set 2-dump routes have virtually zero probability of being weight constrained, thus this expression holds.

Our policy is a sequence of deterministic Markovian decision rules. We derive the optimal policy through backward induction. We start with the vector $\{v_6(1), v_6(2), \ldots, v_6(n+1)\}$, and then work back through each day until finally determining the service level $v_1(1)$. We define $v_6(i) = 0$ if $i \leq n$ and $v_6(n+1) = 1$. We then recursively compute for each $t = 5, 4, 3, 2, 1$,

$$v_t(i) = \max_{a \in \{1,2\}} \left\{ \sum_{j=1}^{n+1} p(j \mid i, a) v_{t+1}(j) \right\} \quad \text{for } i = 1, \ldots, n+1. \tag{1}$$

Equation (1) forms our optimality equation. By recursively applying Equation (1) an optimal set of actions is determined to define the value function $v_1(1)$ (see Theorem 4.5.1 in Puterman 1994).

Thus Equation (1) provides our Dynamic Scheduling Algorithm, where the optimal set of actions specify decisions $d_t(i) \in A$ for each day t and block i whether a truck should follow a 1-dump or 2-dump route. The algorithm requires $O(n^2)$ steps for a fixed number of days.

6.1.1. The Dynamic Scheduling Algorithm: An Example. We present a small example to show how our flexible routes can decrease the trucking capacity required. For this example, a 2-dump route has a weight capacity of 12,000 pounds and a time capacity of 130 minutes. A 1-dump route has a weight capacity of 6,000 pounds and a time capacity of 150 minutes.

We consider a sequence of seven blocks with mean *time* requirements (in min) of

$$\{50.0, 30.0, 23.0, 50.0, 20.0, 30.0, 30.0\},$$

and constant *rates* (in lbs./min.) of

$$\{60.0, 45.0, 70.0, 20.0, 20.0, 60.0, 70.0\}.$$

The mean weight and time for these blocks is 11,060 pounds and 233 minutes, thus we expect at least two truck days of capacity to complete such a sequence with a reasonable service level.

With a standard deviation of $\sigma = 1.3$ we generate the one-step transition probabilities for both 1-dump and 2-dump routes. (A dummy state 8 is added with zero weight and time requirement.)

For a 1-dump route we have the matrix of one-step transitions, $p(j \mid i, 1)$, as

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 0.004 & 0.475 & 0.426 & 0.064 & 0.031 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.031 & 0.622 & 0.345 & 0.002 \\ 0 & 0 & 0 & 0 & 0 & 0.046 & 0.822 & 0.132 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.171 & 0.829 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.003 & 0.997 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.001 & 0.999 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}.$$

Thus a 1-dump route will likely reach state 3 (complete only 2 blocks) with probability 0.475, or will reach state 4 (complete 3 blocks) with probability 0.426 in a single day. This is due to the large rates of the first few blocks resulting in a fast accumulation of weight. The two-step transition probability from state 1 to state 8 is just 0.51 (the [1, 8] element of $\mathbf{P}_1\mathbf{P}_1$), so it requires over two days or more truck capacity to complete the sequence of blocks with 1-dump routes for a reasonable service level.

For a 2-dump route we have the matrix of one-step transitions, $p(j \mid i, 2)$, as

$$\mathbf{P}_2 = \begin{bmatrix} 0 & 0 & 0.020 & 0.903 & 0.071 & 0.006 & 0 & 0 \\ 0 & 0 & 0 & 0.021 & 0.293 & 0.610 & 0.075 & 0.001 \\ 0 & 0 & 0 & 0 & 0.002 & 0.312 & 0.610 & 0.076 \\ 0 & 0 & 0 & 0 & 0 & 0.011 & 0.489 & 0.500 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}.$$

A 2-dump route does better on the first day, most likely to complete 3 blocks (state 4) with probability 0.903. But the transition from state 4 to state 8, to complete the sequence on the second day, has probability only 0.50. Indeed the later blocks pose a problem for a 2-dump route because of the time requirement. The two-step transition probability from state 1 to state 8 is just 0.53 (the [1, 8] element of $\mathbf{P}_2\mathbf{P}_2$). Thus using only 2-dump routes will again require over two days or more truck capacity to complete the sequence of blocks for a reasonable service level.

If, however, we use a flexible route strategy we can complete the sequence in just two days with a service level that exceeds 80%. The optimal decision rule for day 1 starting in state 1 is to choose a 2-dump route,

$d_1(1) = 2$. For day 2, $d_2(i) = (2, 1, 1, 1, 2, 2, 2, 2)$, provides the optimal decision rule for each state i. Thus, in the second day it is best to use a 2-dump route if the truck begins in state 1, 5, 6, 7, or 8, otherwise it is best to use a 1-dump route. In this way, a 2-dump route on the first day is most likely to end in state 4, and thus by switching to a 1-dump route on the second day we increase the probability of completing the sequence from 0.5 to 0.83. Thus the flexibility to adjust dump type increases the service level by at least 0.30 over either pure strategy.

6.2. Weekly Workload Assignment Algorithm

We now consider a fixed sequence of blocks $\{1, 2, \ldots, B\}$ for the entire ward, much more than a single truck can accomplish in a week with a reasonable service level. We split this long sequence into maximal week long subsequences (one for each truck crew) each meeting some specified service level s. Each subsequence represents the weekly workload for a truck. In keeping with their current operation, this sequence maintains a snake-like pattern through adjacent blocks, so that the travel time between blocks is minimal.

We start with the first block in the sequence and increase the number of blocks covered by the first truck until it cannot satisfy the service level s over the week. This is determined by repeated application of the Dynamic Scheduling Algorithm. This defines the weekly workload for truck 1. The next truck starts with the first block in the remaining sequence, and we again determine the maximal subsequence for this truck meeting service level s. We continue the procedure until the entire ward sequence is partitioned into weekly truck workloads, each meeting service level s.

One should note that this service level s is at a truck route level. Thus it specifies that each truck route will complete its route with probability at least s over its five-day work week. If the ward has g truck routes, then the probability that all trucks complete their routes is s^g. But operationally, this is not a clear measure of the service level of a ward, since crews that are behind in their collection can be assisted by other crews. The value s is a parameter that the ward superintendent can choose that will affect the extent of crew reassignment on Friday, i.e., the higher the value of s, the lower the probability that crews will have to be moved between routes at the end of the week.

Weekly Workload Assignment Algorithm

1. Set some sequence $\{1, 2, \ldots, B\}$ of all the blocks in a ward.
2. Specify a service level s for each truck route.
3. Initialize the starting block $i = 1$, and ending block $j = i$ for the first truck.
4. Initialize the number of trucks, $g = 0$.
5. Increment j while $j < B$ and the Dynamic Scheduling Algorithm (called with block sequence $\{i, \ldots, j\}$) returns $v_1(1) > s$.
6. Set $g = g + 1$.
7. Assign sequence of blocks $S(g) = \{i, \ldots, j\}$ to truck g.
8. Set $i = j + 1$ and $j = i$.
9. If $i < B$ GOTO Step 5, otherwise GOTO Step 10.
10. Output: g trucks required.
11. Output: Assign sequence of blocks $S(c)$ to each truck $c = 1, \ldots, g$.

In Step 1, a sequence $\{1, 2, \ldots, B\}$ is specified for all the blocks in a ward. The algorithm can be run for a number of different such starting sequences. The sequences should be chosen to keep the travel time between blocks small, so some snake-like pattern is best. The ward superintendent can create a number of such patterns through his ward that avoids excessive crossings of major roads while keeping inter-block travel times as small as possible.

For a given sequence of blocks $\{1, 2, \ldots, B\}$, the main computation takes place in Step 5. Since the Dynamic Scheduling Algorithm in Step 5 is called at most B times, a worst case bound is $O(B^3)$ steps. (A binary search can be used to set each truck route, but in the worst case this method can exceed B total calls to the Dynamic Scheduling Algorithm over all truck routes.)

7. Empirical Results and Role of the Model

The Weekly Workload Assignment Algorithm and the Dynamic Scheduling Algorithm were coded in the software package *Mathematica* (Wolfram Research Inc. 1991). Routes were computed for ward 13 with 556 blocks in under 7.9 minutes on a Sun Sparcstation 10. Total running time for all five wards was 31.5 minutes. This assumes, however, that the underlying $\Phi_{ij}(a)$

probabilities are given. The $\Phi_{ij}(a)$ are computed for each pair of blocks for a given ward block sequence $\{1, 2, \ldots, B\}$. The time to compute the $\Phi_{ij}(a)$ for either wards 13 or 23 (with around 600 blocks each), using *Mathematica* for the numerical integration, was 3.6 hours on a Sun Sparcstation 10; while the time for wards 12, 14, and 15 (with around 380 blocks each) was about 2.8 hours.[3] However, once the $\Phi_{ij}(a)$ for a block sequence are computed, the Weekly Workload Assignment Algorithm can be repeatedly solved for different service levels s.

We first report the impact of our flexible route strategy for the five pilot wards. We then give the differences in results across the wards, estimate the potential cost savings, and detail a particular truck schedule in ward 12.

7.1. Results for the Five Pilot Wards

For each of the five pilot wards we will generate a capacity required for the flexible route strategy and compare it to the capacity used in their current operation for the winter season. We first modeled the 2-dump route strategy. We selected a sequence of the blocks for the entire ward that matched, as closely as possible, the sequence actually used. We then ran our Weekly Workload Assignment Algorithm (restricting the Dynamic Scheduling Algorithm to use only 2-dump routes), adjusting the service level s until the number of truck routes required matched the actual capacity used by the ward. These service levels averaged 88% across the five wards. The total number of truck routes allocated in the current system for the five wards was 245.

We then used the same service level, and the same sequence of blocks, and ran our Weekly Workload Assignment Algorithm to provide the capacity required under our flexible routing strategy. The number of routes required by a flexible route strategy decreased from 245 to 216, a savings of 12%.

[3] The main computational burden is in evaluating the double integral using numerical methods. To guarantee very high precision, we increased the default recursive subdivision (*MinRecursion* in *Mathematica* from 0 to 4) when the numerical integration detects regions with large changes in value. One could examine univariate normal based approximations when the correlation between weight and time is high in order to decrease computation time.

By introducing 1-dump routes, the number of midday dumpsite visits decrease, thus reducing the congestion at the dumpsite. Streets and Sanitation personnel believe that this reduced congestion plus better dumpsite monitoring would allow the total dumpsite visit time to be decreased to 45 minutes (thus increasing $T(1)$ to 375 and $T(2)$ to 330). With this change, the Weekly Workload Assignment Algorithm showed a decrease in the number of routes required to under 206—a 16% reduction of capacity.

7.2. Differences in the Results Across Wards

The five wards in the pilot were chosen by personnel at Streets and Sanitation. Their justification is that the selected wards are a microcosm of the whole city which includes living units ranging from single family homes to large apartment units.

When we apply the model to these wards for the winter season, it decreases the number of current routes by 16% for wards 12 and 23, 13% for ward 14, 8% for ward 13, and 4.5% for ward 14. The savings varies from ward to ward due to differences in the composition and spatial distribution of block level parameters across wards. When one ward does better than another ward, its because it has a higher probability of exploiting 1-dump routes. Note that, the mix of 1-dump and 2-dump routes actually used in a given week is a function of the realization of the weights and times for that week. This corresponds to an overall reduction in the number of routes of 11.9%. With a reduction of 15 minutes for dumpsite visits, the capacity reduces by 16% over all wards. As we move from winter to summer, the amount of capacity required by the model increases by 12%. This increase is proportional to the increase in capacity from winter to summer currently used at the department of streets and sanitation, thus maintaining the reduction in capacity suggested by the analysis.

7.3. Estimate of Potential Cost Savings

The annual cost estimate to run a truck is $40,000 in fuel and maintenance and $150,000 in wages for the driver and laborers. Thus elimination of a truck saves the system around $190,000 annually. The current annual budget for the five pilot wards is $11.4 million. Using our model, with the savings of 15 minutes for each dumpsite visit, decreases the number of trucks across all five pilot wards by eight. This saves about $1.5 million in capacity related costs—a reduction of 13% of the current budget. A simple system wide projection generates an estimated savings of 13% of the system wide budget of $70 million—a $9.1 million annual reduction in capacity-related costs.

7.4. The Role of the Model: A Sample Schedule for Ward 12

Ward 12 is a ward with a mix of housing unit types; the number of living units per block has a mean of 53 units and a standard deviation of 30 units. The required time to collect a block in ward 12 has a mean of 36 minutes and a standard deviation of 17 minutes. Also, blocks differ in their rates of garbage collection, 57% of the blocks have rates below 61 lbs/min (indicating a large percentage of blocks have rates suitable for 1-dump routes).

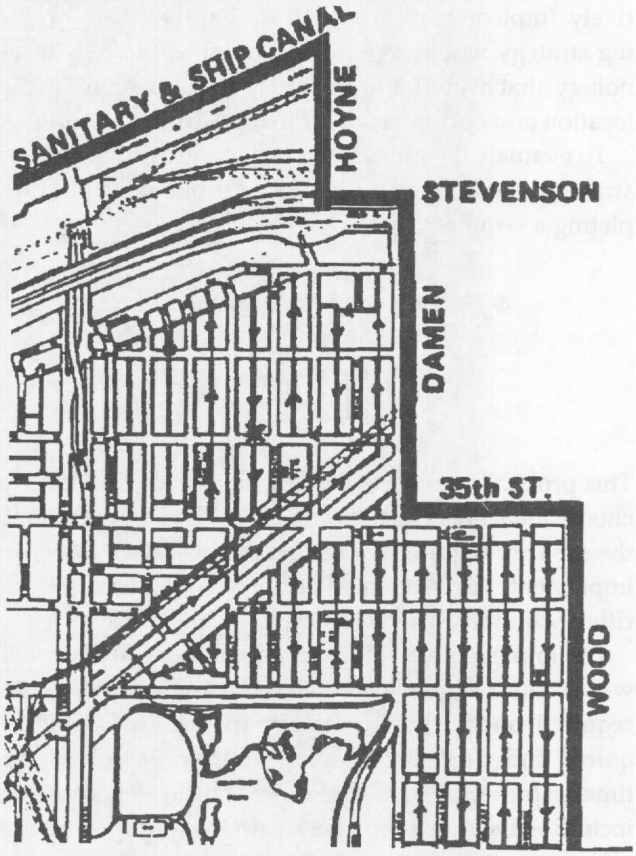
Figure 2 shows a sequence of 41 blocks, the first 36 of which is currently routed by the ward superintendent for collection by a single truck over a week using only 2-dump routes. The rates for these blocks range from 45 to 72 lbs/min. A model run, using only 2-dump routes, indicates an implied truck route service level for the week of 80%. If we permit a flexible routing scheme with the same service level, the model shows that the truck can pick up the full 41-block sequence in a week; thus gaining the collection of an additional 5 blocks.

To implement the Dynamic Scheduling Algorithm the ward superintendent would provide the truck driver the sequence of 41 blocks to collect at the start of the week. The ward superintendent would have the optimal decisions, $d_t(i)$ generated by the model. These decisions would specify for each starting block whether the driver should run a 1- or 2-dump route depending on the day of the week.

The model also allows the ward superintendent to monitor his operation in a new way. The model provides parameters which estimate the mean and standard deviation for the time and weight required to collect any sequence of blocks. Any truck route that falls outside the 3-sigma limits around the mean weight or time suggests further investigation. For instance, too high a weight may require investigation because city trucks are not supposed to be used to transport private commercial garbage (the city has

Figure 2 **The Sequence of a Weekly Schedule is Shown for One Truck in a Portion of Ward 12**

The sequence starts at 35th and Wood. In the original system (a pure 2-dump strategy) this truck completed 36 blocks, but with a flexible strategy this truck can complete 41 blocks with the same service level.



to pay the dumpsite per pound of garbage dumped). Too low a weight may suggest that the garbage was not picked up satisfactorily. Thus the model suggests a control chart like monitoring scheme that can be used by the ward superintendent.

In addition, our model allows the Director of Streets and Sanitation to analyze a number of issues faced by the department, such as the impact of moving to trucks with a different weight capacity, the impact of running longer work days over a four-day work week, and the impact of adding a new dumpsite facility.

7.5. Adaptive Flexible Routes

We noted in §1 that management prefers to run a time stamp approach in the current environment. In this section we show that we can use the models developed to examine the impact of an *adaptive flexible strategy* in which trucks would only visit the dumpsite when they are full, or at the end of the day. Under such a strategy the trucks would follow a 1- or 2-dump route adaptively. Implementation of such an adaptive flexible routing strategy would require an investment in new technology that would allow management to monitor the location and current load of the truck from their offices.

To evaluate the impact of an adaptive flexible routing strategy, we simply redefine the probability of completing a sequence of blocks,

$$\Phi_{ij} = \int_{x=0}^{T(1)} \int_{y=0}^{W(1)} \phi_{ij}(x, y)dxdy + \int_{x=0}^{T(2)} \int_{y=W(1)}^{W(2)} \phi_{ij}(x, y)dxdy.$$

This probability reflects the fact that we do not have to choose an action a at the start of the day. To determine the capacity required for a given truck service level, we implement the Weekly Workload Assignment Algorithm with these probabilities Φ_{ij}.

We implemented this routing strategy for the pilot wards. Our results showed no reduction in the capacity required under such a strategy over the capacity required under a flexible routing strategy that followed a time stamp approach. These results thus suggest that including the time stamp feature in the routing does not affect capacity in the current environment.

8. Conclusions

This paper presents a Markov decision model to dynamically schedule trucks. We explicitly model the variability in the work load faced by a truck and its impact on truck scheduling. This model shows how garbage collection in Chicago can be improved. Our model allows a ward superintendent to improve his collection in a number of ways:

- Maintains the same basic work day and work rules as their current operation.
- Provides a way to use a combination of route types—some routes that visit the dumpsite once a day, and others that visit the dumpsite twice a day.
- Maintains easy to follow, snake-like routes, through the blocks in the ward, but allows the ward superintendent to try different sequences.
- Provides a way to monitor the weight generated by a route using a control chart.

When applied to the data collected from a pilot set of five wards, the model suggests a 12–16% reduction in the number of truck routes required to pick up all the garbage in the system. The implied cost savings is around 13% of the current costs, projecting an overall potential savings of over $9 million. This paper thus shows that flexible schedules can significantly reduce the capacity required to operate a system in the presence of variability.[4]

[4] We thank Claudia Mancini and Douglas Seamen of FRAC (The Financial Research and Advisory Council) for their tremendous assistance in procuring all the relevant data. We also thank from the Department of Streets and Sanitation: Assistant Commissioner Jack Brankin, manager Douglas Ziesemer, the ward superintendents and staff of the 12th 13th, 14th, 15th, and 23rd wards for their assistance. The authors were supported by the Faculty Research Fund at the Graduate School of Business, University of Chicago. We also appreciate the valuable suggestions provided by the reviewers.

Appendix

We model the time to collect the garbage from block k in month m as a normally distributed random variable, X_{mk}, with mean collection time of t_{mk} and a variance in collection time of $\sigma^2 t_{mk}$. The parameter r_{mk} represents the rate of garbage collection in pounds per unit time for block k in month m. Thus the mean pounds of garbage picked up from block k in month m is $r_{mk}t_{mk}$ and the variance in the pounds of garbage collected is $r_{mk}^2\sigma^2 t_{mk}$. (In the body of the paper we omit the subscript m for ease of exposition.)

The procedure uses as input the driver logs from each dumpsite visit d for one week in each month $m = 1, 2, \ldots, 6$ (January, March, June, July, September, and November) through the 1992 calendar year.[5] For each dumpsite visit d in month m, these logs provide the weight, $W(d, m)$, the time, $T(d, m)$, and the number of blocks, $B(d, m)$ collected. Also as input, is the number of living units $L(d, m, k)$ for each block k for $1 \le k \le B(d, m)$. The output of this procedure is an estimate of the garbage collection rate for each block, an average time for each block and variance in collection time for each block.

We first generate a disaggregated weight, time and rate estimate for each block k in dumpsite visit (d, m) as

[5] Since these driver logs were only available in hard copy form, all the data had to be manually entered into the computer. Hence we used only six weeks of detailed data—2200 driver logs.

$$r'(d, m, k) = \frac{W(d, m)}{T(d, m)},$$

$$t'(d, m, k) = \frac{T(d, m)L(d, m, k)}{\sum_{i=1}^{B(d,m)} L(d, m, i))},$$

$$w'(d, m, k) = r'(d, m, k)*t'(d, m, k).$$

The result of this procedure is to generate six values for the weight, time and rate estimates for each block for each of the six weeks of drivers logs (each block in the ward is collected once a week). We let

$$\bar{t}(k) = \frac{\sum_{d,m} t'(d, m, k)}{6}$$

estimate the average time required to collect block k, and

$$\bar{r}(k) = \frac{\sum_{d,m} r'(d, m, k)}{6}$$

estimate the average rate required to collect block k. Examination of the rates across all blocks shows a seasonality pattern. We use the aggregate garbage generated by these pilot blocks over the six different months to generate a seasonality factor defined for month m as

$$s(m) = \frac{6 \sum_d W(d, m)}{\sum_{d,i} W(d, i)}.$$

Thus we now have our estimate of rate for a block k as

$$r_{mk} = s(m)\bar{r}(k).$$

Our estimate of time for a block k as

$$t_{mk} = \bar{t}(k).$$

Our estimate of weight for a block k as

$$w_{mk} = \bar{t}(k)r_{mk}.$$

The last step is to generate an estimate for σ^2. In order to estimate this value, we first used the values $t'(d, m, k)$ to generate a standard deviation in the collection time for block k. This value is used to generate an estimate of σ_k as

$$\sigma_k = \frac{\text{Standard Deviation across } t'(d, m, k) \text{ for block } k}{\sqrt{\bar{t}(k)}}.$$

We generate a block independent estimate of σ as the average across the estimates generated across all the blocks k i.e.,

$$\sigma = \frac{\sum_k \sigma_k}{\sum_k 1}.$$

This average value of σ was estimated from the data as $\sigma = 1.3$.

References

Beltrami, E. J. and L. D. Bodin, "Networks and Vehicle Routing for Municipal Waste Collection," *Networks*, 4 (1974), 65–94.

Biro, M. and E. Boros, "Network Flows and Non-Guillotine Cutting Patterns," *European J. Oper. Res.*, 16 (1984), 215–221.

Christofides, N. and C. Whitlock, "An Algorithm for Two-Dimensional Cutting Problems," *Oper. Res.*, 25 (1977), 30–44.

Eglese, R. W., "Routing Winter Gritting Vehicles," *Discrete Applied Math.*, 48 (1994), 231–244.

Golden, B. L. and A. A. Assad, *Vehicle Routing: Methods and Studies*, North-Holland, Amsterdam, 1988.

—— and R. Wong, "Capacitated Arc Routing Problems," *Computers and Oper. Res.*, 10 (1983), 47–50.

Imai, H. and T. Asano, "Finding the Connected Components and a Maximum Clique of an Intersection Graph of Rectangles in the Plane," *J. Algorithms*, 4 (1983), 310–323.

Maehara, H., "A Digraph Represented by a Family of Boxes or Spheres," *J. Graph Theory*, 8 (1984), 431–439.

Male, J. W. and J. C. Liebman, "Districting and Routing for Solid Waste Collection," *J. Environmental Engineering Division*, ASCE, 1–14, 1978.

Puterman, M. L., "*Markov Decision Processes: Discrete Stochastic Dynamic Programming*," John Wiley, New York, 1994.

Riccio, L. J., "Management Science in New York's Department of Sanitation," *Interfaces*, 14 (1984), 1–13.

——, J. Miller, and A. Litke, "Polishing the Big Apple: How Management Science Has Helped Make New York Streets Cleaner," *Interfaces*, 16 (1986), 83–88.

—— and A. Litke, "Making a Clean Sweep: Simulating the Effects of Illegally Parked Cars on New York City's Mechanical Street-Cleaning Efforts," *Oper. Res.* 34 (1986), 661–666.

Roberts, F. S., "On the Boxicity and Cubicity of a Graph," *Recent Progress in Combinatorics*, Academic Press, New York, 301–310, 1969.

Rohatgi, K., *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley, New York, 1976.

Accepted by Edward H. Kaplan; received April 2, 1995. This paper has been with the authors 3 months for 2 revisions.