Supply Chain Metrics for Dynamic Routing

Douglas A Popken, PhD

Dynamic routing is defined in this article as a shipping strategy that allows the originally planned route of a truck to change after the route has already started. The strategy is motivated by a desire to respond to new transportation demands more quickly. It requires use of a freight tracking system that determines the exact location of the truck so that the ability to service the new demands can be accurately assessed. It provides the greatest benefit when integrated with automated route-building software. Dynamic routing can improve customer service by reducing the lead time between placing an order and pickup of the shipment, thereby improving the chances of rapid delivery. It can also save operating costs by making it unnecessary to dispatch an additional truck when an existing truck can service the demand with perhaps only a small deviation in its current route. However, it makes both route management and performance measurement more difficult.

Dynamic routing is sometimes used by LTL (Less Than Truckload) carriers, particularly when high priority shipments are involved. Truckload carriers can also employ a dynamic routing strategy as a means of avoiding "dead-head" (no load) route segments. Other applications include auto-carriers and "dial-a-ride" type people moving systems (shuttles, taxis, etc.)

A metric is a verifiable measure of company performance. Metrics can be either quantitative, such as percentage of late orders, or subjective, such as providing above-average service according to customer satisfaction surveys. Metrics can be used for:

- learning and understanding causes of variance
- identifying opportunities for improvement
- control and evaluation of employee performance
- reporting to internal and external units
- allocating resources

Dynamic routing applications provide special challenges in developing effective metrics. Measures appropriate for static routing, such as utilization, are insufficient at best, and often misleading. We propose using two related capacity-based metrics and three time-based metrics.

Capacity-Based Metrics

The first question is - how to measure capacity? For dynamic routing applications, we propose that the most relevant capacity measure is ton-miles, which is the product of tons carried times the number of miles traveled. For example, a truck that holds 40 tons and travels 10,000 miles on a route, contributes 400,000 ton-miles to system (fleet) capacity. (Note that for people-moving applications, we could substitute "passenger-miles" for "ton-miles" in the discussion without change to the analysis. For auto carriers, "vehicles-miles" could be used.) An individual shipment uses an amount of system capacity equal to its weight in tons times the actual distance that it travels on a truck. To assess performance, capacity use must be measured with respect to some reference point. The reference points are provided within the definitions of the proposed capacity-based metrics – *utilization* and *efficiency*.

Utilization

Utilization refers to the proportion of system capacity being used, ranging from 0 to 1.0. For example, if all the trucks in a fleet are on average ½ full as they travel about, then the average system utilization would be approximately 0.50. We generally like to have high values for utilization, as this indicates that we are making use of expensive resources. However, shipping is no different than many other industries, where a single-minded pursuit of high utilization would be detrimental. We need to make sure that capacity is also being used efficiently, as discussed below.

Efficiency Ratio

Efficiency refers to actual capacity use in relation to a minimum. Above, we saw that a single shipment uses an amount of capacity (ton-miles) equal to its weight in tons times the distance it travels on a truck. A shipment uses the minimum amount of system capacity if the truck(s) carrying the shipment proceed directly from the shipment origin to its destination. Since a truck may make a number of stops to service assigned shipments, any given shipment may take a relatively circuitous route as it travels to its final destination. We can measure the efficiency of the travel *for a given shipment* by taking the ratio – (actual travel distance / direct travel distance). This metric takes on the value of 1.0 for a perfectly efficient (direct) shipment route and larger values for more circuitous shipment routes. Since a single truck route will likely be relatively efficient for some shipments and less so for others, the efficiency metric is most useful when averaged over a large number of orders.

Another way of looking at the efficiency is as a measure of "excess utilization". Suppose that a shipment contributes .18 to the utilization of a truck on a given route, while the efficiency of the route with respect to that shipment is 1.50. If the efficiency was instead 1.0, then the shipment would have contributed (.18/1.5) = .12 to the utilization. The difference, .18 - .12 = .06 is the excess utilization. The excess utilization is also a form of "opportunity cost"; potentially, it could have been employed to move additional shipments on the same truck had the route been more efficient.

Utilization and Efficiency Tradeoffs

The best situation is routes that have an efficiency ratio and utilization both near 1.0. This is most easily achieved on direct routes between two points with naturally occurring high volume flows - a situation that may not be possible with current customer demand patterns. However, suppose that a truck fleet has good (close to 1.0) average utilization rates, while the average efficiency of the shipments is relatively poor (much greater than 1.0). We would observe that the trucks are traveling nearly fully loaded at all times and would be tempted to conclude that the routes must be near ideal. But high truck utilization does not always imply a well-run system. We may actually observing high levels of "excess utilization" as defined above. With more efficient routing, it may be possible to service the same shipments with lower average utilization. This opens up additional space on board the trucks to take on additional shipments, as we would expect to occur in dynamic routing situation. This can in turn lead to higher customer satisfaction levels as future pickup and delivery are more likely to occur earlier.

Time Based Metrics

Time based metrics provide a way of measuring system performance with respect to customer expectations of shipping timeliness. Three metrics are proposed, two of which are related: Proportion of Late Orders and Average Lateness (of Late Orders). Lateness is determined with respect to the Due Date provided to the customer when the order was placed. The third metric, the Travel Time to Lead Time Ratio, is related to the actual setting of due dates.

Lateness Metrics

Determining whether an order is late is a simple matter of comparing the due date and arrival date at the destination terminal. We also need to consider the priority of an order. Most customers have an expectation that higher priority orders are less likely to be late than low priority orders. Thus we measure both Proportion of Late Orders and Average Lateness within each priority level and overall.

Lead Time Metric

The Lead Time is defined as the difference between the date/time the customer makes a shipment available at its origin and the mutually agreed due date/time at the destination. The due date is used to establish lateness. However, customers are often concerned with the length of the lead time itself, even if the shipment is officially "on-time". There is a natural tendency for the operations division of the company to advocate longer lead times while the marketing division advocates shorter lead times. Long lead times benefit operations since they provide a cushion against the inevitable problems that crop up during shipping. The situation can be likened to navigating treacherous waters, with numerous "rocks" (system problems) lying below the surface. The "high water level" (long lead time) provides a margin of safety against running into the rocks. An alternative strategy is to "remove the rocks" (identify and eliminate problems). That is where the lead time (and other) metrics may be useful.

To monitor the sufficiency of lead times we can look at the ratios of *Transit Time to Lead Time*. A ratio of 1.0 means that due dates are just being met, values < 1.0 indicate that lead times may be too long, while values > 1.0 indicate that lead times are too short. Since lead times and lateness are directly related it is reasonable to compute the Transit Time to Lead Time ratio by priority level as was suggested for the lateness measures. In this way system process changes that differentiate by priority can be fully evaluated.

Conclusions

Relatively simple performance metrics for dynamic routing applications can be derived from basic system historical data. The proposed metrics are relevant because they are linked to company value; capacity-based metrics are linked to operating costs while time-based metrics are linked to customer service levels. The metrics are appropriate because they are not too great in number, but still sufficient to describe critical operating characteristics. This paper does not propose how to improve a system, but how to begin measuring for improvement. Because dynamic routing applications are particularly complex, they have a special need for performance measurement systems to help identify and eliminate problems.