Logistics Systems Design Using a Combination of Exact and Approximative Methods

Tobias Winkelkotte

Deutsche Post Chair of Optimization of Distribution Networks

October 6, 2009

RHEINISCH-
WESTFÄLISCHE
TECHNISCHE
HOCHSCHULE
AACHEN

Outline

Outline

Objective of Research

- We aim to optimize logistics systems where
  - a huge amount of small items
  - is distributed and/or collected
  - regularly (daily?),
  - passing through several stages and facilities.

  Application case: Postal Distribution Network.
- Logistics Systems Design means consideration of:
  - facility location
  - vehicle routing tours
  - tour schedules
- A Standard Network should be implemented:
  - Area is devided into *Regions*
  - Each region is assigned to a facility, where e.g. the sorting takes place
  - Transports from/to the facilities, and between them
- Solution approaches:
  - "Exakt Formulation"
  - "Approximation"

Hybrid Approach

This presentation suggests an model that aims to combine both methods:

- should provide solutions to strategical problem (facility locations)
- But: tactical and operational problems are only considered approximatively

Procedure:

- Given set of potential facilities, from which locations will be chosen
- But: no customer locations assumed, instead the area is devided in *subareas*, each with a different customer density
- Subareas act as "customers" of a Facility Location Problem, so they have to be assigned to the potential facilities
- Cost of allocating several subareas to a facility are estimated using an approximation

Outline

Definitions

Definition: Area

The area the parcels have to be distributed on is denoted by $\mathcal{A}$.
$\mathcal{A}$ is a compact subset of $\mathbb{R}^2$.

Definition: Partition

Let $\mathcal{P}$ a finit set of subsets of $\mathcal{A}$.
$\mathcal{P}$ ist called a *partition* of $\mathcal{A}$, if:

- P is a compact set for all $P \in \mathcal{P}$,
- $\bigcup_{P \in \mathcal{P}} = \mathcal{A}$,
- $\text{int}(P_1) \cap \text{int}(P_2) = \emptyset \ \forall P_1, P_2 \in \mathcal{P}$
- The index set of $\mathcal{P}$ is denoted by $\mathcal{I}$.

Definitions

Definition: Properties of a Subarea

Let $i \in \mathcal{I}$. We define:

- A_i: area size of P_i
- N_i: number of customers in P_i
- λ_i: average demand of a customer in P_i

Definition: Potential Facilities

- $\mathcal{F}$ denotes the set of potential facilities.
- Each potential facility, $j \in \mathcal{F}$, can be opened on different capacity stages: $\mathcal{K}_j$.
- Running a facility, $j \in \mathcal{F}$, on a certain capacity stage, $k \in \mathcal{K}_j$, causes fixed cost of f_j^k.

Decision Variables and Objective Function

Variables

$$y_j^k = \begin{cases} 1, & \text{if facility } j \in \mathcal{F} \text{ is opened with capacity stage } k \in \mathcal{K}_j \\ 0, & \text{otherwise} \end{cases}$$

$$z_{ij} = \begin{cases} 1, & \text{if subarea } i \in \mathcal{I} \text{ is assigned to facility } j \in \mathcal{F} \\ 0, & \text{otherwise} \end{cases}$$

Objective Function

The assignment cost contains cost for distributing the parcels in the region. That means:

- it is non-linear, since it is the approximation of the logistics system,
- it depends an how the subareas are *combined* to a region!

So the assignment cost can – here – only be written as

$$C = C(z)$$

Formulation of the Facility Location Model

So, we formulate the following optimization model, that differs from a CFLP only in the objective function:

$$\min \; C(z) + \sum_{j\in\mathcal{F}} \sum_{k\in\mathcal{K}_j} f_j^k y_j^k \tag{1a}$$

$$\text{s.t.:} \sum_{j\in\mathcal{F}} z_{ij} = 1 \;\forall i \in \mathcal{I} \tag{1b}$$

$$z_{ij} \leq \sum_{k\in\mathcal{K}_j} y_j^k \;\forall i \in \mathcal{I}, j \in \mathcal{F} \tag{1c}$$

$$\sum_{k\in\mathcal{K}_j} y_j^k \leq 1 \;\forall j \in \mathcal{F} \tag{1d}$$

$$\sum_{i\in\mathcal{I}} N_i \lambda_i z_{ij} \leq \sum_{k\in\mathcal{K}_j} K_k y_j^k \;\forall j \in \mathcal{F} \tag{1e}$$

$$z_{ij},\; y_j^k \in \{0,1\} \;\forall i \in \mathcal{I}, k \in \mathcal{K}_j, j \in \mathcal{F} \tag{1f}$$

Evaluation

- This model describes enables the choice of facilities and the assignment of subareas to them.
- But:
  - It does not contain the cost of the long-distance transports, that probably cause very high cost (and can therefore not be neglected)
  - It is possible that the regions are not connected. But this is an important requirement to a lot of logistics networks.
- In the following we will derive linear constraints to consider these two aspects in the optimization program.

Description of the Constraints

- Let c_{j_1,j_2} be the cost of the long-distance transport from sorting center j_1 to j_2 ($j_1, j_2 \in \mathcal{F}$).
- Then the total cost of the long-distance transport can be written as:

$$\text{cost} = \sum_{j_1,j_2 \in \mathcal{F}} c_{j_1,j_2} y_{j_1} y_{j_2},$$

$$\text{or: cost} = \sum_{j_1,j_2 \in \mathcal{F}} c_{j_1,j_2} \min\{y_{j_1}, y_{j_2}\}.$$

- Both formulations are non-linear.
- So we want to find another formulation.

A Linear Formulation of the Constraints

- Introduction of a new set of variables:

$$u_{j_1,j_2} := \begin{cases} 1, & \text{if facilities } j_1 \text{ and } j_2 \text{ are both opened} \\ 0, & \text{otherwise} \end{cases}$$

- The variable having the correct value is assured by:

$$y_{j_1} + y_{j_2} - 1 \leq u_{j_1,j_2},\ j_1, j_2 \in \mathcal{F}.$$

- Constraint with capacity stages:

$$\sum_{k \in \mathcal{K}_{j_1}} y_{j_1}^k + \sum_{k \in \mathcal{K}_{j_2}} y_{j_2}^k - 1 \leq u_{j_1,j_2},\ j_1, j_2 \in \mathcal{F}.$$

- Additional term in the objective function:

$$\sum_{j_1,j_2 \in \mathcal{F}} c_{j_1,j_2} u_{j_1,j_2}.$$

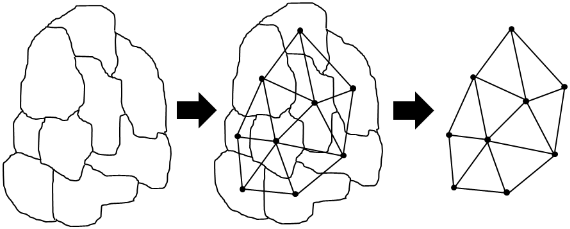
Connectivity

Definition: Connectivity Graph

Let $G_\mathcal{P} := (V_\mathcal{P}, E_\mathcal{P})$ be an undirected unweighted graph with the following properties:

- Each node $v \in V_\mathcal{P}$ is assigned to one and only one $P \in \mathcal{P}$.
- $e := (v_{i_1}, v_{i_2}) \in E$ if and only if P_{i_1} and P_{i_2} border to each other.

Then, $G_\mathcal{P}$ is called the *Connectivity Graph* of $\mathcal{P}$.
$G_\mathcal{P}$ is represented by its adjacency matrix, $A_\mathcal{P} = (a_{ij})$.



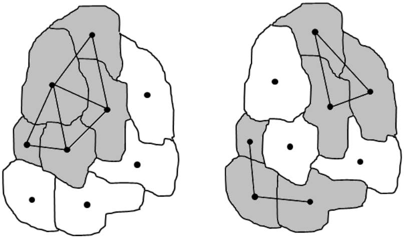
Connectivity and Regions

Definition: Region

Let $R \subseteq \mathcal{P}$ and $G_R := (V_R, E_R)$ a subgraph of $G_{\mathcal{P}}$, where:

- $V_R := \{i \in V_{\mathcal{P}} : P_i \in R\}$
- $E_R := \{(i,j) \in E_{\mathcal{P}} : i,j \in V_R\}$

R will be called *connected*, if G_R is a connected graph.
If R is connected it will be called a *Region*.
$\mathcal{R} := \{R : R \text{ is a region}\}$.



Connectivity Constraints

- Connectivity of regions has to be stipulated by additional constraints in the linear program
- One possibility is:

$$\sum_{\substack{i \in S_1 \cup S_2, \\ j \in \overline{S_1 \cup S_2}}} a_{ij} z_{jk} \geq 1 + \sum_{i \in S1 \cup S_2} z_{ik} - |S_1 \cup S_2| \tag{2}$$

$$\forall S_1, S_2 \subsetneq V_{\mathcal{P}} \text{ where: } S_1 \cap S_2 = \emptyset, a_{ij} = 0 \text{ for } i \in S_1, j \in S_2 \ \forall k \in \mathcal{F}$$

(no a proof here)

- This set of constraints is exponentially big.

Outline

Idea

It is hard to solve the problem, because of:

- exponential number of constraints:
  - $\longrightarrow$ results from stipulating connectivity for all regions simultaneously
- non-linear objective function:
  - $\longrightarrow$ arises from non-linear logistics cost functions
  - $\longrightarrow$ makes it impossible to use linear programming solvers
- integrality

So the idea is:

- to create regions in a subproblem, where only *one* region has to be considered at the same time,
- and to find the best combinations of them in a *mixed integer linear* master program.

$\longrightarrow$ Decomposition of the original program into a master program an a subproblem.

Decomposition

- One subproblem for each potential facility
- Decisions to make for each potential facility:
  - opened or closed?
  - if opened: which capacity stage?
  - and assignment of which region?
- A region, R, allocated to a potential facility, j, will be represented by the binary vector, $x^{R,j}$, where:

$$x_i^{R,j} = \begin{cases} 1, & \text{if subarea } i \text{ belongs to region } R, \\ & \text{that is allocated to potential facility } j \\ 0, & \text{otherwise} \end{cases}$$

- The subproblem will also evaluate its cost: $C(x^{R,j})$

Master Program

- The master program determines the best combination of regions, such that $\mathcal{A}$ is covered completely and non-overlapping.
- Connectivity is now a *property* of the regions, so the connectivity constraints are not required here.

$$\min \sum_{R \in \mathcal{R}} \sum_{j \in \mathcal{F}} C(x^{R,j}) \lambda^{R,j} + \sum_{j_1, j_2 \in \mathcal{F}} c_{j_1,j_2} u_{j_1,j_2} \tag{3a}$$

$$\text{s.t.:} \sum_{j \in \mathcal{F}} \sum_{R \in \mathcal{R}} x_i^{R,j} \lambda^{R,j} = 1 \ \forall i \in \mathcal{I} \tag{3b}$$

$$\lambda^{R,j_1} + \lambda^{R,j_2} - 1 \leq u_{j_1,j_2} \ \forall j_1, j_2 \in \mathcal{F} \tag{3c}$$

$$\sum_{R \in \mathcal{R}} \lambda^{R,j} \leq 1 \ \forall j \in \mathcal{F} \tag{3d}$$

$$\lambda^{R,j} \in \{0, 1\}, u_{j_1,j_2} \geq 0 \ \forall R \in \mathcal{R}, \forall j, j_1, j_2 \in \mathcal{F} \tag{3e}$$

Developement of new Regions

- There is one subproblem for each potential facility, j.
- It searches for the column with least negative reduced cost.
- These are given by the dual prices of the master problem, π_i: Each subarea, i, is associated with (positive or negative) cost, π_i.
- So the program for facility j can be described as[1]:

$$\min\ C(x^{R,j}) - \sum_{i \in \mathcal{I}} \pi_i x_i^{R,j} \tag{4a}$$

$$\text{s.t.: } x^{R,j} \text{ represents s a feasible region} \tag{4b}$$

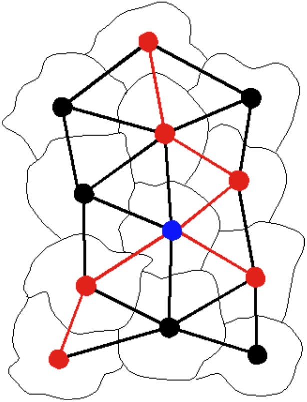
- Objective function is non-linear

  $\longrightarrow$ replace (4) by a surrogate program, evaluate result with original cost function.

[1] some constants neglected

Characterization of the Subproblem

- Feasibility of regions means: Connectivity.
- That means: Each subarea has to be obtainable from the facility, only passing other subareas of the same region, R.
  $\rightarrow$ Nodes of R have to be connected (at least) by a *Spanning Tree*
- For a facility, j, we search for a sub-graph, T_j, of $G_{\mathcal{P}}$, that fulfills:
  - T_j contains the node of the subarea where j is located,
  - Total demand of $\{P_i \in \mathcal{P} : i \in T_j\}$ does not exceed the capacity of j,
  - T_j is a (at least) Spanning Tree of selected nodes,
  - T_j is the cost-minimal graph with these properties



Profit Maximizing Spanning Tree Problem

- The problem to be solved is not known from literature.
- So we formulate a new variant of Spanning Tree Problems.

Definition of the Underlying Graph

Let $G = (V, E)$ be a connected directed graph and

- $w : V \to \mathbb{R}_+$, $w_i := w(i)$, $i \in V$ a *demand function* on the nodes
- $c : E \to \mathbb{R}$, $c_{ij} := c((i, j))$, $(i, j) \in E$ a *profit function* on the edges
- C a *capacity*

Node $v_0 \in V$ is called the "origin" of G.

Variables

$$x_i \in \{0, 1\} : \text{node } v_i \in V \text{ is selected } (1) \ / \text{ not selected } (0)$$

$$y_{ij} \in \{0, 1\} : \text{edge } (v_i, v_j) \in E \text{ is selected } (1) \ / \text{ not selected } (0)$$

Formulation of the PMSTP

The general case of the PMSTP can be formulated as follows:

$$\max \sum_{i,j \in V} c_{ij} y_{ij} \tag{5a}$$

$$\text{s.t.: } x_0 = 1 \tag{5b}$$

$$\sum_{i \in V} w_i x_i \leq C \tag{5c}$$

$$\sum_{i \in V} y_{ij} = x_j \ \forall j \in V \setminus \{0\} \tag{5d}$$

$$y_{ij} \leq x_i \ \forall i, j \in V \tag{5e}$$

$$\sum_{i,j \in S} y_{ij} \leq |S| - 1 \ \forall S \subseteq V \setminus \{0\} \tag{5f}$$

$$x_i, y_{ij} \in \{0, 1\} \ \forall i, j \in V \tag{5g}$$

Relevant Special Case

- Special case: *Edges pointing to the same node have the same profits:*

$$c_j := [c_{ij}].$$

- Now the profits the profits can be moved to the nodes, and the problem is reduced to the selection of nodes:

$$\max \sum_{i \in V} c_i x_i \tag{6a}$$

$$\text{s.t.: } x_0 = 1 \tag{6b}$$

$$\sum_{i \in V} w_i x_i \leq C \tag{6c}$$

$$\sum_{i \in V} y_{ij} = x_j \ \forall j \in V \setminus \{0\} \tag{6d}$$

$$y_{ij} \leq x_i \ \forall i, j \in V \tag{6e}$$

$$x_i, y_{ij} \in \{0, 1\} \ \forall i, j \in V \tag{6f}$$

Solution of the PMSTP-Formulation of the Subproblem

- Choice of edges does not influence the objective function value
  $\longrightarrow$ Model does not contain constraints eliminating cycles
- Cost coefficients are given by dual prices of master program:

$$c_i := -\pi_i \; (\in \mathbb{R}!).$$

- Solution procedure:
  - Relaxation by neglecting constraints (6d) and (6e),
    Result is a knapsack problem, that can be solved in short time.
  - Computation of optimal solution with branch and bound
  - In case of several capacity stages:
    Computation of optimal solution for each of these stages, and selection of the one with maximal profit

Outline

Solution Procedure

1. Find initial regions using a heuristic
2. Solve master program, evaluate dual prices
3. Solve subproblem (PMSTP) for each facility
4. Evaluate costfunction of found region
5. Add column representing the region to master program
6. Resolve master program

Summary

- Formulation of a model to optimize complex logistics systems
- Model is generic an can be used different logistics system
- The only thing to do is to find the correct logistics cost function of the current system.
- A logistics cost function is also derived for the application.
- Solution procedure of the subproblem is implemented
- For solving the master problem we use the SCIP framework from Zuse Institute Berlin.

ANY QUESTIONS?