Real-Time Order Acceptance in Transportation Under Uncertainty

By: Hiral Nisar and Joshua Rosenzweig
Thesis Advisor: Dr. Chris Caplice

Summary: Many industries require the capability to accept, reject, and prioritize incoming orders under uncertain future conditions. Some industries, like make-to-order manufacturing and transportation, use order acceptance criteria often based on capacity constraints and order characteristics to make these real-time decisions. In the transportation domain, orders that are in close spatial and temporal proximity are often aggregated to achieve economies of density and scale. The problem is that companies don’t know if future demand will allow them to effectively combine shipments. The objective of our research was to create and validate a model to determine if historical demand data can be used by retail firms operating private fleets to make effective real-time order acceptance/rejection decisions with the purpose of eliminating unprofitable orders in a short-haul transportation setting. A Java tool was developed to instantaneously decide whether or not to accept an order depending on the order location and time of receipt. The tool was then tested against several alternative models.

Hiral Nisar received a B.S. in Computer Engineering. Prior to the SCM program, she worked as a technology consultant in the IT industry in India for 8 years, developing software and implementing products.

Joshua received a B.S. in Mechanical Engineering. Prior to the program, Joshua worked for Teach for America as a high school mathematics and science teacher in Waianae, Hawai‘i.

KEY INSIGHTS

1. Comprehensive order acceptance criteria are critical for creating an effective decision making model.

2. Demand probability distribution is an important component for use in real-time operational decision making processes.

3. Logistic regression analysis is not appropriate for an order acceptance model, as it does not provide the flexibility required to incorporate capacity constraints.

The Problem

In transportation, cost is minimized by aggregating shipments in close spatial proximity. Accepting orders indiscriminately can put unwanted constraints on the ability to combine shipments. Firms may have to reject incoming orders due to limited capacity. Many industries use order acceptance criteria based on order characteristics, customer priority and capacity constraints. However, minimal research has been done on order acceptance criteria involving probabilities of order deliveries in sub-regions based on historical demand. In this paper, we develop a decision making model based on the probability distribution of orders across service regions, marginal cost of accepting incoming orders, and truck capacity.

Methods

To formulate an order acceptance model based on probabilities, we first determined the criteria that have the largest impact on transportation cost. We found that an order’s destination region most directly affects the transportation cost. Profitability of an order requires that marginal revenue exceed marginal cost for a region; therefore a certain amount of order revenues must be received to offset the cost of delivering to a specific region. Using a transportation cost approximation formula, the breakeven number of orders for a region, where revenues equal costs, was calculated. The model accepts orders for a certain time period, which we call the order acceptance period. After the end of this period, the actual delivery of the orders is executed. However, exact costs from vehicle routing are not explored in this thesis.
The two major criteria for the model are the breakeven number of orders and the probability of receiving that many orders during the time left in the order acceptance period. Available truck capacity is also an important consideration when developing the model. Figure 1, below, demonstrates the structure of the probabilistic model algorithm.

![Decision making tree](image)

All incoming orders for regions that are already being serviced will be accepted if there is available capacity, as the majority of costs have already been incurred with the acceptance of an initial order. Incoming orders in new regions must be checked for expected profitability. The node labeled “Profitable?” in the decision making tree above will determine the profitability of deploying a truck to a new region. In the event that a region has no prior orders, the model finds the likelihood of receiving breakeven number of orders in the remainder of the order acceptance period. If the probability of receiving the breakeven number of totes is greater than a predetermined threshold and capacity is available, the model will accept the order and subsequent orders in the same region until capacity is fully utilized.

There are several key calculations at the core of the probabilistic model. The breakeven number of orders requires a calculation of cost. Total cost comprises purchase cost, transportation cost, and a fixed operating cost. Purchase cost and operating cost are considered constants in our model, with a fixed purchase cost per tote and a fixed operating cost per truck.

The nature of the model requires that orders are constantly being added to regions, which will ultimately have an effect on the actual route driven at the end of the acceptance period. Because we could not include route optimization in our model due to data constraints, exact transportation cost was impossible to determine. For this reason, we modified an equation from Daganzo (2010) to calculate an approximation of transportation cost upon addition of each order. The equation uses number of customers and customer density of a region to estimate the costs associated with loading and unloading totes, and traveling to and from a region.

After the breakeven number of orders is known for each region, the likelihood of receiving that number of orders must be determined. For the purposes of this model, the frequency of deliveries per customer was assumed constant. This assumption is crucial in allowing the model to determine the number of totes a region can expect to receive in a certain time period. By knowing the total number of customers in a region, a Poisson distribution analysis was done to compute the expected number of totes in a certain time interval. The Poisson output tells us the probability of receiving exactly \( k \) orders in a time period given an expected value \( \lambda \). In this situation, however, the model determines the probability of receiving at least \( k \) orders.

Combining the Poisson distribution analysis with the breakeven analysis, the model is able to determine the likelihood of receiving enough orders to reach the breakeven point for each region. Orders for regions that are not expected to reach the breakeven number of totes in the time left will be rejected.

**Alternative Models**

Determining the effectiveness of the probabilistic model required a comparison of results of various alternative strategies and a baseline model. The models developed range from overly simple to complex and are described below.

**Optimal Model Overview**

In our problem, decisions must be made instantaneously upon order arrival. In an optimal situation, we would have complete knowledge of all daily orders before making decisions. This model optimizes capacity utilization and maximizes profit and provides a baseline for model comparison. By comparing our model against results from the optimal situation, we can determine the effectiveness of the decision making process in our model.

**Myopic Manager Model Overview**

In this model, orders are accepted sequentially until capacity is filled. As in our probabilistic model, decisions are again made instantaneously upon receipt of orders. This model reflects the actions of a firm with no intelligent operational decision making process and simulates a worst-case scenario.
Logistic Regression Model Overview

The binary logistic regression uses explanatory variables and previous decisions to predict the discrete outcomes of future decisions in order to make instantaneous order acceptance decisions. To determine the previous decisions to use for the regression formulation, we used binary outputs from the optimal solution model on a dedicated dataset. In our model, explanatory variables used to create the model are the incoming order’s distance (miles) from the distribution center and order size (number of totes).

Results

The results in Figure 2, below, show that the maximum daily profit earned for the optimal strategy is much higher than that of any other strategy. This is because it has complete information of the demand and only chooses those orders that earn higher revenue and maximize truck utilization.

The maximum daily profit earned by the probabilistic model is lower than that of the optimal model. The profits increase linearly with fleet size until it reaches a peak and then starts decreasing due to the penalty of not utilizing all the available trucks. This is because the model is highly influenced by time and does not deploy new trucks after a certain point in the acceptance period.

The daily profits for the myopic manager model are even lower than the probabilistic model. This is because the model deploys trucks in the regions for earlier incoming orders. As a result, it can accept an order for a region which does not have a high customer density and will cause most of the truck’s capacity to be underutilized. Therefore, the company will lose money in these unprofitable regions and report lower profits.

The logistic regression analysis offers the poorest performance. One factor causing low profits is that the model does not adequately describe the situation with the available explanatory variables. This attributes to model accuracy of approximately 76% correct decisions. The regression model actually performs worse than the myopic manager model because the nature of the regression parameters allows for orders with small numbers of totes to be automatically rejected even if there is already available capacity to the same region. The inability to incorporate capacity into the regression analysis adds unfavorable inflexibility to the model.

Table 1: Comparison of average daily profit with fleet size of 13

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Average Daily Profit</th>
<th>Difference from Optimal</th>
<th>Percent Difference from Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>2894</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Myopic Manager</td>
<td>1792 (1102)</td>
<td>(38.1)</td>
<td></td>
</tr>
<tr>
<td>Probabilistic</td>
<td>2691 (203)</td>
<td>(7.0)</td>
<td></td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>1652 (1242)</td>
<td>(42.9)</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion

Inputting the same simulated demand data sets into each Java tool, we were able to determine that our probabilistic model provided approximately 8% less profit than the optimal solution under a scenario of flexible fleet size; however, it outperformed all other instantaneous decision making models. Average daily profits for the myopic manager model and the logistic regression model were about 11% and 34% lower than the probabilistic model, respectively. We determine that the myopic manager and logistic regression models simply do not offer the insights required to match supply with demand in an effective way.

We are able to conclude that demand probabilities are an important variable for use in real-time operational decision making processes. This research shows that demand probabilities determined by historical demand patterns should be considered by companies with private fleets that make acceptance/rejection decisions under capacity constraints.
Future Research

To understand how the model would function in a more realistic setting, more research must be done to amend the model to remove naïve assumptions.

First, assumptions surrounding customer order frequency, order size, and order revenue should be refined through the use of additional data in related fields. This would allow for the employment of more realistic probability distributions in the creation of demand samples. Second, the model should be adjusted to permit trucks to deliver to more than one region. This would allow the model to be more favorable to orders for low customer density regions that are in close proximity to high density regions. Third, the cost for order rejection must be included in the model. The cost associated with outsourcing tote delivery should be incorporated in the regional break even analysis.

In addition to relaxing assumptions, future research should be done to create accurate models of current practices of retail companies operating private fleets. In our thesis, we simplify the actions of these companies with the myopic manager model, a worst-case scenario in which companies will simply accept every order until capacity is unavailable. This is most likely not true for many companies, as they may adapt strategies and develop decision making heuristics over time. By investigating the real processes used by companies, we can more accurately assess the true benefits of the model.