AN ANALYSIS OF THE ASSIGNMENT OF DELIVERY ROUTES TO VEHICLE
DRIVERS IN STOCHASTIC VEHICLE ROUTING OPERATIONS

Michael A. Haughton
School of Business and Economics
Wilfrid Laurier University
Waterloo, Ontario, Canada N2L 3C5
Telephone: (519) 884-1970 (Ext. 6205)  Fax: (519) 884-0201
Email: mhaughto@wlu.ca

Abstract

Random day-to-day fluctuations in customer demands extend the range of decisions to be made by managers of vehicle routing/dispatch operations. For one, dispatch/routing managers must decide how responsive the delivery routes should be to the stochastic demands. But even with that decision settled –often by using daily route reoptimization to maximize responsiveness– the assignment of drivers to the reoptimized delivery routes must also be determined. In the interest of customer service, managers may use driver-to-route assignment rules that ensure that the driver who is historically most familiar with a given customer will most likely be chosen to continue serving the route that that customer is on. Using data from several vehicle routing scenarios, this paper presents a statistical analysis of one such decision rule, and uses the analysis to derive managerial implications of rules that seek to maximize customer-driver familiarity. The paper also provides some preliminary insights on the potential for Markov Chains in modeling driver-to-route assignment decisions.

Abstract Number: 002-0007

Introduction

This paper compares two approaches for assigning drivers to delivery routes in situations where daily route reoptimization is used in dealing with random day-to-day fluctuations in customer demands. One approach, which can be thought of as a baseline approach, is that for any given route (group of customers) on any given day each available driver has the same probability of being assigned to it. This may reflect a situation of uniform competence across drivers, making it irrelevant which driver gets which route. The second assignment approach is to ensure that each day, the driver who is historically most familiar with a given customer will most likely be chosen to serve the route that that customer is on. A supporting logic for this latter approach is that increasing a driver’s familiarity with the specifics of a given customer, not only increases the efficiency with which the customer can be served but is also enhances customer service.

Examination of these approaches is relevant in several practical settings. For example, in supply chains that feature stochastic customer demands, the day-to-day decision problem of driver-to-route assignment typically occurs at the retailer/end-customer interface or echelon. Another context of random customer requirements that gives rise to the decision problem is among courier companies in both their delivery and pick-up of small packages. These issues relate to two broad segments of the scientific literature: vehicle routing/dispatch under stochastic demands and workforce scheduling. The former category has largely focused on developing solution methodologies for the mathematical programming representation of the routing problem. Thus, many relevant matters concerning the assignment and scheduling of drivers remains unexplored. For example, the potential problem of instability in the day-to-day duties of drivers, while long acknowledged in the literature –e.g., Bertsimas (92), Waters (89)– has only recently been subjected to formal measurement –Haughton (2002).

The literature segment on workforce scheduling literature is also dominated by efforts to develop heuristic/algorithmic solutions. See, for example, Xu and Chu (2001), Ahire, Greenwood, Gupta, Terwilliger (2000), Gans and Zhou (2002), Brusco (1998). In that literature, issues concerning the link between the workforce assignment/schedule and the
worker’s *learning* (via, e.g., the familiarity that comes with frequent performance of a task) and *forgetting* (via, say, long intervals between working on the task) appear to have been broached only peripherally and/or as a matter for future research; e.g., Quintana and Ortiz (2002). Thus, neither of these two literature categories has dealt with the core question of this paper: characterizing the impact of driver assignment policy on three items in a stochastic vehicle routing context: (a) *learning* (through frequent delivery visits to a customer), (b) *forgetting* (through long intervals between visits to the customer), and (c) customer service.

The remainder of this paper further articulates both the problem context and the two aforementioned driver-to-route assignment policies, describes the simulation experiments used to compare them, presents and discusses the comparative results from the experiments, and outlines some planned extensions of the work.

**Problem Context**

Consider a situation in which $N$ customers, each with delivery requirements (demands) that fluctuate each day according to some known distribution. The customers are to be served by a fleet of vehicles, each with capacity $Q$. For simplicity, assume that the number of available vehicles is enough to handle the maximum possible value of total demand across all customers on any day; i.e., the fleet size is equal to $\max\left[\sum_i r_i\right] \div Q$ (rounded up to the next integer), where $r_i$ is the $i^{th}$ customer’s demand. The number of available delivery vehicle drivers, denoted $J$, is assumed to be similarly adequate. Each day, once a vehicle routing algorithm has been run to ascertain the optimal number of routes (customer groupings), the specific customers on each route, and the optimal intra-route sequences of visits to customers, attention can then turn to deciding which driver should serve each route. That is, denoting each driver (and vehicle) as $j, j = 1$ to $J$ and each route as $k, k = 1$ to $K, (K \leq J)$, the issue is to select from among the $J! \div (J-K)!$ possible assignments. An implicit assumption is that the vehicle assignment is a not of immediate concern; which is equivalent to assuming a pre-existing 1-1 link between
vehicle and driver. Of possible interest is how the chosen assignment rule affects statistics concerning:

(a) The proportion of visits a customer gets from each driver
(b) The duration of continuous visits by a driver to a given customer (perhaps as a gauge of the driver’s opportunity to become familiar with relevant information on the customer and thus be proficient in making his/her deliveries; i.e., the benefit of continuously learning)
(c) The duration of periods between visits by the driver to a given customer (perhaps to gauge the risk that a driver might forget relevant information on the customer and thus lose some proficiency in completing his/her deliveries; i.e., the problem of forgetting)

Driver Assignment Rule 1: Random
This baseline rule views the decision problem as equivalent to tossing a coin. That is, for any given route, \( k; k = 1, 2, \ldots K \), each driver has the same probability of being assigned to it. Equivalently, the probability that a particular customer will be served by a particular driver is \( 1/J \). With Rule 1, the day-to-day situation for each customer can be modeled as a \( J \)-state Markov process in which State \( j \) \( (j = 1, 2, \ldots J) \) is reached when the customer is served by a driver \( j \). From this simple process, it can be easily shown that the steady-state results for the statistics of interest are:

(i) The probability of the customer being in any one of the \( J \) states is \( 1/J \) (i.e., the proportion of visits that the customer gets from each driver \( \cong 1/J \)); thus, for a sufficiently long horizon of \( T \) days, the mean number of visits that the customer gets from each driver \( \cong T/J \), and variance \( \cong T(J-1)/J^2 \).
(ii) The customer’s mean duration of continuous visits (i.e., unbroken sequence of visit) from any specified driver is \( J/(J-1) \) and variance \( \cong J/(J-1)^2 \).
(iii) The customer’s mean time between visits from any specified driver will average \( J \) days with variance \( \cong J(J-1) \).
Driver Assignment Rule 2: Priority

This priority rule can be formally described by first defining the following variables:

\[ x_{kj(T+1)} = \begin{cases} 1, & \text{if route } k \text{ is served by driver } j \text{ on day } T+1, \\ 0, & \text{otherwise} \end{cases} \]

\[ x_{ij(T+1)} = \begin{cases} 1, & \text{if customer } i \text{ is served by driver } j \text{ on day } T+1, \text{ (i.e., customer } i \text{ is on route } k), \\ 0, & \text{otherwise} \end{cases} \]

\[ c_{ijT} = \sum_{t=1}^{T} x_{ijt} \quad \text{cumulative number of visits to customer } i \text{ by driver } j \text{ up to day } T \]

\[ c_{kjT} = \sum_{i} c_{ijT} \quad \text{cumulative number of visits to all customers on route } k \text{ by driver } j \text{ up to day } T; \ n_k \text{ defines the set of customers on route } k; \ i.e., \ i \in n_k \]

Since the problem on the current day, \( T+1 \), is to maximize the familiarity between customer and driver by assigning each route to the driver who is most familiar with (has had the most previous visits to) the set of customers on that route, it can be formulated as:

\[
\text{Maximize} \sum_{k} \sum_{j} c_{kjT} x_{kj(T+1)}
\]

\[ \sum_{k} x_{kj(T+1)} \leq 1; \ \forall j \quad (2a) \]

\[ \sum_{j} x_{kj(T+1)} = 1; \ \forall k \quad (2b) \]

The constraints in (2a) and (2b) ensure that, respectively, each driver serves no more than one route, and each route gets served by exactly one driver. In the experiments described in the next section, the problem was solved each day by using a greedy algorithm. A partial illustration of the procedure is illustrated by the example in Table 1. In this example, the dispatcher (router) is planning for day 6, and must therefore decide how to assign three routes across three drivers. A greedy algorithm which starts by selecting which of the \( K \times J \) driver-route pair has the largest familiarity measure \( (c_{kj}) \) then continues by sequentially selecting the pair with the largest remaining familiarity measure would yield B-3, C-2, A-1 as the assignment for day 5. The greedy algorithm would be done for each route, and be repeated each day to maximize overall driver-customer familiarity. The emphasis is on overall because, as the illustrative example shows, Andy is not familiar with all four customers on the route he is assigned to on day 6 (prior to day 6, he would have not yet visited customer D) even though Bobby is familiar with all four customers.
Table 1: Illustration of Assignment Rule 2

<table>
<thead>
<tr>
<th>Driver</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Totals = $c_{1j5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andy</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Bobby</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Charlie</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Driver</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Totals = $c_{2j5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andy</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Bobby</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Charlie</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Driver</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>Totals = $c_{3j5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andy</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Bobby</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Charlie</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Simulation Experiments: Design and Results

The experimental conditions for the simulation comprised the following:

(a) an actual 100x100 km² area covering a road network comprising several adjoining cities in Southwestern Ontario, Canada.

(b) 100 customers, randomly selected from 1000 actual customer addresses that are approximately uniformly distributed throughout the area (the experiment was replicated 5 times by repeating the procedure of randomly selecting 100 addresses from the total of 1000.

(c) Each customer’s daily demand is independent and follows an identical normal distribution with $(\mu, \sigma) = (100, 30)$ units.

(d) Vehicle Capacity ($Q$) = 1000 units (12 vehicles/drivers available so $J = 12$)

(e) Simulation length ($T$) = 50 days
The routing was done with a commercial grade software program called Roadshow® from Descartes Corporation. The software does not have any goal-oriented driver-to-route assignment routine such as maximization of driver-customer familiarity so that aspect of the experiments was handled after the program generated the routes. The routines for calculating the output statistics of interest were coded using Dev C++. Table 2 shows the summary results for each assignment rule. For completeness, the table also shows the expected steady state (Markov based) results for Rule 1 in parentheses beside the observed simulation results. The steady state and the observed results are nearly identical except for the duration of breaks between delivery visits, where the observed statistics are below the steady state statistics. The explanation for the conspicuous difference lies in the fact that the duration of breaks between delivery visits, which is approximately exponentially distributed, is subject to great variation across driver-customer pairings (i.e., extremely large values). Thus, while the steady state parameters would account for the occurrence of these extremes (e.g., breaks exceeding 50 days), a simulation of length $T = 50$ days will not adequately capture the extreme values. Since the present work is exploratory, and the transitory state results are enough to educe genuine differences between the two driver assignment policies, the increase in the run length is left as a matter for the follow-up paper containing more complete analysis.

<table>
<thead>
<tr>
<th>Statistic (across all non-zero driver-customer pairs)</th>
<th>Rule 1</th>
<th>Rule 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean proportion of delivery visits per driver-customer pair</td>
<td>0.083 (0.083)</td>
<td>0.084</td>
</tr>
<tr>
<td>Standard deviation in the proportion of delivery visits</td>
<td>0.044 (0.038)</td>
<td>0.174</td>
</tr>
<tr>
<td>Mean duration of continuous delivery visits</td>
<td>1.051 (1.091)</td>
<td>1.790</td>
</tr>
<tr>
<td>Standard deviation in the duration of continuous visits</td>
<td>0.303 (0.315)</td>
<td>1.673</td>
</tr>
<tr>
<td>Mean duration of breaks between delivery visits</td>
<td>10.16 (12.00)</td>
<td>12.09</td>
</tr>
<tr>
<td>Standard deviation in the duration of breaks</td>
<td>8.636 (11.49)</td>
<td>12.41</td>
</tr>
</tbody>
</table>

Notes: (1) Steady-state Markov-based estimates for Rule 1 are shown in parentheses
(2) Customer-driver pairs for which no visits occurred are excluded from the calculations

With respect to the mean values, a noticeable (and statistically significant) difference between the two policies is that Rule 2 gives a customer the benefit of a longer run of continuous visits from the same driver. That is, drivers might have benefit of ongoing
learning about the customer through greater continuity in delivery visits. However, against the benefit of learning, drivers under Rule 2 may face the contrasting problem of "forgetting" since the mean duration of the break once a run of visits is broken is longer by about two days (12.09 versus 10.16). Thus, if these transitory state differences between Rule 1 and Rule 2 hold in steady state then, in a practical setting, the problem of deciding between the two assignment rules might hinge on how the impact of learning forgetting compares with the impact of continuous learning.

As the standard deviations show, the variability across driver-customer pairings is noticeably greater for Rule 2. The intuitive explanation is that under Rule 2, each driver tends to have extremely good mean values for a relatively small subset of customers (more visits, longer unbroken visit sequences, and shorter breaks between visits) but extremely poor values for a larger subset of customers. As depicted in Figure 1, this means that the distributions under Rule 2 (bottom panel) exhibit greater variation and are more heavy-tailed than those under Rule 1 (top panel).

Figure 1: Graphical Depiction of Rule 1 (top panel) and Rule 2 (bottom panel)
As a practical matter, this imbalance associated with Rule 2 might raise an important concern. That is, if the driver that makes the bulk of the visits to a particular customer is unavailable then would the driver that is next most familiar with the customer be an adequate substitute? An indirect examination of this question involved analysis of the distribution of visits across each customer’s three most familiar drivers. The simulation results showed that typically, a customer’s most familiar driver served that customer on 32 of the 50 days (64% of the time in a range of 40%-80%), with the second, and third most familiar drivers visiting the customer, respectively, only 4 and 3 days (i.e., 8% and 6% of the time). The outcome of the simulation was that the other drivers would typically visit that customer no more than once. Thus, while by design and outcome, Rule 2 ensures that each customer has a tight relationship with a particular driver; it could incur the risk of not allowing other drivers to develop an adequate knowledge of that customer.

Interestingly, while this would mean that a driver’s unavailability would adversely affect the driver’s most favourite customer, Rule 2 may be able to mitigate the effect on a group of customers to a greater degree than Rule 1! Consider, for example, the group of eight customers that a driver is most familiar with. The simulation results showed that the driver typically visited that group an average of 27 days (54% of the time), while the drivers that are, respectively, the second and third most familiar with that group visited them 18% and 8% of the time. Now, although 18% is a far cry from 54%, it is larger than the estimate of 8.33% for Rule 1 (i.e., 1/J for all drivers since no tight customer-driver relationships are formed under Rule 1). This result suggests that, under Rule 2, the unavailability of one driver might not have serious adverse effects across a group of customers. True, because the second best driver would not have a uniformly high level of familiarity with all customers in the group (on the route) some inefficiency in serving the route can result. But again, the magnitude of that inefficiency may be greater with Rule 1.

---

1 The reason for using 8 to illustrate the point is that 8 is the mean number of customers on a route; i.e., a given group of approximately N/J customers can be viewed as comprising the driver’s most favourite route.
Conclusion

Using transitory state results from simulating a particular set of experimental conditions, this paper explores two rules for assigning drivers to routes in stochastic customer demand settings. The rules were viewed in terms of statistics and related considerations that are likely to be of practical interest in vehicle routing/dispatch operations; e.g., the ease with which efficient delivery operations can be maintained when a driver is unavailable. The results suggest some superiority of the rule of engendering tight driver-customer relationships (Rule 2) over the rule of random (balanced) rotation of drivers across customers (Rule 1). Further work is being undertaken to see how the results hold up to changes in the experimental conditions; e.g., greater demand variability, larger vehicle capacity and thus larger groupings of customers on each delivery route, larger number of customers, and thus a larger number of delivery routes, etc. In particular, the results on Rule 2 suggest a particular pattern in the percentage of visits that a customer gets from each driver (from 40%-80% for the customer’s most familiar driver to about 8% for the next most familiar driver). It will be interesting to see how these results are affected by changes in experimental conditions. If these percentages remain robust then, being estimates of the steady state $\pi_j$ (proportion of time a customer is served by driver $j$), they can be used to readily generate the Markov statistics of interest; e.g., duration of breaks between visits, etc. for each driver that the customer encounters. Table 3 shows what the steady state results might be for a modified version of an observed pattern from the simulation experiments: $\pi_1 = 0.70$, $\pi_2 = \pi_3 = \ldots = \pi_{12} = (1-0.70)/11$.

Table 3: Steady-state Markov results for Rule 2

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Favourite Driver</th>
<th>Other Drivers</th>
<th>All Drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean duration of continuous delivery visits</td>
<td>3.333</td>
<td>1.028</td>
<td>1.220</td>
</tr>
<tr>
<td>Standard deviation in the duration of continuous visits</td>
<td>2.789</td>
<td>0.555</td>
<td>1.156</td>
</tr>
<tr>
<td>Mean duration of breaks between delivery visits</td>
<td>1.429</td>
<td>36.67</td>
<td>33.73</td>
</tr>
<tr>
<td>Standard deviation in the duration of breaks</td>
<td>0.782</td>
<td>19.81</td>
<td>21.32</td>
</tr>
</tbody>
</table>

Another extension of the work involves exploring variations of Rule 2. One example is a structured rotating assignment scheme to ensure that multiple drivers (say 2 to 4) are very familiar with each customer. Many key insights of this goal of worker flexibility via task
rotation (as well as related initiatives such as cross-training) are well understood in the literature; e.g., Brusco and Johns (1998). However, the potential impacts of assignment rotation in vehicle routing contexts, particularly impacts on the efficiency of route designs, warrant exploration.

References
7 Quintana, Rolando, Juan G Ortiz, "Increasing the effectiveness and cost-efficiency of corrective maintenance using relay-type assignment", *Journal of Quality in Maintenance Engineering*: 2002. Vol. 8, Iss. 1; pp. 40-61