物流系统分析
Logistics System Analysis

第 12 周　带转运的一到多配送问题（2）——一次转运
One-to-Many Distribution with Transshipments — One Transshipment

葛乾

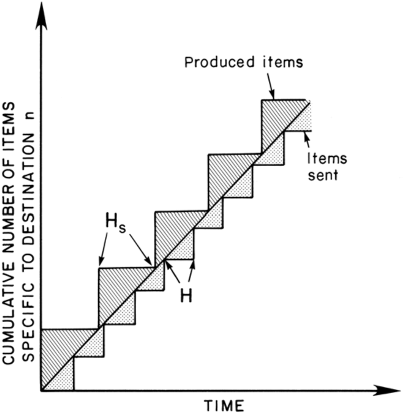
西南交通大学　系统所 & 交通工程系

目 录 |Outline

The Problem

- We will focus first on the problem with only **one transshipment** (finding $I_0(\boldsymbol{x})$). This most common case is also useful as a building block toward multiple transshipment solutions.
- The one transshipment problem is similar to the classical facility sizing and location problem; it is slightly more complicated, however, because in addition to **facility sizes**, **service schedules** need to be determined.

- Consider an imaginary subregion of $\boldsymbol{R}$ that is located r distance units away from the depot and exhibits a constant, stationary demand rate density (λ items per unit time and unit area) and a constant spatial customer density (δ customers per unit area).
- We will find the **optimal dispatching frequency** and **the size of the influence area** I^* in the imaginary subregion, assuming that vehicle routes are constructed as described in 1-to-N distribution systems — the subscript "0" is not used to index "I" because only level-0 influence areas are being considered in this talk.



If no effort is made to coordinate the inbound and outbound (进站与出站) schedules at a terminal, but the inbound and outbound headways (H^i; H^o) are constant, the accumulation of items at the terminal for a specific destination is given by the **vertical separation between step curves** such as those of the figure. The average inventory cost per item is then $(c_i/2)(H^i + H^o)$

Holding costs at the terminal

- The maximum accumulation of items of any type cannot exceed the maximum vertical separation between the two curves. Since the item flow through the terminal is $D' = \lambda I$, the maximum vertical separation is $\lambda I(H^i + H^o)$. Thus, a conservative estimate for the holding costs per item at the terminal (the terminal serves an area of size I), is:

$$(\frac{c_i}{2} + c_r^t)[H^i] + (\frac{c_i}{2} + c_r^t)[H^o] + (c_i + c_r^t)H^t$$

where H^t represents a (fixed) transfer time that an item must spend in the terminal even if H^i and H^o were zero, and c_r^t is the terminal rent cost coefficient (in monetary units per item-time).

Holding costs at the terminal (cont.)

The waiting costs per item at the terminal

$$(\frac{c_i}{2} + c_r^t)[H^i] + (\frac{c_i}{2} + c_r^t)[H^o] + (c_i + c_r^t)H^t$$

are a sum of three separable components:

1. a first term which only depends on H^i and is identical to the term that would have existed if the terminals had been the final destinations;
2. a second term which only depends on H^o and is identical to the term that would exist if the terminal had been a depot producing items at a constant rate
3. and a third term which is a constant penalty

Discourage small terminals

- For more realism we may also want to include a minimum rent to be paid per unit time c_r^o, even if the maximum accumulation is zero. This will discourage the operation of very small terminals. Prorated to the items served in one time unit, the minimum rent is $c_r^o/(\lambda I)$; thus, the third term becomes:

$$(c_i + c_r^t)H^t + \frac{c_r^o}{\lambda I}$$

- This expression only accounts for the holding costs specific to the terminal; i.e. costs added by the transshipments, and not included in the sum of costs of distribution to the terminals and the cost of distribution from the terminals

Handling cost

- In addition, items passing through the terminal must pay a **handling cost** penalty, which will have three terms: the cost of *unloading the vehicle*, the cost of *sorting and transferring the items internally* and the cost of *loading the outbound vehicles*.
- The 1st term is the same that would have to be paid if the terminal was a final destination, and the 3rd term the same as if the terminal was the depot; these two terms will be captured later.
- The 2nd term is terminal-specific. Its magnitude, on a daily basis, should grow roughly linearly with the number of items handled λI; expressed as a cost per item, it should be of the form:

$$c_f^o/(\lambda I) + c_f^t.$$

where c_f^o and c_f^t are handling cost constants that depend on the nature of the items and the terminals.

中转枢纽的总成本

- The total (motion plus holding) cost specific to the terminal is the sum:

$$\text{terminal cost per item } \approx \alpha_5 + \alpha_6 / I$$

  where $\alpha_5 = (c_f^t + c_i H^t + c_r^t H^t)$ and $\alpha_6 = (c_r^o + c_f^o)/\lambda$.
- Note that this expression is independent of H^i and H^o. It captures the costs not included in the sum of the costs of distributing to the terminals (inbound costs z^i), and the costs of delivering from the terminals (outbound costs z^o)

Inbound Costs

- The total logistic cost, in addition to the terminal cost, must include all inbound and outbound costs. These costs already have been studied, as in the 1-to-N systems $z = \alpha_0 + \frac{\alpha_1}{n_s v} + \frac{\alpha_2}{v} + \alpha_3 n_s + \alpha_4$ s.t. $n_s v \leqslant v_{\max}; n_s \geqslant 1$ 包含了处理成本、运输成本和管道库存成本
- The inbound cost would be given by the minimum of total motion cost as applied to a problem where the terminals are the final destinations. Thus, $v_{\max}$ is the capacity of the vehicles used to feed the terminals, and the spatial density of customers δ becomes the density of terminals I^{-1}. 把中转枢纽看做目的地，则 1 到多配送系统中的移动成本公式也可以用到进站成本的计算
- Care must be exercised in solving the equations. For large I, constraint $n_s \geqslant 1$ may be binding. It may be optimal for vehicles to visit only one terminal at a time ($n_s^* = 1$). Other constraints for route length or number of stops may also have to be considered

$\alpha_0 = c_s' + c_i r/s + c_i t_s/2; \alpha_1 = 2rc_d + c_s; \alpha_2 = c_d k\delta^{-1/2} + c_s; \alpha_3 = 1/2c_i(k\delta^{-1/2}/s + t_s); \alpha_4 = c_h/D'$

k 的取值

- In solving the problem we may also want to alter the value of k (the VRP dimensionless constant for the distance added by each stop) to reflect the fact that stops will now be (roughly) on a lattice. 此时顾客大致分布在网格上，也应该响应修改 k 的取值
- This coefficient declines a little, but the change is only on the order of 15%*. When there are more stops per tour than tours (this is highly unlikely when distributing to terminals) the change in k' is also small. 当每个旅程所服务的顾客数多于旅程数时，k 的变动也比较小

*See Problem 5.2

进站成本的性质

- In any case, the minimum inbound cost will be a function of decision variable I only. This function will decrease with I because the more concentrated the demand becomes at fewer terminals ($I \to \infty$) the cheaper it is to serve it. 需求点越集中在一些中转枢纽，服务这些点的成本越低
- Note that the minimum cost per item can depend on parameters r and λ *but not on* δ. It will be denoted: $z^i(\lambda, r, I)$. The cost per unit area and per unit time, λz^i, will share the same properties

出站成本

- The outbound cost per item depends on the density of destinations, but not on the distance from the depot. It can be calculated with the continuous approximation method, as if the terminal were producing items for the customers in its influence area, and averaging the result across the influence area in the usual way.

与进站成本的异同

- Let $z_o(\lambda, r, \delta)$ denote the per-item cost of serving without transshipments a set of customers located r distance units away from a depot (the terminal) when the demand rate density is λ and the destination density is δ .
- This function is also similar to the logistics cost in the 1-to-N distribution system, but it may be somewhat different than for inbound costs because:
  1. customers may be randomly scattered (not on a lattice like the terminals)
  2. vehicles may have smaller capacities
  3. travel speeds may be lower
  4. perhaps all the customers do not need to be visited with each dispatch

出站成本的构成

- According to the continuous approximation approach, the cost per item delivered from the terminal can be approximated by averaging $z_o(\lambda, r, \delta)$ over r, where r is now the distance from points in the influence area to its terminal. We will denote this average, independent of r but a function of I, by a capital "Z" superscripted by "zero" — the level of the influence area — Z^0. Thus:

$$Z^0(\lambda, \delta, I) \cong E_r[z_0(\lambda, \delta, I)]$$

- z_0 increases with r, and that in some cases (e.g. when the vehicles are filled to capacity*) it does so linearly. It is thus reasonable to substitute $E_r[z_0(\lambda, r, \delta)]$ by $z_0(\lambda, E[r], \delta)$, and to approximate $E(r)$ by a simple function of I.

*total combined cost per item$\approx [c_s + 2c_d E(r)]/v_{\max} + c_s' + 2\{c_r[c_s + c_d k E(\delta^{-1/2})]/\bar{D}'\}^{1/2}$ and total motion cost per item$\approx \frac{\alpha_1}{v_{\max}} + \frac{\alpha_2}{v_{\max}} + \alpha_4 v_{\max}$

简化公式

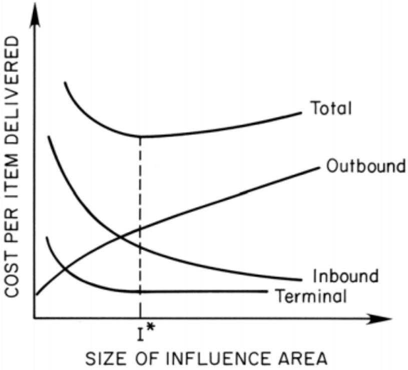
Since influence areas will be drawn to approximate circles and the density of destinations is approximately uniform, we can assume that $E(r)$ is $2/3$ the maximum distance from the terminal, $(I/\pi)^{1/2}$*, and thus:

$$Z^0(\lambda, \delta, I) \cong z_0\left(\lambda, \frac{2}{3}\sqrt{\frac{I}{\pi}}, \delta\right) = z_0\left(\lambda, 0.38I^{1/2}, \delta\right)$$

which increases with I, (linearly with $I^{1/2}$ in some important cases

*圆内任意一点到圆心的距离的期望值是 $2R/3$，此时圆的半径为 $(I/\pi)^{1/2}$

The Design Problem

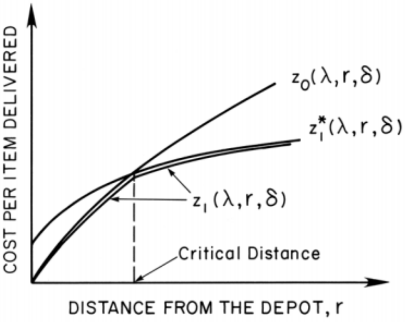


The next step consists in writing a logistic cost function that relates the total cost per item distributed to the decision variables of the problem. In our particular case, the total cost per item distributed is the sum of the terminal, inbound and outbound costs:

$$\underbrace{\alpha_5 + \alpha_6/I}_{\text{terminal}} + \underbrace{z^i(\lambda, r, I)}_{\text{inbound}} + \underbrace{Z^0(\lambda, \delta, I)}_{\text{outbound}}$$

- The value of I that minimizes this expression is the size of the influence area which we would like to use. Values of I larger than the service region size, $|\boldsymbol{R}|$, do not need to be considered. The optimum influence area size, I^*, should usually grow with the distance from the depot but it can also be independent of r, e.g., as occurs with the "cheap item" scenario leading to $\frac{\alpha_1}{v_{\max}} + \frac{\alpha_2}{v_{\max}} + \alpha_4 v_{\max}$.
- The minimum cost obtained with the above expression, denoted $z_1^*(\lambda, r, \delta)$ because one transshipment is used, should be compared to the cost of distribution without transshipments, $z_0(\lambda, r, \delta)$. Only if $z_1^* < z_0$ should transshipments be used. The cost per item with up to one transshipment z_1 is the minimum of z_1^* and z_0: $z_1 = \min\{z_0, z_1^*\}$.

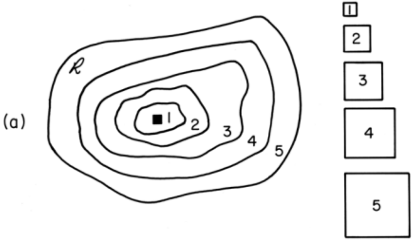
引入中转枢纽后的成本变化



The figure depicts this relationship as a function of *r* for constant λ and δ. As we have indicated, z_0 increases with *r*; z_1^* also increases with *r*, but at a lower rate for large *r*. If the curves don't intersect, then terminals don't have the potential for reducing cost. We have already seen that terminals are beneficial if **there are restrictions to the size of a local delivery vehicle and/or route length limitations**, but in the absence of such limitations transshipments are likely to be unnecessary

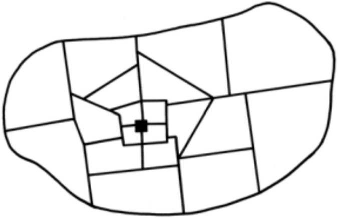
子区域的单位时间期望成本

- The expected total cost per unit time over $\boldsymbol{P}$, any subregion of $\boldsymbol{R}$, can be obtained even before a solution scheme is constructed, by integrating $\lambda z_1(\lambda, r, \delta)$ over $\boldsymbol{P}$. Expressed per unit time, the total cost, again denoted by a capital "Z" , is:
$$Z_T^1(\boldsymbol{P}) \cong \int_{\boldsymbol{P}} \lambda z_1(\lambda, r, \delta) \mathrm{d}\boldsymbol{x},$$
where λ, r ,and δ can be slow varying functions of $\boldsymbol{x}$. The subscript "T" alludes to "total cost per unit time" and the superscript to the maximum number of transshipments allowed.



This figure depicts the loci of points in $\boldsymbol{R}$ for which level-0 influence areas have five different sizes. This could be the result of solving the idealized model for different points in $\boldsymbol{R}$, with different λ, r and δ. These sizes were chosen to increase relatively fast to make the partitioning more difficult. Points in between the curves require intermediate sizes.

(b)



This figure shows a possible partition of $\boldsymbol{R}$ that conforms fairly well with the stated requirements.

- In general, the complete design can be obtained as follows. First carve out "round" influence areas that pack and conform to the calculated sizes $I(\boldsymbol{x})$ as well as possible, as we have just shown. Then locate the terminals near their middle, obeying any local constraints that may exist. Finally, determine the optimal operating strategy within each influence area using the techniques for the 1-to-N distribution systems, separately from the others. 完整的设计步骤：首先，分割出尽量规则的符合计算出的 $I(\boldsymbol{x})$ 大小的影响区域；然后，在满足区域约束的前提下，将中转枢纽放置于形状的中心；最后，基于 1 对多配送系统中所总结的技术，为每个影响区域计算最优的运营策略。

一些注意事项

- Note from the figure that while many points in $\boldsymbol{R}$ do not belong to an influence area of the right size, few have to be enclosed in areas that are off by more than 50% from the target size. Larger discrepancies should be rare in practice. Discrepancies of typical magnitude introduce little error into the resulting cost, $Z^1_T(\boldsymbol{R})$, since the logistic cost function is usually rather flat around its minimum with respect to I

影响区域的大小与成本的关系

$$\underbrace{\alpha_5 + \alpha_6/I}_{\text{terminal}} + \underbrace{z^i(\lambda, r, I)}_{\text{inbound}} + \underbrace{Z^0(\lambda, \delta, I)}_{\text{outbound}}$$

- The solution to problem 3.10 illustrates this fact by examining cost functions of the common form: $\alpha I^a + \beta I^{-b}(a, b \leqslant 1)$. For this kind of expression the chosen value of I can depart from the optimum by as much as 50%, and the resulting cost will still be within a few percent of the optimum. When a and b are smaller than 1 the solution is even more robust than the EOQ expression (the case with $a = b = 1$). We can be reasonably sure as a result that demand points do not have to be enclosed in influence areas of the precise size for a solution to be near-optimal

影响区域的大小与成本的关系 (续)

- For example, if (i) moderately priced goods have to be delivered to fixed retail outlets, (ii) vehicles can make multiple stops, and (iii) no terminal economies of scale exist ($\alpha_6 = 0$), then the cost function consists of a constant, a term proportional to $I^{1/2}$ and a term proportional to $I^{-1/4}$.
- Then, I could be 1.5 times larger or smaller than I^* and cost would only increase by about 1%. Although not quite so robust, the example about to be introduced exhibits a similar behavior.
- Among those problems explored (involving various underlying metrics, deliveries of people and goods, routes with and without multiple stops, deliveries to fixed retail outlets, and individually located customers, etc...), the example corresponds to the set of conditions that makes the cost most sensitive to I.

- Here we consider a region $\boldsymbol{R}$ with constant λ and δ. Line-haul vehicles shuttle between a distribution center and consolidation terminals.
- Neither local nor line-haul vehicles are allowed to make multiple stops because the cost (and delay) of a stop is large compared with that of the moving portion of the trip. This could happen for air transportation of valuable goods.
- In our case, local transportation vehicles pick up their loads at the consolidation terminals and distribute them (non-stop) to destinations scattered over the terminals' influence areas. Local vehicles are assumed to have a small capacity, $v_{\max}$, and to travel full; i.e, the solution to the minimum combined cost is $n_s = 1$ and $v = v_{\max}$.

出站成本

- To make things easier we also assume that the pipeline inventory cost and rent costs can be neglected; i.e., $c_h = c_i$. We then see that the minimum cost is of the form:

$$z_0(\lambda, r, \delta) = \frac{\alpha_1 + \alpha_2}{v_{\max}} + \alpha_4 v_{\max} = \text{constant} + \frac{2rc_d}{v_{\max}} + \frac{c_h \delta}{\lambda} v_{\max}.$$

  The expression is a direct result since the local distance vanishes as $n_s = 1$ and the average customer demand rate is $D' = \lambda/\delta$

出站成本

To simplify the notation, we will ignore the constant term and introduce two constants "a" and "b" ($a = c_h\delta$ and $b = 2c_d/2.7$) so that:

$$z_0(\lambda, r, \delta) = \frac{a v_{\max}}{\lambda} + (\frac{2.7b}{v_{\max}})r.$$

The first term is the stationary holding cost and the second term, the component of transportation cost that is sensitive to distance. For this example, z_0 is independent of δ, and so is the outbound cost function*:

$$Z^0(\lambda, \delta, I) \cong \frac{a v_{\max}}{\lambda} + \frac{b}{v_{\max}} I^{1/2}.$$

* $Z^0(\lambda, \delta, I) = z_o(\lambda, 0.38 I^{1/2}, \delta)$. 出站成本可以通过将 r 替换为 $0.38 I^{1/2}$ 获得。

进站成本

- Inbound transportation to the terminals is assumed to take place on larger vehicles, of capacity $v'_{\max} > v_{\max}$ and cost per mile c'_d, operated at capacity so that the cost $z^1(\lambda, r, I)$ will be (for a demand rate $D' = \lambda I$):

$$z^i(\lambda, r, I) = \frac{\alpha_1 + \alpha_2}{v'_{\max}} + \alpha_4 v'_{\max} = \text{ constant } + \frac{2rc'_d}{v'_{\max}} + \frac{c'_h}{\lambda I} v'_{\max}$$

- Using $a' = c'_h$, $b' = 2c'_d$ again and ignoring the constant we can write:

$$z^i(\lambda, r, I) = \left(\frac{a' v'_{\max}}{\lambda}\right) \frac{1}{I} + \left(\frac{b' r}{v'_{\max}}\right)$$

The first term of this expression represents inventory cost, and the second the cost of overcoming distance. Inventory cost must increase with the number of destinations; as such it is proportional to I^{-1}. Other costs (handling, etc.) that don't depend on I, r, or λ would appear as part of the omitted additive constant.

- Let us assume that terminal costs are proportional to flow ($\alpha_6 = 0$). Then they can be ignored, and the optimal influence area size is the result of a trade-off between the cost of overcoming outbound distance from the terminals ($\frac{b}{v_{\max}} I^{1/2}$) and the stationary inventory cost from inbound distribution ($\left(\frac{a' v'_{\max}}{\lambda}\right) \frac{1}{I}$); the solution is:

$$I^* \cong \left[\frac{2a' v_{\max} v'_{\max}}{b\lambda} \right]^{2/3}.$$

Therefore the one-transsshipment cost is:

$$z_1^* \cong a \frac{v_{\max}}{\lambda} + \frac{b' r}{v'_{\max}} + 1.89 \left(\frac{b}{v_{\max}} \right)^{2/3} \left(\frac{a'}{\lambda} v'_{\max} \right)^{1/3}.$$

最优解

$$l^* \cong \left[\frac{2a' v_{\max} v'_{\max}}{b\lambda} \right]^{2/3}.$$

- The optimal size of the influence area increases with the 2/3 power of the vehicle capacities and decreases with the 2/3 power of the demand density; it does not depend on the distance, r, from the distribution center.
- This is logical, because changing r does not alter the terms traded off.
- These qualitative conclusions, however, are specific to the conditions of the example.

- To see how they would change, assume that the inbound vehicles, still restricted to making one stop, now can carry as many items as desired ($v'_{\max} = \infty$). Then, the loads carried would be the result of an EOQ tradeoff, and instead of $z^i(\lambda, r, I) = \left(\frac{a' v'_{\max}}{\lambda}\right)\frac{1}{I} + \left(\frac{b' r}{v'_{\max}}\right)$ we would have:

$$z^i(\lambda, r, I) = 2\left(\frac{a' b' r}{\lambda I}\right)^{1/2} \text{ and } I^* \cong \left(\frac{2 v_{\max}}{b}\right)\left(\frac{a' b' r}{\lambda}\right)^{1/2}.$$

- The optimal solution is no longer insensitive to r; it grows with r as indicated earlier. It also varies with a smaller power of λ and a larger power of $v_{\max}$. The optimal cost also depends on r and λ, although somewhat differently:

$$z_1^* \cong a\frac{v_{\max}}{\lambda} + 2.83\left(\frac{b}{v_{\max}}\right)^{1/2}\left(\frac{a' b' r}{\lambda}\right)^{1/4}.$$

敏感性分析

- It should be easy to design a system with influence areas close to I^* for most points. Failure to select an I equal to I^* does not result in large increases in cost. For both examples a 30% deviation from I^* results in a cost increase below three percent; for 20% deviations cost increases less than 1%.
- These percentages refer only to the two cost terms that depend on I; otherwise, the percentages would be even smaller. The dependence of cost on I (and its sensitivity to errors in λ and δ) tends to weaken even more when multiple stops are allowed; the conditions of the example are unfavorable.

目录 |Outline

Refinements and Extensions

This section addresses the following subjects which are extensions to the simple model of the previous section:

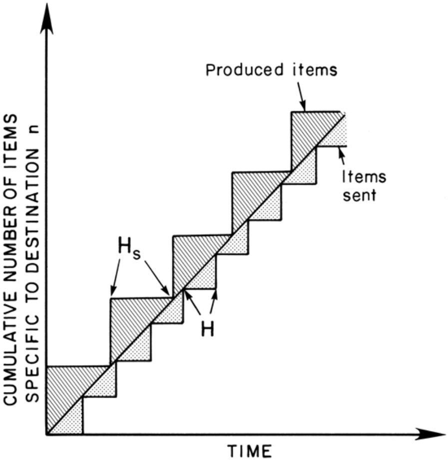
1. synchronization of the inbound and outbound transportation schedules to reduce terminal holding costs;
2. treatment of location/routing constraints cutting across distribution levels;
3. consideration of time-varying demand, with and without uncertainty;
4. development of discriminating strategies when conditions warrant

- The analysis addressed previously was possible because inbound and outbound vehicle routes and schedules from the terminal could be set independently of each other. This decomposition allowed the results of the 1-to-N system without transshipment to be invoked, yielding simple inbound and outbound cost expressions.
- Because some of the extensions explored in this section link the inbound and outbound operations, a conditional decomposition method is used repeatedly. It entails the identification of suitable decision variables, conditional on which the problem decomposes across levels. A similar approach is recommended whenever inbound and outbound operations are coupled.

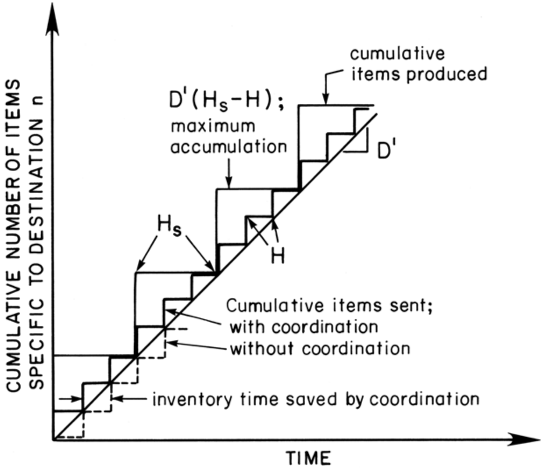
Schedule Coordination

- It was assumed that inbound and outbound operations were independent. Yet, terminal holding costs can be reduced through synchronization. (Our previous lectures showed how synchronization of transportation and production schedules could reduce holding costs; something similar happens here).

Without coordination in production and distribution



With coordination in schedules



The departure curve had been shifted to the left by an amount H.

同步时的枢纽成本

- If we restrict the inbound headway to the terminal H^i to be an integer multiple of the outbound headway H^o, or the other way around, then it is possible to synchronize the arrivals and departures
- This synchronization allows the average time in the terminal to be reduced by the smaller of the two headways: $\min\{H^i; H^0\}$. Then, the maximum accumulation is reduced by $\lambda I \min\{H^i; H^0\}$, and the terminal cost per item becomes:

$$\alpha_5 + \alpha_6/I - \min\{H^i; H^0\}[c_i + c_r],$$

which no longer is independent of H^i and H^o.

Conditional decomposition

- The method we are about to present works even if the outbound headways from the terminal are not equal for all the delivery districts; but we shall assume for the moment that they are.
- The total cost will be expressed as a function of I, H^i and H^o. Conditional on these three variables (instead of only one, I), the total logistic cost per item decomposes in three independent components: (i) an inbound motion cost, z^i_m; (ii) an outbound motion cost, z^o_m; and (iii) the terminal costs plus all holding costs. Thus, the new logistic cost function is expressed as follows:

$$z^i_m(\lambda, r, I, H^i) + z^o_m(\lambda, \delta, I, H^o) + (c_i + c_r)\max(H^o; H^i) + (\alpha_5 + \alpha_6 I^{-1}),$$

where we have assumed that the rent costs only need to be considered for the terminal.

Conditional decomposition

- If rent costs at the origin and the destinations cannot be neglected then a term of the form $c_r(H^i + H^o)$ should be added to the expression.
- In this case, the headway choices may differ from those recommended below. In both cases, however, the choices arise from the minimization of a simple logistic cost function of three variables: I, H^i and H^o.

Conditional decomposition

$$z^i_m(\lambda, r, I, H^i) + z^o_m(\lambda, \delta, I, H^o) + (c_i + c_r)\max(H^o; H^i) + (\alpha_5 + \alpha_6 I^{-1}),$$

- The inbound motion cost term assumes that the inbound routes have been optimized for the given set of terminals and inbound headways, independently of all outbound decisions. This cost can be estimated by the minimum of the first three terms of $\alpha_0 + \frac{\alpha_1}{n_s v} + \frac{\alpha_2}{v} + \alpha_3 n_s$ with respect to n_s for agiven v since the delivery lot size to a terminal v is fixed by $H^i : v = \lambda I H^i$.
- We are pretending here that the terminals are the final destinations. The outbound motion cost for delivery to the customers can be obtained in a similar way, also conditional on the delivery lot size to the customers, $v = (\lambda/\delta)H^o$.

Conditional decomposition

- For most problems the inbound and outbound cost per item, z^i_m and z^o_m, will be decreasing (or non-increasing) functions of H^i and H^o respectively.
- This is logical since with longer headways more goods will have accumulated with every dispatch and they can be distributed more efficiently*.

---

*This assumes that vehicles make many stops and therefore it is only true up to a point. Once H^i (or H^o) is so large that each destination requires a full vehicle load on each visit, increasing H is no longer beneficial.

Conditional decomposition

$$z^i_m(\lambda, r, I, H^i) + z^o_m(\lambda, \delta, I, H^o) + (c_i + c_r)\max(H^o; H^i) + (\alpha_5 + \alpha_6 I^{-1}),$$

- It follows from these properties that the least cost is achieved when $H^i = H^o$; it should be clear that if the smaller of the two headways is not equal to the other, increasing the smaller one until it equals the largest will reduce cost: clearly, the holding cost does not change, and we have already said that the motion cost declines with an increasing H. Thus, we let $H^i = H^o = H$, so that the expression becomes:

$$z^i_m(\lambda, r, I, H) + z^o_m(\lambda, \delta, I, H) + (c_i + c_r)H + (\alpha_5 + \alpha_6 I^{-1}),$$

whose minimum (I^*, H^*) is the desired solution.

Conditional decomposition

- To find it we can hold I constant and minimize the first three terms, the inbound plus outbound costs, with respect to H; the result is of the form:

$$z_m^{i\,*}(\lambda, r, I) + z_m^{o\,*}(\lambda, \delta, I) + (c_i + c_r)H^*(I) + (\alpha_5 + \alpha_6 I^{-1}),$$

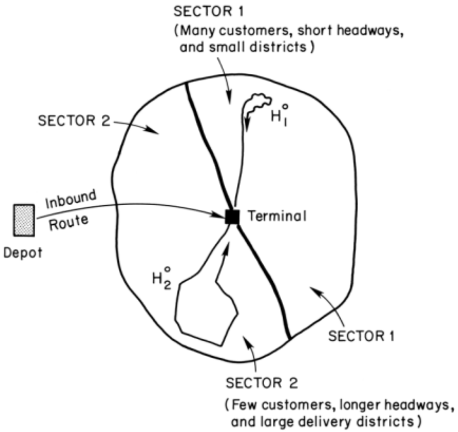
where $H^*(I)$ is the optimum headway for a given influence area size. This expression can now be used for design purposes.

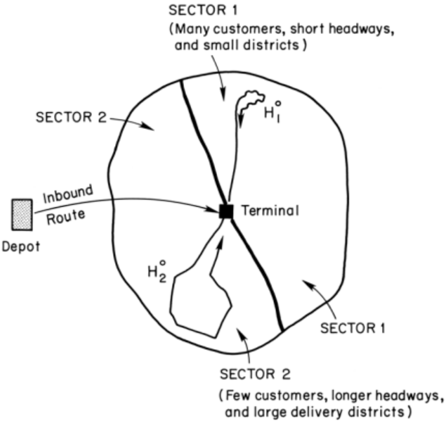
Conditional decomposition

- Schedule synchronization takes some effort and may add to the total cost because the operation of the system is more complex.
- Obviously, it should only be used if the gains outweigh the complexity penalty. The higher the time value of the items the larger the gain and the more desirable synchronization becomes.

Different outbound headways

Suitably modified, the approach we have described can be applied when the influence area is not homogeneous; e.g. if λ and δ change within the influence area





- The vehicle routes to parts of the influence area with different characteristics (or sectors, j) ought to have different numbers of stops and different headways, H_j^o.
- These sectors should be no smaller than the districts covered by one vehicle. Of course, this restriction is irrelevant if vehicles make only one stop.

$$z^i_m(\lambda, r, I, H^i) + z^o_m(\lambda, \delta, I, H^o) + (c_i + c_r)\max(H^o; H^i) + (\alpha_5 + \alpha_6 I^{-1}),$$

- The decision variables are I, H^i and $\{H^o_j\}$. If the outbound headways are multiples or submultiples of H^i, then the expression holds with the following modifications: (i) the outbound motion cost is the demand weighted average of the costs of each sector, $z^o_{m,j}(\lambda, \delta, I, H^o_j)$, and (ii) the waiting cost is the demand weighted average of $(c_i + c_r)\max\{H^i, H^o_j\}$.
- As before the $z^o_{m,j}$ are decreasing functions of H^o_j , and thus **outbound headways should be no smaller than the inbound headway**. Most likely, and even if the demand is quite heterogeneous within I, the solution with $H^o_j = H^i = H$ will be close to optimal, and we can obtain $H^*(I)$ in the same manner as we described.

- A more accurate approach described in the above mentioned references would find first the optimal $\{H_j^o\}$ conditional on I and H^i. This is easy because each H_j^o can be obtained independently of the others as the result of a trade-off between the outbound transport cost in its sector alone, and the corresponding waiting cost: $(c_i + c_r) \max\{H^i, H_j^o\}$. One could then find the optimal H^i and I, either numerically or analytically.
- Even if the H_j^o are not restricted to be integer multiples of H^i, the optimal solution is an integer multiple of H^i.

Further reductions to holding cost

- In the discussion of dispatching strategies it was assumed, even when the schedules were coordinated, that the space needs at the terminal were the sum of the space needs for all the outbound destinations from the terminal.
- This assumption is conservative because it ignores that the need for storage space can be reduced if one staggers the delivery schedules

- If vehicles depart the terminal once every two days for points in the influence area, we implicitly assumed that the maximum accumulation occurred at the same time for all outbound routes.
- But if some of the routes (1/2 of them, say) depart on even days and the other half depart on odd days, then the maximum accumulation will be reduced.
- It will be more difficult, though, to coordinate the deliveries to the terminal with the improved staggered schedule from the terminal. Obviously, the advisability of staggering outbound dispatches will depend on the specific situation. In any case, the methods we have introduced still apply.

Constrained Design

- Here we address two types of design restrictions, which influence the solution approach in different ways: (i) constraints to individual decision variables, and (ii) constraints to sets of variables.
- The first type of restriction arises in connection with tactical problems, which are used as an illustration for the solution approach to (i).

Tactical problems

- If some of the decision variables are fixed, the optimization process is often simplified. This situation is common for short term problems, where the terminal locations are given, but the vehicle schedules and routes need to be determined.

- If so few terminals are available that all should be used, then each influence area will be greater than ideal. One could then easily carve the region into influence areas around each terminal, perhaps allocating every customer to the nearest terminal.
- Ideally, one would like to allocate customers using a marginal cost rule, ensuring that customers near the boundaries are not better off in the neighboring influence area, but this is quite laborious and unlikely to change the size of the influence areas enough to matter for cost calculations. (A marginal allocation is reasonable because outbound distribution costs per unit area $\lambda z_0(\lambda, r, \delta)$ increase with r and as a result the total cost is convex in the zone dimensions).
- Once the partition has been completed, a cost estimate — as well as the optimal headways and routes — can be obtained for each influence area by minimizing an appropriate logistic cost function (consists of terminal, inbound and outboutn) with a fixed I.
- Finally, the detailed solution can be finetuned by computer by testing whether marginal customers near the boundaries should switch terminals.

- If there are so many terminals that we don't know beforehand which ones should be operated, a preliminary step should be taken to make this decision. Based on the given arrangement of terminals, we would define a minimum feasible influence area size, $I_{\min}(\boldsymbol{x})$, as a function of position. We would then obtain for different values of $\boldsymbol{x}$ an ideal influence area size, $I^*(\boldsymbol{x})$, by minimizing the total cost subject to $I(\boldsymbol{x}) \leqslant I_{\min}(\boldsymbol{x})$; the result would then be used to decide which terminals to operate.
- Of course, if there are considerably more terminals than needed this constraint plays no role. Then, the terminals can be selected based on the solution to the strategic problem.

- In the short term we may also have to account for restrictions in the flow through some of the terminals, which essentially impose a limit on the size of their influence areas. If such flow restrictions can be translated into an upper bound restriction to $I(\boldsymbol{x})$, $I(\boldsymbol{x}) \leqslant I_{\max}(\boldsymbol{x})$, then the preliminary step can still be carried out as indicated. The detailed allocation of customers to terminals during fine-tuning, however, must recognize the existence of the flow restrictions. The optimal allocation can be obtained by linear programming.
- If one thinks of terminals as sending a flow equal to their capacity to destinations requesting a flow equal to their demand, and we include an extra destination to which the slack capacity is sent at zero cost, then the flow allocation problem reduces to the Hitchcock transportation problem of linear programming.

- For our problem the costs have a special structure that relates to the geographical distribution of customers; i.e., customers that are close geographically have similar costs from all the terminals.
- As a result, it is not difficult to prove that the set of customers to be served from any terminal should be a well defined region around the terminal. Thus, if the terminal capacities change, only the boundaries to the regions should move.

Multilevel constraints

- The problem just discussed was viewed as a design problem with a constraint on the size of the influence areas. Constraints affecting only inbound, or only outbound, logistic operations are also easy to incorporate; one simply needs to make sure that the expressions for inbound and outbound motion costs, z_m^i and z_m^o, properly reflect the effect of the constraints.
- We have already seen how to develop these expressions under a variety of conditions in the 1-to-N problems without transshipment.

- Constraints that cut across levels are a different matter. This occurs, for example, if there is a maximum time allowed for an item between the origin and the final destination, or a limited transportation budget and/or fleet size for both distribution levels.
- Multilevel constraints like these, can be captured with an extra level of decomposition. In addition to I, H^i and H^o, one should include one or more conditioning variables that will decompose the logistic cost function into independent subcomponents.

- For an example in which total time is limited to an amount tmax , e.g. for the distribution of perishable items, we could use maximum times for the inbound and outbound operation, $t_{(1)}^{\max}$ and $t_{(2)}^{\max}$, and write the equivalent of $z_m^i(\lambda, r, I, H^i) + z_m^o(\lambda, \delta, I, H^o) + (c_i + c_h)\max[H^o; H^i] + (\alpha_5 + \alpha_6 I^{-1})$ also as a function of $t_{(1)}^{\max}$ and $t_{(2)}^{\max}$. The resulting 5-variable logistic cost function can be minimized subject to constraints on $t_{(1)}^{\max}$ and $t_{(2)}^{\max}$ that will ensure $t^{\max}$ will not be exceeded (e.g. $t_{(1)}^{\max} + t_{(2)}^{\max} + \max(H^i, H^o) + H^t \leqslant t_{\max}$).
- A solution to this problem, for a newspaper delivery network, is discussed in Han (1984) and Han and Daganzo (1988). For this problem, in contrast to most other applications, as one moves farther away from the depot both the size of the influence areas and the length of the delivery routes decline

We now discuss stochastic and deterministic demand variations

Stochastic demand

- Here we examine the implications of random (unpredictable) demand at the destinations. Random demand requires extra inventories at the destinations and also at the terminals.
- We will first see that the decision to hold a certain safety stock at a destination can be separated from the routing decisions, conditional on the inbound and outbound terminal headways. A similar decomposition had been already introduced for one-to-one distribution problems

- We will then examine the need for inventories at the terminals (warehouses).
- We assume that the safety stock carried by a customer (destination) depends on the time between deliveries and requests. With deliveries to many customers, however, it is unreasonable to assume that one would dispatch on request, at the precise time when the customer request arrives, lots of specific sizes; it would then be impossible to construct "peddling" delivery tours.
- Rather, one would attempt to coordinate deliveries to all the customers by establishing a dispatching schedule from the terminal at headways H^o — a decision variable — and allowing customers to decide whether or not they desire a delivery on any given dispatch as well as the size of the delivery, $v > D'H^o$.

- For this operating scheme the inventory accumulation at a destination depends on the fixed lead time and on H^o, but it is independent of the transportation routing decisions.
- The solution to problem 5.6 reveals that the holding cost per item includes: a term proportional to v, $c_h v/D'$, that represents the load make-up cost, a safety stock component that increases slightly with H^o but is independent of v, and a new term, $c_h H^o$, that captures the discreteness of the transportation schedule.

- Then, conditional on I, H^i and H^o, the customers' decisions about v are independent of the routing decisions; the problem decomposes. Transportation costs decrease with H^o, while holding costs increase. To enforce rational customer behavior, e.g. discouraging small orders, we will pretend that an amount c_p is charged to the customer for each delivery.
- Whether real or fictitious (if the customers are part of the same firm) this charge can be used to control the customer lot sizes and in the process achieve some overall goal such as maximizing profit, minimizing the sum of costs to the supplier and the customers, etc... We show below how the various logistic cost components can be expressed as functions of I, H^o, H^i and c_p.

- Aside from additive terms independent of v, the motion cost per item paid by a customer will be c_p/v and the holding cost will be $(c_h/D')v$. The optimal lot size chosen by the customer is thus the result of an EOQ tradeoff between those two costs: $v^* = [c_p D'/c_h]^{1/2}$, provided v^* is greater than $D'H^o$; it is $D'H^o$ otherwise. Such a customer would place an order on one out of every $v^*/(D'H^o)$ dispatches, on average.
- If all customers were roughly alike, the reciprocal of this ratio would also represent the fraction of customers requesting service. The effective density of delivery stops δ_e in the region would then be:

$$\delta_e = \delta\{\min[1, D'H^o(c_p D'/c_h)^{-1/2}]\}$$

- The average cost paid by a customer per item delivered is also the result of the EOQ trade-off:

$$\text{customer cost/item} = \begin{cases} 2(\frac{c_p c_h}{D}')^{1/2} + c_h H^o + \text{constant}, & \text{if } v^* \geqslant D' H^o \\ \frac{c_p}{D' H^o} + c_h H^o + c_h H^o + \text{constant}, & \text{otherwise} \end{cases}$$

- Note that this expression is an increasing function of cp and Ho . In other cases (e.g. with different customers or other reorder strategies) the lot size, stop density and cost relationships would be similar; in particular, the effective stop density would still be a non-decreasing/non-increasing function of H^o and c_p.

- This effective density, together with H^o and I determines the supplier's outbound motion costs from the terminal, z_m^o. Inbound motion costs are also determined from H^i and I.
- Both the inbound and outbound motion costs are insensitive to fluctuations in the number of items dispatched from the terminal every H^o because these costs only depend on the average terminal throughput λI

- The supplier's holding costs at the terminal would increase with the fluctuations in the number of items demanded per dispatching headway. Of order $[\lambda I H^o]^{1/2}$, however, the fluctuations should be small compared with the mean $[\lambda I H^o]$ because influence areas contain many customers.
- The added warehousing costs should be small compared with the deterministic holding costs, which in turn are small compared with the motion costs. As a result warehousing costs can be neglected as a first approximation, and terminal holding costs can be approximated by a simple function of H^i, H^o, and I.

- The sum of the inbound, outbound and terminal costs captures the supplier's logistics cost per item as a simple function of c_p, I, H^o and H^i. It is then a simple matter to obtain the c_p, I, H^o and H^i that minimize any desired combination of the supplier and customer costs

- If desired, the analysis can be refined by including warehousing costs at the terminal. These costs are likely to be significant only if the demand is highly variable and unpredictable, and order response time is critical. Although, there is an extensive body of literature on inventory control for a hierarchical system of warehouses very simple models should suffice for our purposes; sensible decisions can be reached without resorting to very detailed models.

- Problem 5.7, an illustration, describes a situation where warehousing costs must be traded-off against transportation costs. In the problem, order response time is so critical that individual items are delivered immediately upon request from warehouses by a very expensive and expeditious transportation mode. Cheap transportation is used to feed the warehouses. The influence area size is the key decision variable for this problem because warehousing costs decrease with I (with larger I the fluctuations in throughput are smaller) but the number of expensive vehiclemiles traveled increases with I (since the distance traveled per item is proportional to $I^{1/2}$)

When the ratio of inventory cost to local distribution cost is sufficiently high warehouses should pool risk by operating in clusters that can share inventory by coordinating their local distribution. This has the potential to reduce cost even more. Two different cooperation methods are possible:

1. periodic redistribution of goods as per a transportation problem of linear programming (TLP);
2. continual re-balancing by serving customers roughly equidistant from two warehouses by the one with the least inventory. The automobile industry essentially uses an extreme version of this method — since automobile purchasers are usually served by the nearest dealer having the desired car.

Hybrids of the two strategies can also be used. Hierarchical schemes, where a lead warehouse holds extra inventory for potential redistribution, can be used as well, but they are less efficient than (1).

- Strategy (1) should be optimized by treating the influence area size, cluster size, safety stock level and re-balancing period as decision variables. The objective function for scheme (1) is easy to write using the expected distance formula for the stochastic TLP*.
- In case (2), the re-balancing period is not an issue. The objective function for (2) is more complicated - since the safety stock level affects the frequency deliveries from the second or third-nearest warehouse in a nontrivial way - but this too can be done

*Daganzo and Smilowitz (2004)

Non-stationary demand

- At the tactical level, i.e. when only vehicle routes and schedules can be changed, non-stationary conditions do not introduce major difficulties. If the average demand rate $\lambda(t, \boldsymbol{x})$, the customer density $\delta(t, \boldsymbol{x})$ and the given set of terminals vary slowly with time the CA approach can be used. First we divide the time line into intervals with quasi constant conditions. Then, we find the optimal customer allocations, vehicle routes and frequencies for each interval independently as if the number of terminals, λ and δ, didn't change with t.
- The chosen solution and the cost estimated for each interval should recognize stochastic effects if they are deemed important. For this to be possible, however, the system should nearly reach a steady state in each interval.

- Each solution is then adopted for its time interval. The average cost over time can be approximated by the weighed average of the costs for the intervals.
- The strategic problem, including the number and location of the terminals can be addressed in a similar way if terminals can be opened, closed and relocated with little cost. If this is not the case the problem is considerably more complicated because small changes in I from one time interval to the next, as would result from the CA approach, might require that most terminals be relocated, and the relocation cost is hard to define.

- If relocation is expensive we would like to relocate few terminals if I changes little. In fact, we would like to change only the absolute minimum number of terminals; i.e. "x" percent of them if I changes by "x" percent.
- This can be achieved if terminals are located approximately on a square lattice, and in every time interval the influence area size is restricted to take a value from the set $\{2^K I_0\}$ for some integer K. (Note that if terminals are located on a square lattice, one can obtain another square lattice with twice the I oriented at 45° with the old, by eliminating $1/2$ of the terminals.)
- With the I restricted in this manner, it is a simple matter to define relocation costs as a function of the change in I from interval to interval. Conditional on I, then, the remaining costs can be obtained as the solution to the tactical problem. Thus, it should be possible to use a dynamic programming formulation to determine the best sequence of I's for a set of consecutive time intervals, where the dynamic programming stage is the time interval and the state is the I for the current interval.

- More sophisticated methods don't seem necessary because long term forecasts as would be needed for strategic analysis are not likely to be reliable. Perhaps we should take a cue from nature in seeing how a logistic structure should adapt to changes in its environment without a forecast for future conditions. As a tree grows taller, new branches overshadow old branches, which may atrophy and die, but the larger older established branches survive.
- Because of the "cost" of growing new branches (opening a terminal), the tree does not totally redesign itself with each change in the environment; rather it preserves a large portion of its structure and builds on it. Moreover, the tree adapts to the future without knowing it — at best it uses the recent past experience as an indication of things to come.

- In light of this, and since the logistic cost is not very sensitive to the specific location and number of terminals, it should be possible to respond to a change in demand by opening and closing only a small fraction of the total number of terminals, and still obtain a configuration that will yield near minimal cost for the new conditions and the anticipated immediate future.

Discriminating Strategies

- So far in these lectures we assumed that customers in the same general area received the same type of service in terms of delivery frequency and type of vehicle route.
- No attempt was made to discriminate across customers based on their individual characteristics. We had seen in the 1-to-N problem that if some customers are much larger than others, or request substantially different items, it might be cost-effective to treat them differently, perhaps even serving them with different transportation systems.
- The same phenomenon can be expected of systems with terminals but the question is now whether or not all customers should be served through the terminals; large ones may be better off with direct service.

Approach

- The conditional decomposition method introduced in this section can also be used to explore this possibility. Given a set of terminals, the tactical problem could be solved by dividing the set of customers into a set that is served through the terminals and another set which is served without a transshipment, organizing the distribution process for the two sets separately, calculating the cost, and then comparing the results for different customer partitions.
- For the decomposition based on customer partitions to be successful one needs to focus only on a few partitions that have a chance of being optimal because the number of arbitrary partitions can be astronomical. Depending on the specific situation at hand, a set of candidate partitions should not be difficult to identify based on physical considerations.

- For example, if as in the discriminative strategies for the 1-to-N problem without transshipment, the items have different values and the customers have different sizes, one may prove that the destinations with the largest dollar demand per unit time should be served direct and the rest through the terminal. This happens because one can reduce the holding cost by swapping customers between the two shipping methods without changing the transport routes and cost.
- This property of the problem allows us to use the fraction of customers that are served without transshipments, f_o, to define the two customer classes and decompose the problem. The fo leading to the least cost is optimal

- The proposed method also applies to passenger transportation problems, although in this case it is somewhat simpler since the partitions are determined by the passengers and not the analyst. A good example of this type of application is Wirasinghe et.al. (1977).
- This reference examines an idealized situation in which passengers traveling to a city from its outlying suburbs have the option to travel either directly by bus (no transshipments) or indirectly by a faster transit system whose stations can be accessed by means of feeder bus lines. Passengers are assumed to use the fastest travel option so that proximity to the transit stations and distance from the city are the main determinants of their choices. Therefore, the resulting partitions are purely geographical.
- The reference is noteworthy because it appears to be the first application of the CA approach to this type of problem.

Items with different densities

- The decomposition approach can also be used when customers differ in other ways. We have assumed so far that an "item" is a given volume of a commodity and that a vehicle can hold a fixed number of items, v_{max}. Although this is a fair description for most freight, for some commodities a vehicle will exceed the roadway axleweight limitations before it is filled.
- To side-step this problem we can define an item as a unit of weight when the commodity being handled is denser than the ideal density for the vehicle (the ideal density is the ratio of the vehicle's weight and volume capacities). All the discussion, theory and methods presented up to this point also hold for dense commodities without any modifications.

- This of course assumes that all the destinations request items of similar density, or at least denser than ideal. If some customers are "light" requesting items lighter than ideal) and others are "heavy" (denser than ideal) it may be advantageous to use an asymmetric treatment to exploit the differences.
- This situation is more likely to arise for collection problems from many suppliers than for distribution problems; e.g. for the collection of the many different parts needed at an automobile assembly plant such as foam for seats, nuts and bolts.

- If a single origin produces many different commodities for a single destination, it is not difficult to see that the number of vehicle loads needed to carry the amounts produced in a given time is minimized if one of the following two conditions is satisfied (Daganzo and Hall, 1985): (i) either all vehicles reach their weight capacity, except possibly the last one which may be partially filled, or (ii) all the vehicles reach their volume capacity. Note that one of the conditions is sure to be satisfied if all the loads are as large as possible while roughly containing the same mixture of items.
- This should be clear since all the loads will then be either below or above the ideal density. As a corollary of this observation we note that if a single destination is fed without transshipments from many small suppliers producing different items, then the symmetric collection strategies described in Chapter 4 also minimize the number of vehicle tours.

- This happens because if both “”heavy” and “light” suppliers are uniformly scattered over the area, then the collecting vehicles will automatically tend to pick up item mixtures with approximately the same density. Without transshipments, thus, there seems to be no incentive to discriminate across customers.
- An exception occurs if light and heavy suppliers tend to form separate clusters, as illustrated in the above reference; in that case it may be advantageous to increase the length of some tours to enhance their load composition, thereby reducing their number. This is in general a complicated problem, whose accurate solution depends on details such as the relative proximity of light and heavy supplier clusters

- Another exception occurs if transshipments are allowed. A case of particular interest occurs if the vehicles are only allowed to make one stop, but collection can take place with a transshipment.
- Suppliers with the lightest and densest commodities have the most to gain from sending their shipments through the terminal since, combined with complementary commodities at the terminal, they can be carried to the destination in ideal density loads requiring fewer vehicle-miles.

- An asymmetric treatment of customers would then be in order. Problem 5.9, based on Daganzo (1988), encourages the reader to develop an optimal asymmetric shipping strategy where the rent for space and the items are so cheap that holding costs can be ignored; only transportation and handling costs need to be considered.
- The solution is obtained by decomposition, conditional on the number of ideal density truckloads sent through the terminal. As part of the solution, the reader needs to determine which suppliers — and how much of their production — should be shipped through the terminal to obtain the conditioning flow through the terminal

Any questions?

Readings

- Daganzo. Logistics System Analysis. Ch.5. Page 171-194