物流系统分析
Logistics System Analysis

模块 4　带转运的一到多配送问题
One-to-Many Distribution with Transshipments

葛乾

西南交通大学　系统所 & 交通工程系

- A transshipment is the act of taking an item out of a vehicle and loading it onto another. Typically, transshipments take place at fixed facilities, i.e., terminals.
- For modeling purposes, these can be viewed as a set of berthing gates connected by an internal sorting, storage and transfer system. The berthing gates accommodate the vehicles while they are being loaded and unloaded; the sorting-storage-transfer system moves the items from one vehicle to another. 从建模的角度，这些中转枢纽可视为一系列的泊位闸门。这些闸门由内部的“分拣-储存和转运系统”相连接。车辆在装卸货时由闸门处理，分拣-储存和转运系统负责将货物在车辆上移动

中转枢纽的职能

- Although many different technologies exist depending on the freight that is being moved, conceptually this makes little difference. (The internal transfer system, for example, can use: carts on rails 轨道运送小车, forklift trucks 叉车, conveyor belts 传送带, idler rollers 传送支持滚轮 or gravity chutes 重力滑槽.) The emphasis at efficient terminals is on moving the freight quickly with little allowances made for long term storage.
- But if there is a need to accommodate seasonal fluctuations in demand, or to hold inventories closer to the points of demand when response time is critical and demand cannot be anticipated, **terminals can also provide a warehousing function**.

We study

- qualitative properties of **near-optimal systems**, which allow the problem to be treated analytically;
- how **system**s where items are transhipped no more than once can be **design**ed, using an uncomplicated scenario as an illustration.
- modifications to the procedure able to capture the following complicating features: **schedule synchronization**, **variable and uncertain demand**, **asymmetric strategies**, as well as **constraints on locations, routes and schedules**.
- multiple transshipment problem.
- how to computerize the design guidelines

目 录 |Outline

1 Introduction

2 Distribution with Transshipments

3 The One Transshipment Problem

4 Multiple Transshipments

5 Automatic Discretization

- After reviewing the reasons for transshipments in one-to-many logistics systems, we will show that finding the **ideal spatial arrangement of terminals is the critical step in designing a system**.
- The rest is easy because, for a given arrangement, there is a well defined set of **item paths**, **vehicle routes**, and **schedules** that (nearly) minimize total cost.

目录 |Outline

2. Distribution with Transshipments
   - The Role of Terminals in 1-to-N Distribution
   - Design Objectives and Possible Simplifications

何时发生转运

- Items are often transshipped when there is an incentive to change transportation modes or vehicle types.
- While **geographical barriers** such as coastlines invariably require a modal change (e.g. at seaports), purely **economical considerations** may also encourage changes in vehicle type.

- We realized that vehicles should be filled to capacity for the distribution of "cheap" freight; i.e., where pipeline inventory cost is negligible compared to the other logistic cost components.
- Because the optimal cost was a decreasing cost of v_{max}, we argued that (if there is a choice) one should use the largest vehicles that the local roads and the destination loading/unloading facilities can accommodate.
- If vehicle size is limited in the immediate vicinity of the customers, trans-shipments at terminals in the general neighborhood of the customers may be attractive, as this could allow larger vehicles to feed the terminals. 假设两个直接相邻的两个顾客之间对车辆容量限制不同，则在二者之间设置中转枢纽比较有利，因为可以用大容量的车辆满足枢纽需求

Effect of a transshipment on vehicle-miles traveled

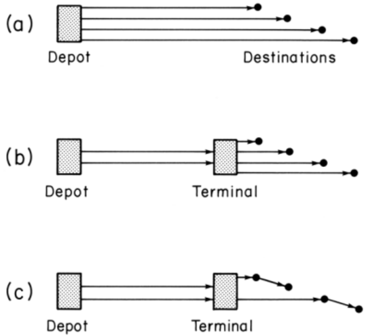
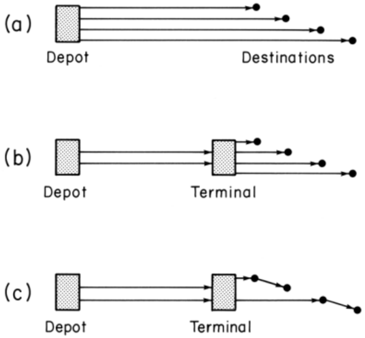


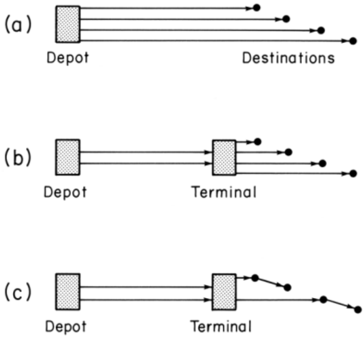
Figure (a) depicts one origin (the depot) and four customers that receive direct service once a day. Each daily trip is represented by one arrow joining the beginning and end of the trip. Let us assume that the pattern of the figure is optimal for the situation at hand, and that the trips are made by **delivery vans**, due to the small access roads leading to the customers.

Effect of a transshipment on vehicle-miles traveled



- If a terminal is introduced, the transportation cost can be reduced without changing the service frequency to the customers (i.e. the waiting cost at the destination).
- The depot-to-terminal roads could accommodate trucks, and terminal-to-destination roads could be served by vans. $\text{Capacity}_{\text{truck}} = 2 \times \text{Capacity}_{\text{van}}$. Only 2 trucks will be dispatched every day. Destinations can still be served daily by vans from the terminal
- The transportation cost can be cut by a factor of ≈ 2 while the holding costs at the destinations keep unchanged

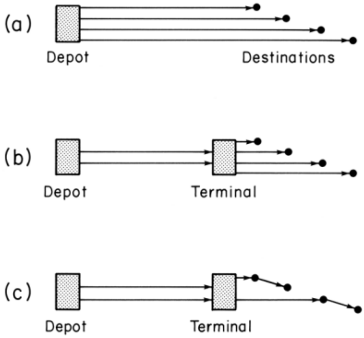
Effect of a transshipment on vehicle-miles traveled



- On the other hand, introducing a terminal may increase holding cost at the origin — items now leave the origin in larger batches — and introduces new handling and holding costs at the terminal.
- Whether the distribution scheme of Figure (b) is advantageous will depend on the magnitude of the transportation cost savings, which grow with **the distance** between the terminal and the depot, and with **the size difference** between vehicles delivering to the terminal and the customers.

- Recall that pipeline inventory considerations, in addition to operating restrictions such as the duration of a work shift, may **restrict delivery route length**; very valuable items should not be delivered on many-stop routes.
- A benefit from transshipments may be derived even if, due to route length limitations, vehicles cannot travel full.

- To illustrate this benefit, imagine that the system in Figure (a) is optimal, and that its vehicles leave the depot only 1/2 full. 假设图 a 中的最优配送批量是 1/2 车容量
- In other words, we are assuming that increasing (or decreasing) the delivery lot size is not desirable because holding costs at the destination would then be too large (or too small). 多之则降低配送频率，提高目的地的保管费用；少之则提高配送频率，提高运输费用；
- Furthermore, although one could presumable reduce costs by using delivery routes with two stops without changing the delivery frequency, we also assume that the loading/unloading operation is so slow that there is no time in a work shift to make more than one stop and return to the depot. 同样，某些情况下可通过让车辆每次服务两个顾客，以在不影响配送频率的前提下降低运输成本。为避免该种情况出现，假设装卸货很慢，以至于无法在一次轮班内停靠多于一个顾客并返回到配送中心。
- Thus, without transshipments, the arrangement can be assumed to be optimal. 在不存在转运时，该排班可设为最优。



- Clearly, the introduction of transshipments as in Figure (b) allows matters to be improved, since the terminal allows the routes to be broken into shorter segments. Although deliveries still take place in half filled vehicles, the terminal is supplied by full vehicles. Further improvement is possible. 图 b 中引入的转运可使得长距离的配送被分割为更短的小段。尽管最终的配送仍然由半满的车辆实现，转运站处的需求可以由满载的车辆供给
- Because the deliveries now start from a place closer to the destinations, there may be time to make more than one stop and reduce even more the daily distance traveled for local delivery, as illustrated in Figure (c). No change in delivery lot sizes results 现在最终端的配送的起点距离终点更近，因此最终的本地运输每次可访问多于一个顾客，更进一步降低每日的运输距离。如图 c 所示，此时终端配送的顾客批量没有发

- In summary, terminals allow us to **decouple the line-haul transportation and local delivery operations**, enabling us to use larger vehicles for linehaul than are used for delivery; they also increase the number of delivery stops that can be made without violating route length limitations.
- We will see in N-to-N problems that terminals can also play a "break-bulk" role.

2 Distribution with Transshipments
- The Role of Terminals in 1-to-N Distribution
- Design Objectives and Possible Simplifications

- The structure of a distribution system is defined by the **number and location of the transshipment points**, the **routes and schedules of the transportation vehicles**, and by the **paths and schedules followed by the items**.
- Usually, the number and location of the transshipment points cannot be changed as readily as routes and schedules. The latter are tactical level variables, and the former strategic variables. Since customers are usually not affected by routing changes, the vehicle routes and item paths can often be viewed as operational level variables, which can be changed even more readily than the delivery schedules.

关于枢纽选址的假设

- For long term (strategic) analyses, decisions at all levels (operational, tactical and strategic) need to be made. For this type of problem we will **develop optimal system configurations assuming that the terminals can be opened, closed and relocated without a penalty**.
- This simplification is not as restrictive as it may seem because, if conditions change slowly with time, locations do not need to be changed often. If $\lambda(t, \boldsymbol{x})$* changes slowly with time, near-optimal terminal locations will be shown also to change slowly with time (this dependence is even more sluggish than the dependence of headways and number of stops on t) 假设需求函数随着时间和区域的变化不大，则近似最优的枢纽选址随着时间的变化也比较慢；这种缓慢变化的幅度，相比配送频率和每次访问的顾客数的变化更小。

*demand per unit time unit area.

选址问题

Because the overall cost is not overly sensitive to the specific locations, one can keep a given set of terminals for a long time before some need to be opened, closed or relocated. In any case, relocation costs are likely to be greatly reduced by current trends in the logistics industry, such as the advent of "third-party logistics" firms that furnish full service terminal/warehousing facilities; 由于最终的成本与特定选址的关系不强，在中转枢纽开放、关闭、移址之前，其存在时间较久。现在物流行业的趋势是移址的费用被大大降低，例如，提供全部中转/仓储服务的第三方物流开始出现。

需求函数的假设

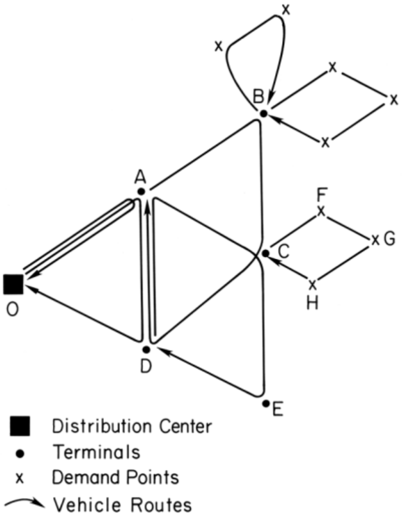
- Unless the changes in $\lambda(t, \boldsymbol{x})$ arise from policy decisions (e.g. expanding the service region over time), the timing of changes to λ may be hard to predict. Without reliable information on them it might be reasonable to design the system as if the changes occurred gradually, using a **smooth forecasted** $\lambda(t, \boldsymbol{x})$ demand density, or else adapting to the **current circumstances as time passes**.
- In either case one would rarely expect the optimal distribution of terminals to change rapidly with time, and it should be possible to design a strategy for opening, closing and relocating terminals that maintains a near-optimal distribution of terminals without large relocating costs.

短期规划

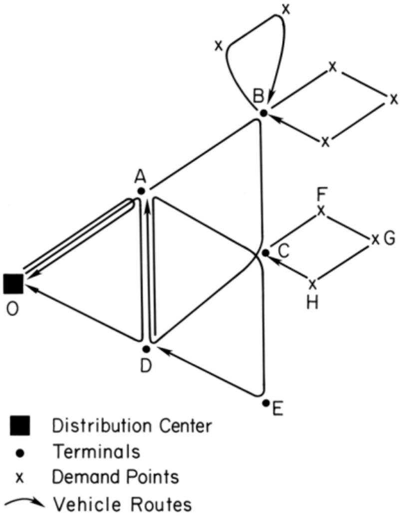
- For the short term one may be interested in adjustments to the tactical and operational decision variables. We may want to determine the best set of **vehicle routes and frequencies for a given set of terminals**; including, of course, the possibility of not using some of the terminals. These (tactical) problems will also be discussed in the talk, although strategic analyses will be its main focus.

- Obviously, the design problem is very complicated if considered with all its details. Our immediate goal, thus, will be to reduce it to a form involving little data and few decision variables, yet capturing the essence of the logistical costs.
- The remainder of this section is devoted to this endeavor; it describes some properties of near-optimal distribution systems with terminals that allow the formulation to be greatly simplified.

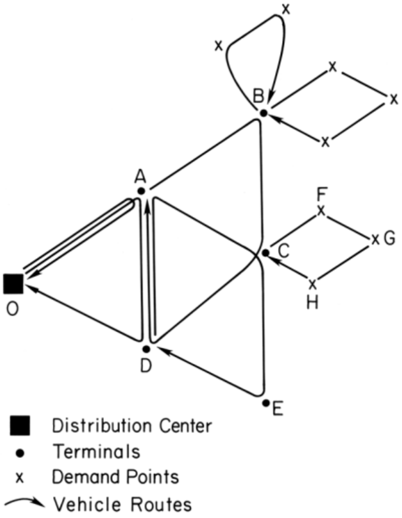
配送系统示意图



- The figure depicts a physical distribution network to carry items from one depot to multiple customers. The network includes terminals (dots on the figure), and multi-stop vehicle routes (looping arrows) that may stop at terminals and customers (***x***'s).



- Because we are only looking at distribution, we shall assume that a vehicle only loads items at the beginning of its route and only distributes them in succeeding stops. This is a reasonable assumption because the within-vehicle sorting complexity and stowage/restowage costs would increase substantially otherwise during a tour. Even collection/distribution systems, for which the savings of interspersing pick-ups and deliveries are obvious, tend to segregate them on individual tours.



- An item that needs to be taken to destination F in figure may use vehicle routes (OAO, ABCDA, and CFGHC), or (OADO, DACED, and CFGHC) to get there. In the first case, it would use path OABCF and in the second case OAD-ACF. If redundant network structures, where some destinations can be reached by more than one path (such as those of the figure), can be shown not to be necessary, we would like to rule them out before starting any analysis. This is done next.

Near-optimality of non-redundant networks

Here we show that, in many situations, networks providing redundant paths are not needed because total cost is concave on flow.

Near-optimality of non-redundant networks

- For the proof we focus on an operational problem, where the terminal locations, vehicle routes and schedules are fixed but one can choose the **item paths and vehicle sizes**. Then, the daily cost of transportation will be the sum of the transportation costs on each route.
- Each one of these route costs should only depend on the size of the vehicle used on the route. Furthermore, the relationship should be **concave** and **increasing** because of the economies of scale in vehicle size. Clearly then, on each route we should choose the smallest vehicles able to carry the load.
- Because **the size of the vehicle must be proportional to the flow of items on the first link of its route**, and **these link flows are linear functions of the item path flows**, transportation cost must be a concave function of the path flows. 车辆的规模应该与它路径中第一个路段的流量成比例，而路段的流量是使用该路段的货物路径流量的线性函数，因此运输成本是路径流量的凹函数。

Near-optimality of non-redundant networks

- Assuming that the path of each item is chosen at the distribution center 0, **independently of the time at which it becomes available for shipment** and of **the characteristics of the item**, we see that the average time that items are waiting outside vehicles on a specific path is not affected by the path selection strategy at 0; the average time is fixed.
- Since travel times are also fixed, total inventory costs must be linear in the path flows. Therefore, the total distribution cost (if rent costs are ignored) must be concave in the path flows*. (We recognize that rent costs are not concave. These costs, however, are typically small compared with transportation costs and should, thus, be unable to reverse the effects of concavity.)

*配送成本由库存成本和运输成本构成，且凹/凸函数加上一个线性函数不改变凹凸性。

两种例外情形

- Before discussing the implications of concavity, it is worth clarifying the two exceptions that were made in the above argument.
- If **the selection of a path** for an item is allowed to **depend on the time it becomes available for shipment** (e.g., passengers using public transportation systems will often choose the first of several lines to depart, if there is a choice) the stationary inventory cost depends on flow; examples can be built where total inventory cost is convex in the path flows. **Even in the (rare) case where dynamic path selection is an option, it is unlikely that one would provide multiple paths to exploit such dynamics**. 配送时间可选

两种例外情形

- The second exception refers to **items of different characteristics**. As shown in the 1-to-N distribution system without transshipment, sometimes it is advantageous to send items of widely varying prices per unit weight on different paths (e.g., expensive items by air freight and cheap goods by land). 异质货物
- In such cases, the pipeline inventory cost is not linear in the path flow; it depends on which items are sent on specific paths. The cost concavity argument does not hold either.

- If all customers are treated alike — asymmetric strategies where this is not the case will be discussed later — and rent costs are not dominant, then **total costs are concave in the flows**; in other words, there are scale economies.
- In this case, as we showed in the 1-to-1 distribution system, only one path should be used to reach each destination.

- These arguments also apply if the destination is an intermediate terminal because intermediate path flows are linear functions of path flows and concavity is preserved. Consequently, path redundancy to either intermediate or final destinations is not needed.
- If follows that each terminal, or final destination point, needs to be served by **only one vehicle route**. Otherwise, the stop could be bypassed by all vehicle routes carrying no flow to it for a reduction in transportation cost.

- This implies that **each destination point should be on only one route from only one terminal**. That is, if we define the level-n influence area of a terminal as the set of points that are served from it with n or less transshipments at succeeding terminals, level 0 influence areas must form a partition of the service area. 一个中转枢纽的第 n 级影响区域为从它出发需要 n 次或者更少转运次数到达的点集。第 0 级影响区域构成其服务区域的一个分割。
- Since each terminal can only be on one vehicle route starting at another terminal, the influence areas at every level must also form a partition. 由于每个中转枢纽仅位于另一个其他中转枢纽的路径上，每级影响区域都构成一个分割。

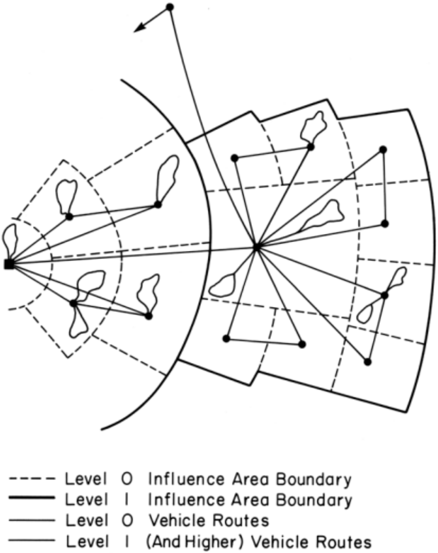


Figure: A possible structure where influence areas are simply connected sets (with no holes). We will reasonably assume from now on that influence areas are simply connected.

Near-optimal operations

- Given the dispatching frequency from every terminal, we describe here which stops should be served from which terminals, and the structure of the vehicle routes based at each terminal.
- To build such routes in a near-optimal way for a given set of stops their length should be minimized. Otherwise, a reduction in length could reduce total cost through decreases in the pipeline inventory cost, and the transportation cost. Thus, it seems logical to construct the routes with a VRP technique

Near-optimal operations

- We also need to decide which stops are to be served from which terminal. It will be assumed that vehicle routes do not stop at both terminal and final destinations.
- This is reasonable (and common practice) because otherwise **sorting and scheduling work would increase substantially in size and complexity**.
- For systems with more than one level of terminals it will be assumed that vehicle routes only stop at one level of terminal. This is also reasonable because substantially different flows pass through terminals at different levels, and it just doesn't seem economical to serve them equally frequently with the same tour.
- Thus, the routes from any level-j terminal* will be assumed to serve all the level-$(j-1)$ terminals in the level-j influence area, or the customers if $j = 0$.

*We say that a terminal is of level$-j$ if its items are transshipped a maximum of j times after passing through the terminal. The terminal serves a level$-j$ influence area.

Near-optimal operations

- As a result, a set of influence areas and terminals (at all levels) defines the stops served from every terminal. Since the VRP solution defines the routes, the overall strategy is defined by a set of **influence areas** and a set of **dispatching frequencies**

Near-optimal operations

- Because level-$(n-1)$ influence areas are usually contained in much bigger level-n influence areas (otherwise terminals would not be cost-effective), the flow through a terminal usually is considerably smaller than the flow through the terminal feeding it.
- This, among other reasons such as restrictions to heavy vehicle travel on local streets, makes it economical to distribute items in loads smaller than those used to feed the terminal. Thus, **each item-mile requires more vehicle-miles during distribution from the terminal than while being fed to the terminal**.
- Consequently, in order to minimize vehicle-miles of travel, terminals should be centrally located within their influence areas. This is true for influence areas of all shapes.

Near-optimal operations

- The same location principle was applied to the one-dimensional terminal location problem. Although the optimal terminal locations obtained in the one-dimensional problem were not exactly in the center of each interval, the displacements were slight.
- Not surprisingly, the CA approximation with centered terminals was found to be quite accurate. A two-dimensional analysis confirms that this simplification leads to negligible errors.

- Unlike VRP zones, influence areas should not be elongated toward the depot; their shape should be selected to be **as close to a circle centered around its terminal as possible**, because this minimizes vehicle-miles.
- Of course, perfect circles cannot be used because they would not fill the space, but non-elongated shapes — "round" we call them — that approximate circles (e.g. squares, hexagons, and triangles) should be appropriate. The specific round shape used does not matter much

- It is thus possible to describe a near optimal system structure by **the sizes of the various level influence areas**, $I_j(\boldsymbol{x})$, as a function of position — together with the **dispatching headways** used at each level.
- As stated earlier, this reduces the very complex design problem to the determination of just a few decision variables.
- Building on this result, the following sections show how to estimate cost and develop a system design for various scenarios.

目 录 |Outline

1 Introduction

2 Distribution with Transshipments

3 The One Transshipment Problem

4 Multiple Transshipments

5 Automatic Discretization

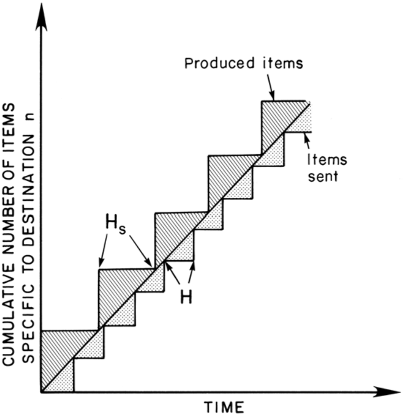
The Problem

- 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.

目录 |Outline

3 The One Transshipment Problem
- Terminal Costs
- Inbound Costs
- Outbound Costs
- The Design Problem
- Example



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)

目 录 |Outline

3 The One Transshipment Problem

- Terminal Costs
- Inbound Costs
- Outbound Costs
- The Design Problem
- Example

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，也应该相应修改 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

目录 |Outline

3 The One Transshipment Problem
- Terminal Costs
- Inbound Costs
- Outbound Costs
- The Design Problem
- Example

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:

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

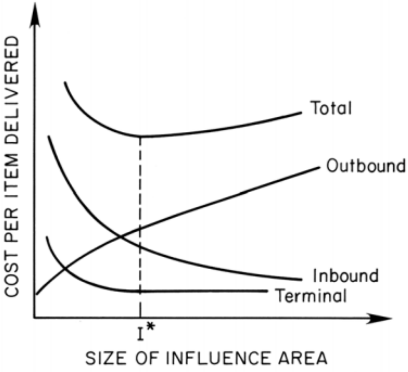
which increases with I, (linearly with $I^{1/2}$ in some important cases

*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 $(I/\pi)^{1/2}$

目 录 |Outline

3. The One Transshipment Problem
   - Terminal Costs
   - Inbound Costs
   - Outbound Costs
   - The Design Problem
   - Example

The Design Problem



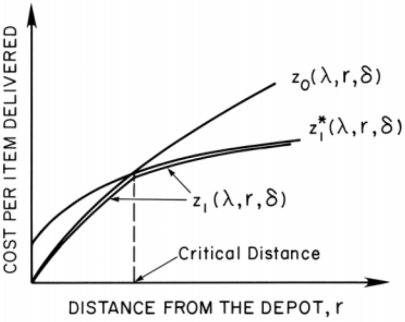
The next step consists in writing a logistic cost function that relates the total cost per item distributed to the decision variables of the problem. In our particular case, the total cost per item distributed is the sum of the terminal, inbound and outbound costs:

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

最优影响区域划分

- 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

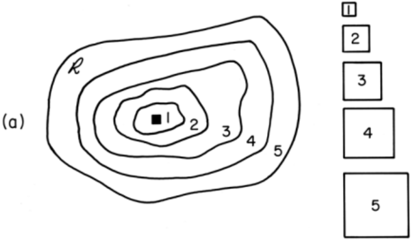
子区域的单位时间期望成本

- 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.

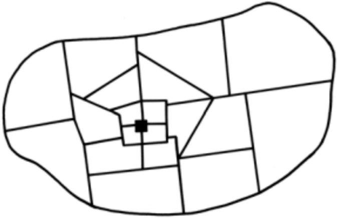
设计方法的应用



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.

设计方法的应用 (续)

(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.

目录 |Outline

3 The One Transshipment Problem

- Terminal Costs
- Inbound Costs
- Outbound Costs
- The Design Problem
- Example

案例

- 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 $D' = \lambda/\delta$

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:

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

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*:

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

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

忽略进站车辆容量约束时的最优解

- 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:

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

- 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

1 Introduction

2 Distribution with Transshipments

3 The One Transshipment Problem

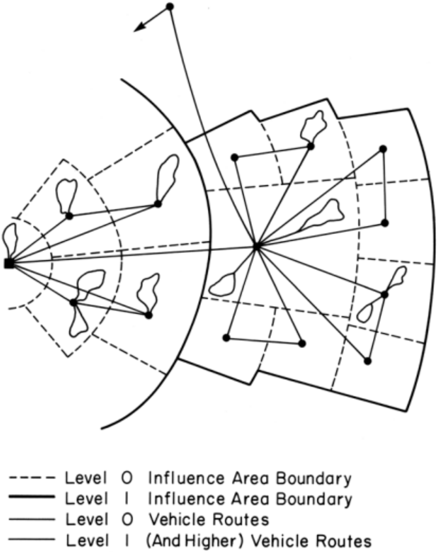
4 Multiple Transshipments

5 Automatic Discretization

目 录 |Outline

4 Multiple Transshipments
- Influence Area
- Example

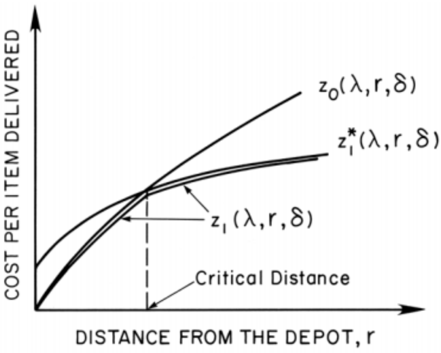
View of the Level-I Influence Area



View of the Level-I Influence Area (cont.)

- The figure depicts a level-1 terminal and its influence area, whose size is now denoted $I_1(\boldsymbol{x})$. Recall that all the customers in a level-1 area are served from the level-1 terminal with at most 1 transshipment, not including the one at the level-1 terminal, and that the level-1 terminals themselves are served without transshipments from the depot.
- This structural organization makes it easy to express, conditional on I_1, the inbound, outbound and terminal costs for a level-1 terminal; the logistic cost function is now:

$$\text{cost/item} = \underbrace{(\alpha_5 + \alpha_6/I_1)}_{\text{terminal cost}} + \underbrace{z^i(\lambda, r, I_1)}_{\text{inbound cost}} + \underbrace{Z^1(\lambda, \delta, I_1)}_{\text{outbound cost}}$$



- The terminal and inbound costs assume the same functional form as in the expression for distribution systems with one terminal, since the cost of delivering and passing through the level-1 terminals does not depend on how the items are treated once they leave them.
- The outbound cost is superscripted by "1" since Z_1 should now represent the average of $z_1(\lambda, r, \delta)$ instead of the (larger) $z_0(\lambda, r, \delta)$

不建议多次转运

- Multiple transshipments are unlikely to be advisable for most physical distribution applications, because **each additional transshipment generates additional handling costs** and **the vehicle economies** (v_{max} vs. v'_{max}) **can be achieved with just one transshipment**.
- In any case, systems that allow multiple transshipments can be designed, using the one-transshipment results as a building block.

This lecture presents a simple recursive technique to this effect, and illustrates it with an example. The technique uses the function $z_1(\lambda, r, \delta)$ to construct a function $z_2(\lambda, r, \delta)$ representing the minimum cost per item with at most two transshipments.

Outbound cost of the level-1 influence area

- We may want to approximate the average cost by the cost of the average:

$$Z^1(\lambda, \delta, I) \cong z_1(\lambda, 0.38 I^{1/2}, \delta)$$

  but the accuracy of this approximation will now have deteriorated because z_1 is more highly non-linear as a function of r than z_0.
- One may instead opt for using the exact definition:

$$Z^1(\lambda, \delta, I) = E_r[z_1(\lambda, \delta, r)]$$

- Either one of these expressions can be used to find the minimum of the logistics cost with respect to I_1. The result should be a function of λ, r, and δ.

- $z_2^*(\lambda, r, \delta)$, as a function of r, should start higher and be flatter than either z_0 or z_1^*. 使用二级转运起始成本可能较大，但是相比无转运或者仅一次转运，曲线更“平”。
- As a result, we may find a second critical distance beyond which two transshipments are needed ($z_2^* < z_1$). For most practical problems, this distance is likely to be large compared with the distance between the depot and the farthest reaches of $\boldsymbol{R}$.

- It is theoretically possible, but practically unnecessary, to iterate this procedure to obtain the optimal size of higher level influence areas. The technique can also be applied if shipments are to be synchronized at the level-1 terminals and also if constraints require a more extensive list of conditioning variables for the decomposition principle to apply.
- In this case one would minimize $z_m^i(\lambda, r, I, H^i) + z_m^o(\lambda, \delta, I, H^o) + (c_r + c_h)\max[H^o; H^i$ $(\alpha_5 + \alpha_6 I^{-1})$ holding H_i constant, and this variable would appear in the expression for Z_1. The new expression would then include the inbound and outbound headways as decision variables, in addition to I_1

- In order to design the system one would carve out the service region into influence areas approximating the ideal size $I_1(\boldsymbol{x})$. Of course, this only needs to be done for the portion of $\boldsymbol{R}$ lying beyond the second critical distance.
- The headways at the level-1 terminals, a byproduct of the optimization, can be used to construct the level-1 feeder routes and schedules.
- Within each level-1 influence area, the system can be designed as previously. An example illustrates the procedure.

目录 |Outline

4. Multiple Transshipments
   - Influence Area
   - Example

- The example that led to z_1^* in the primal one transshipment problem is continued here.
- To simplify the notation we will give some arbitrary values to the constants that appeared: $v_{\max} = b = b' = a/\lambda = a'/\lambda = 1$, and will then eliminate these variables from the notation. We assume that the demand and customer density do not depend on location or time, and use the case with $v'_{\max} = \infty$.

一次转运的计算

Recall the cost expressions:

$$z_0(\lambda, r, \delta) = \frac{av_{\max}}{\lambda} + [\frac{2.7b}{v_{\max}}]r = 1 + 2.7r$$

$$Z^0(\lambda, \delta, l) \cong \frac{av_{\max}}{\lambda} + \frac{b}{v_{\max}} l^{1/2} = 1 + l_0^{1/2}$$

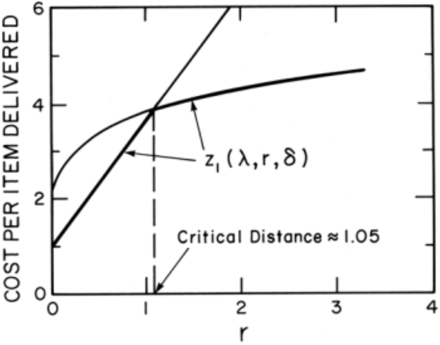
$$z^i(\lambda, r, l) = 2\left(\frac{a'b'r}{\lambda l}\right)^{1/2} = 2(r/l_0)^{1/2}$$

$$l_0^* \cong \left(\frac{2v_{\max}}{b}\right)\left(\frac{a'b'r}{\lambda}\right)^{1/2} = 2r^{1/2}$$

$$z_1^* \cong a\frac{v_{\max}}{\lambda} + 2.83\left(\frac{b}{v_{\max}}\right)^{1/2}\left(\frac{a'b'r}{\lambda}\right)^{1/4} = 1 + 2.8r^{1/4}$$

Putting in the given values, we know

$$z_0 = 1 + 2.7r,$$
$$Z^0 = 1 + l_0^{1/2},$$
$$z^i(r, l_0) = 2(r/l_0)^{1/2},$$
$$l_0^* = 2r^{1/2},$$
$$z_1^* = 1 + 2.8r^{1/4}.$$



Thus

$$z_1(r) = 1 + \min\{2.7r, 2.8r^{1/4}\}$$

Note that when $r \geqslant 1.05$, transshipments become necessary.

二次转运中第一级中转枢纽出站成本的计算

To calculate $Z^1(I_1)$, one should take the expectation of $z_1(r)$ for the r values that arise in an influence area of size $I_1 : r \in [0, (I_1/\pi)^{1/2}]$.

- For small influence areas $(I_1 \leqslant 0.29^{-1})$, $z_1(r) = 1 + 2.7r$ and

$$Z^1(I_1) = 1 + I_1^{1/2}, \text{ if } I_1 \leqslant 0.29^{-1}.$$

  In this case, the level-1 influence area is not large enough to require another transshipment.

- For $I_1 \geqslant 0.29^{-1}$, we find

$$\begin{aligned} Z^1(I_1) =& E_r[z_1(\lambda, r, \delta)] \\ Z^1(I_1) =& 1 + 2.7\int_0^{1.05} \frac{r \times 2\pi r}{(I_1/\pi)^{1/2}} \mathrm{d}r + 2.8\int_{1.05}^{(I_1/\pi)^{1/2}} \frac{r^{1/4} \times 2\pi r}{(I_1/\pi)^{1/2}} \mathrm{d}r \\ =& 1 + 2.17(I_1^{1/8} - I_1^{-1}) \text{ if } I_1 \geqslant 0.29^{-1}. \end{aligned}$$

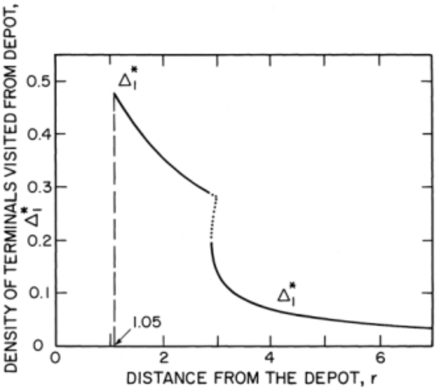
二次转运总成本的计算

Applying the optimal I_1^*, cost per item now becomes (remember that we assumed $\alpha_5 = \alpha_6 = 0$):

$$z_2^* = \begin{cases} 2(r/I_1)^{1/2} + 1 + I_1^{1/2} & \text{if } I_1 < 0.29^{-1} \\ 2(r/I_1)^{1/2} + 1 + 2.17(I_1^{1/8} - I_1^{-1}) & \text{if } I_1 \geqslant 0.29^{-1}. \end{cases}$$

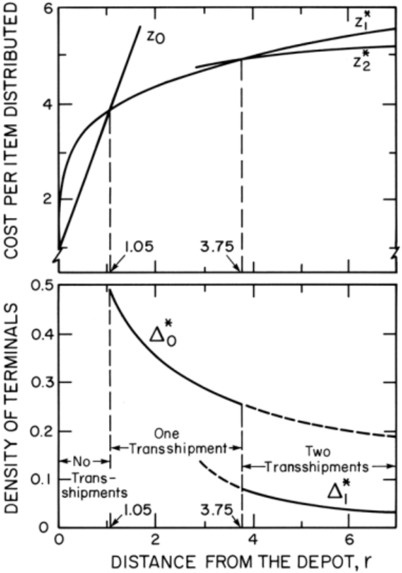
二次转运系统中一级影响区域与 r 的关系

- This expression should now be minimized for all values of r. For this particular problem the task is easy. One can find for every I_1^*, the value of r that makes it optimal — and one can be plotted against the other.
- The right figure plots the reciprocal (关联) of I_1^* as a function of r.



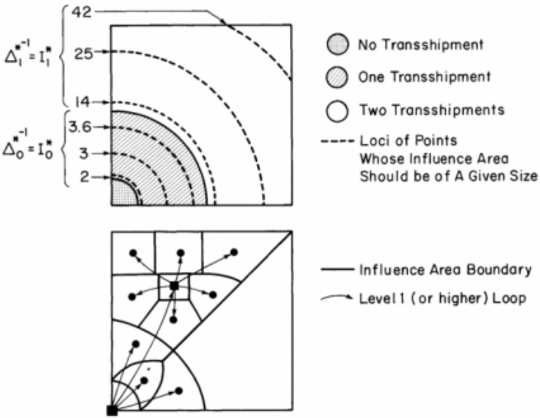
Optimal density of terminals as a function of distance

- The right figure plots the minimum cost as a function of r as well.
- When r reaches 3.75, the cost, z_2^*, equals z_1. For larger values, two terminal shipping is best



Partition of service zone into influence areas

- This figure depicts a possible configuration of influence areas for a square of side that attempts to be true to the density of terminals shown in the previous page
- For this particular problem the task is easy. One can find for every I_1^*, the value of r that makes it optimal — and one can be plotted against the other.



Partition of service zone into influence areas (cont.)

- Unfortunately, the size of the influence areas forces them to include points that would be better served with larger or smaller influence areas. For example, the level-1 influence zones have an area of approximately 20 units, but they include points that optimally would require $I_1 = 13$ to $I_1 = 42$, plus a few corners with even more stringent requirements.
- Inspection of the expression for cost per item, reveals that variations from the optimal I_1 by a factor of 2 only increase the objective function by about 1%.
- This robustness is even more pronounced than that observed for level-0 influence areas because the exponents of the objective function are now closer to zero

$\Rightarrow$ the departures from optimality observed in the previous page should not matter much.

Fine-tuning

- The exact location of the boundaries and terminals can be fine tuned if desired, but since they are fairly round and centered, respectively, the configuration shown should be nearly optimal.
- In fact, even the precise location of the boundary between 2 and 1 transshipment service areas is not particularly crucial. The following section describes an automatic way to fine-tune, or even develop a design.

目 录 |Outline

1 Introduction

2 Distribution with Transshipments

3 The One Transshipment Problem

4 Multiple Transshipments

5 Automatic Discretization

- Before starting, we should mention that the design problem has also been treated in the literature as a pure optimization exercise - without resorting to the CA approach. In the applied mathematics literature the problem is called the "**optimal resource allocation problem**"*
- Pertinent works seek cost-minimizing locations for **point-like service facilities in a space continuum**, among a continuum of customers. Unfortunately, these optimization problems turn out to be "easy" only when cost is defined as a simple function of a distance norm.
- This cost structure, e.g., with the translational symmetry implied by a norm, is unrealistic for typical logistics problems where costs are complicated and almost invariably location-dependent.

*see Okabe et al. (1992) and Du et al. (1999)

选址问题 (cont.)

- More realistic cost scenarios can be analyzed by considering discrete versions of the problem with only a finite number of locations*.
- Problems of this type are usually solved with mixed-integer programming techniques, where the terminal locations and customer allocations are decision variables.
- But unfortunately, existing programming methods can only deal effectively with small problems if they have complicated cost structures.

*An extensive operations research literature explores this line of inquiry; see e.g., Daskin (1995) and Drezner and Hamacher (2002).

CA 的选址方法

- The manual method overcomes these drawbacks. It succeeds, e.g., as in the example of multiple transshipments because it decomposes the problem in two manageable parts.
- We first look for a continuous target $I^*(\boldsymbol{x})$ without paying attention to the discrete locations, and then delegate the difficult but non-crucial task of finding the specific locations to the human mind. As explained in the design problem, the human designer is simply asked to partition the service region into "round" influence areas $\{I_i\}$ of a size consistent with the CA target $I^*(\boldsymbol{x})$, and a set of centrally-located terminals $\{\boldsymbol{x}_i\}$.
- The remainder of this lecture * shows that this second step can also be performed automatically, even for large problems.

*Ouyang, Y., & Daganzo, C.F. (2003). Discretization and Validation of the Continuum Approximation Scheme for Terminal System Design. Transp. Sci., 40, 89-98.

自动分区算法

- Because roundness is important, we first look for a set of nonoverlapping circular disks contained within the service region, of individual sizes as close the ideal $I^*(\boldsymbol{x})$ as possible.
- The number of disks is given by the CA procedure: $N = \int I^*(\boldsymbol{x})^{-1} \mathrm{d}\boldsymbol{x}$. More specifically, if we characterize the disks by their centers $\boldsymbol{x}_i$ and their radii r_i (for $i = 1, 2, \ldots N$), we look for a set of $(\boldsymbol{x}_i, r_i)$ that satisfy: $I^*(\boldsymbol{x}_i) \approx k\pi r_i^2$ for $i = 1, \ldots N$, for a value of k as close to 1 as possible.
- Once this is done, we generate influence areas by allocating each point in the service region to the nearest $\boldsymbol{x}_i$. This is the right thing to do because it guarantees that the influence areas so generated contain one disk a piece. Therefore, they must be "round" - assuming that a solution with $k \approx 1$ has been found.

自动分区算法 (cont.)

- To find a set of disks, we assign some initial values to the $(\boldsymbol{x}_i, r_i)$ and model the disks as if they were physical particles that (i) are repelled when they overlap either with each other or with the boundary, and (ii) change radius as they move over the service region with the recipe: $r_i \approx [I^*(\boldsymbol{x}_i)/k\pi]^{1/2}$. If k is sufficiently large, a discrete-time simulation of this system quickly leads to an equilibrium where all forces vanish and there is no overlap*.
- The simulation is then repeated with a smaller k. A step-wise gradual reduction in k is continued until an equilibrium cannot be found. This will happen before $k = 1$, since circles do not partition Euclidean space. The procedure is then terminated.
- This procedure can quickly find good designs to problems of practical size†.

*This assumes that the service region is "simply connected", in the sense that a disk of proper size can always be slid between any two points in the service region without touching the boundary. No generality is lost by this assumption, because complex areas (e.g., Japan) can usually be partitioned into simply connected components to which the model can be applied separately

†reported in Ouyang and Daganzo (2004)

Algorithm 1: 自动分区算法

$m \leftarrow 1$; 初始化中转枢纽位置 $\boldsymbol{x}_i$ 和影响区域半径 r_i; 预设参数：收敛容许参数 ϵ，步长 μ_m，扰动参数 δ，设置影响区域的面积系数 $k \approxeq 1$，面积系数的变动参数 Δk

while $F_T \neq 0$ 或 $F_B \neq 0$ **do**

    计算每个 disk 的大小：$r_i = \sqrt{\frac{I^*(x_i)}{k\pi}}$

    计算重叠 disk 的斥力 F_T（依赖于 $r(\boldsymbol{x}_i) + r(\boldsymbol{x}_j)$ 与 $||\boldsymbol{x}_i - \boldsymbol{x}_j||$ 之间的关系）和边界的斥力 F_T（依赖于 $r(\boldsymbol{x}_i)$ 与中转枢纽到边界距离的关系）

    if $F_T \neq 0$ 或 $F_B \neq 0$ **then**

        将中转枢纽沿着斥力方向移动 μ_m，同时随机在某个方向增加一个扰动 δ

    end

    if $\mu_m < \epsilon$ **then**

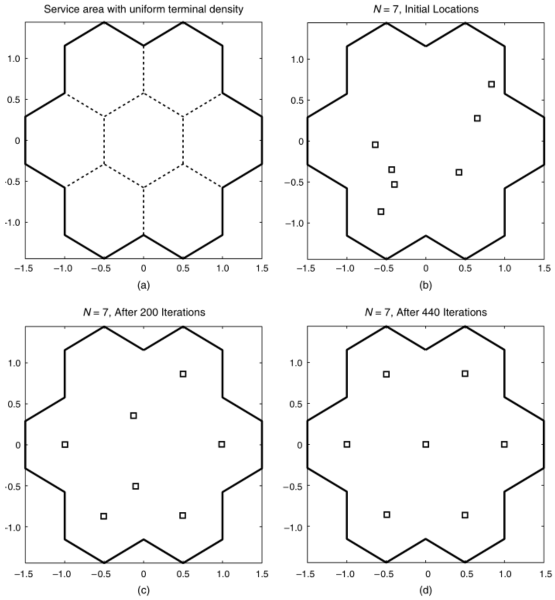
        重置 $m = 1$，$k = k + \Delta k$

    end

    $m = m + 1$

end

基于带权重的 Voronoi 曲面细分方法（weighted–Voronoi tessellation, WVT），划分中转枢纽的影响区域，$i = \arg\min_j \frac{||\boldsymbol{x} - \boldsymbol{x}_j||}{r(\boldsymbol{x}_j)}$。

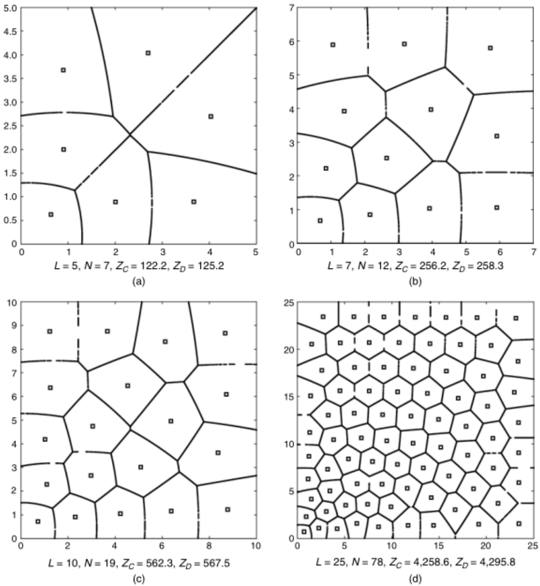


Service area with uniform terminal density
(a)
N = 7, Initial Locations
(b)
N = 7, After 200 Iterations
(c)
N = 7, After 440 Iterations
(d)

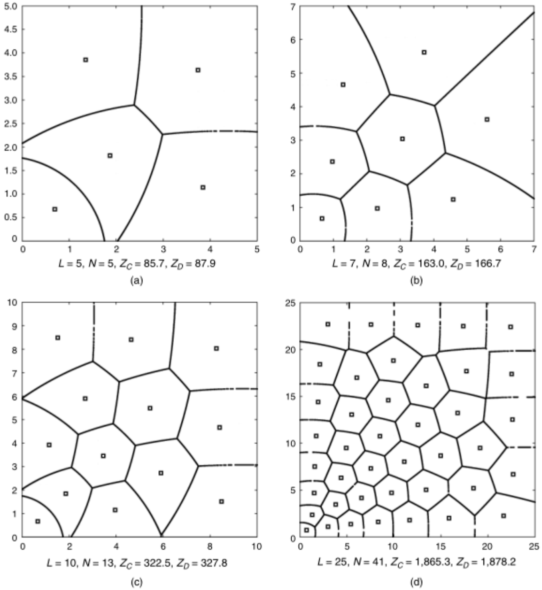
- The figure shows how the method converges in a case where the best design is known. The region is poly-hexagonal with $N = 7$, and the target area size $I^*(\boldsymbol{x})$ is independent of location. The best design is shown in (a), and (b)-(d) show the production of the algorithms of (1) initial locations, (2) location after 200 iterations, (3) equilibrium after 440 iterations respectively.

- The algorithm has also been applied to the example in lecture on the one transshipment problem using $l^* \cong \frac{2v_{\max}}{b} \times (\frac{a'b'r}{\lambda})^{1/2}$ as the target function with $a = b = a' = b' = v_{\max} = 1$; i.e. $l^*(\boldsymbol{x}) = 2[r(\boldsymbol{x})/\lambda(\boldsymbol{x})]^{1/2}$. (Recall that $r(\boldsymbol{x})$ was the Euclidean distance to the depot, and $\lambda(\boldsymbol{x})$ the demand density.)
- Two cases were considered: (a) uniform demand, where $\lambda = 1$ and $l^*(\boldsymbol{x}) = 2r(\boldsymbol{x})^{1/2}$; and (b) declining demand, where $\lambda(\boldsymbol{x}) = r(\boldsymbol{x})^{-1/2}$ and $l^*(\boldsymbol{x}) = 2r(\boldsymbol{x})^{3/4}$. The following two slides show the results for four square regions of sides $L = 5, 7, 10$ and 25 when the customer demand is homogeneous and inhomogeneous, respectively.

Solution for homogeneous customer demand



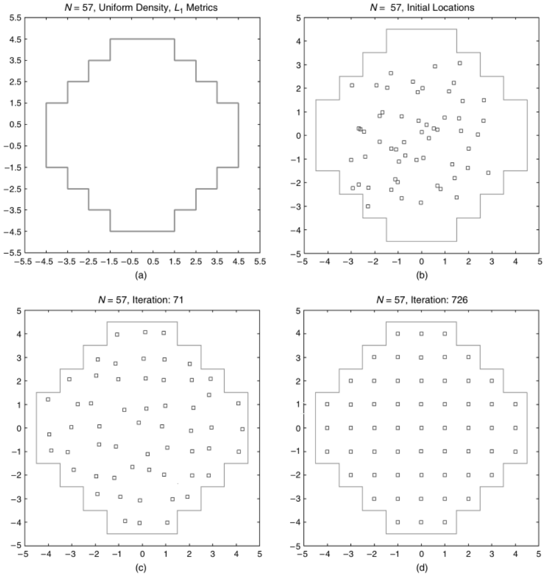
Solution for inhomogeneous customer demand



- In the uniform demand case the difference between the CA cost prediction for the variable costs −the integral $2.83(\frac{b}{v_{\max}})^{1/2}(\frac{a'b'r}{\lambda})^{1/4}$ over the service region — and the variable costs arising from the design is quite small: 2.4% for $L = 5$, 0.8% for $L = 7$, 0.9% for $L = 10$, and 0.9% for $L = 25$.
- In the variable demand case the cost differences are 2.6%, 2.3%, 1.6%, and 0.7% respectively. All these differences are exaggerations because they ignore fixed costs, such as $a\frac{v_{\max}}{\lambda}$, which are large and can be predicted without error by the CA method.
- In all cases, the CA prediction was lower than the actual cost. This is not a coincidence. The CA predictions for our examples should be lower bounds to the optimum solution. Thus, the percentage differences we observed can be interpreted as optimality gaps.

- Note that in both scenarios, and in agreement with theory, the accuracy of the CA formulae and the efficiency of the proposed design method improves with problem size considerably.
- It means both, that the CA formulae describe well the optimum costs of large complex problems, and that the CA discretization algorithm can complement conventional optimization methods when they would have the most difficulty.

Although the discretization procedure was illustrated with Euclidean metrics, it can also be applied to other metrics by deforming the disks during the simulations, and using true distances in the tessellation step. For example, designs for L_1 metrics should use square "disks" with the same repulsive forces as before, and the L_1 distance formula. An example is shown in next slide.


N = 57, Uniform Density, L1 Metrics
(a)
N = 57, Initial Locations
(b)
N = 57, Iteration: 71
(c)
N = 57, Iteration: 726
(d)

Any questions?

Readings

- Daganzo. Logistics System Analysis. Ch.5.