Deutsche Post World Net
MAIL EXPRESS LOGISTICS FINANCE

RWTH RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

Integration of Knowledge and Analytical Model Analysis in the Area of Facility Location in Postal Logistics

Hans – Jürgen Sebastian

Deutsche Post Chair of Optimization of Distribution Networks

RWTH Aachen University

21st Workshop on Methodologies and Tools for Complex System Modeling and Integrated Policy Assessment

IIASA, Laxenburg, Austria, August 27 – 29, 2007

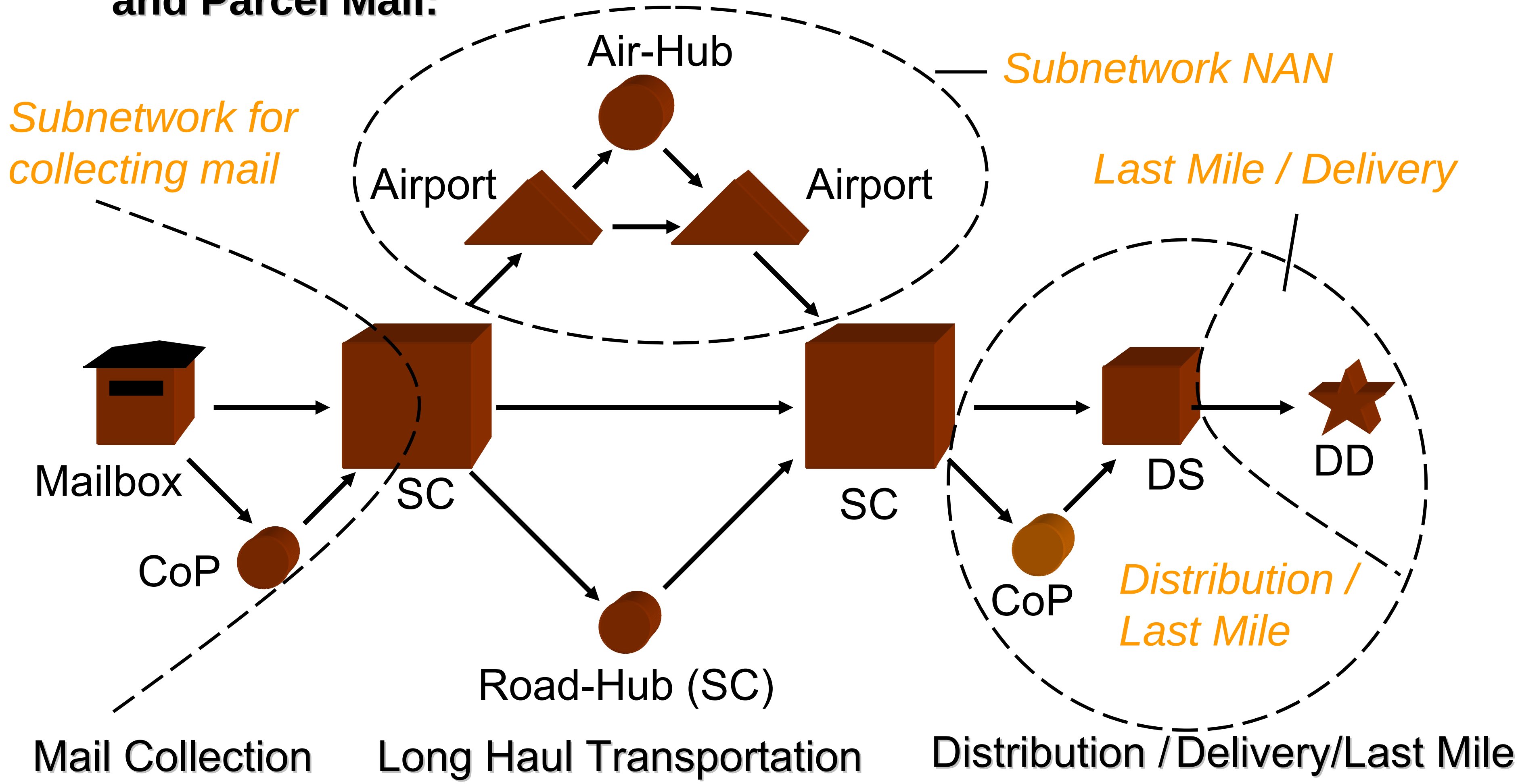
Integration of Knowledge and Analytical Model Analysis in the Area of Facility Location in Postal Logistics

Content

1. Modeling and Model-Based Problem Solving at the Deutsche Post Chair of Optimization of Distribution Networks
2. Selected Analytical Models for Facility Location and Location Routing Problems: Knowledge about Analytical Models and their Solution
3. Combining the Domain Knowledge in the Facility Location Domain with the Solution of Analytical Models: a Case from Postal Logistics

The area of Postal Logistics

Components / Subnetworks of a **Distribution Network for Letter and Parcel Mail:**



Distribution Network for Letter Mail of the DPWN in Germany

40 Million Final Destinations

3 Million Business Customers

72 Million Letters (throughout Germany each working day)

108 000 Mail Boxes

82 Sorting Centers + International Postal Center in Frankfurt

3 300 Delivery Stations (DS) + 12 000 sales offices

Last Mile:	61 000 Delivery Districts:	on foot	: 10 500
	81 000 Postmen	by bycicle	: 25 500
		by car	: 25 000

Important Classes of Optimization Problems in the Field of Distribution Network Planning for Letter and Parcel Mail

Network-Level	Planning Phases: *Strategic*	*Tactical*	*Operational*
Overall Network	Facility Location/Allocation Location Routing	Transportation Network Design Vehicle Routing	Operations Center
Subnetworks:	Decomposition / Integration of Networks and Services (exact or heuristic)		
- Collection	Location / Allocation Location Routing	Vehicle Routing and Scheduling	On-line Optimization for Transporation Subnetworks
- LHT • Hubs • Direct Links • NAN • Airportfeeder	Location / Allocation (Hub Selection) (Airport Selection)	Service Network Design Multi-Depot Vehicle Routing and Scheduling	
- Distribution	Location / Allocation (DS) Location Routing	Vehicle Routing and Scheduling	
- Delivery/Last Mile	Capacitated Arc Routing, Postmen Problems		

Sorting Centers **Simulation of Transportation Systems Layout**

Four classes of optimization problems

PC 1: Facility Location Problems

The Problem is to find facilities, their numbers, types, sizes and locations in order to minimize overall costs.

PC 2: Location Routing Problems

The problem is to determine locations and routes simultaneously.

PC 3: SND – Service Network Design Problems

There are given the facilities by number, size, type, location etc. The problem is to find the Service Network (SND), which minimizes overall costs.

PC 4: Vehicle Routing and Scheduling Problems

In each class, the service quality requirements are used as constraints.

Main characteristics of models in the Postal Logistics and SCM areas

- Problem instances are „large to very large scale“
- Most of the optimization problems in the field are NP-hard (it‘s important to know the exceptions)
- Special structures are present (patterns of constraints) in particular application scenarios

→ Exact methods are successful in special cases and smaller instances only

→ Commercial solvers can not deal with the very large scaled real world applications in a reasonable way

⇒ Heuristics and Metaheuristics complemented by sharp lower bounds (minimization) are needed to solve the practical problems in reasonable time and with reasonable accuracy.

Projects in collaboration with Deutsche Post World Net (DPWN)

- Sorting center location optimization
- Delivery station location optimization project (TOPAS)
- Optimization of the Night-Airmail Network for Letter Mail in Germany
- Delivery station optimization (capacitated arc-routing). Optimal cutting and sequencing!
- Optimal routing within the subnetworks for collecting and distribution
- Optimization of Mail-box Locations
- DHL freight transportation problem of swap body containers
- Distribution Network Design for electronic consumer goods and for fashion articles in Eastern Europe

- Project work is completely funded by the DPWN
- IT-Tools as e.g.
  - database management systems (ORACLE)
  - optimization solvers (ILOG)
  - GIS

  are used

Modeling tools such as e.g. AIMMS, OPL-Studio, AMPL, GAMMS are not used

The reasons:

- The existing modeling environments are not needed to do the implementation work!
  (use of C++, Java, DBMS, Opt-solver, GIS to build the applications)

- They do not support the core activity of the modeling process, which is creative rather than routine or technical activity.
  Modeling is a combination of art and science.

→ A real challenge would be to develop a Modeling Environment which supports the modeling process by making the state of the art expert knowledge of the considered domain available.

Questions arising during the elaboration of a model

For example, there are many important decisions a person who elaborates a model has to make, depending on the answers to the following questions:

- Why the model is needed and which are the reasons to use it? (to describe the problem quantitatively and to classify it, to check feasibility of solutions which have been generated without the model, to prove properties of optimal solutions, to prove properties of algorithmic approaches, to generate solutions model-based)
- What type of model fits best the respective problem description? (optimization vs. simulation, simultan-model (OR-model), algebraic model, differential equations, static or dynamic model, dealing with uncertainty within the model)
- How the model can be changed over time (manipulated) and who is responsible for doing this?
- Which kind of solvers are available for different type of models? (LP, MIP, NLP, heuristics and meta-heuristics, commercial or public domain, available on the web)

RWTH RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

2. Selected Analytical Models for Facility Location and Location Routing Problems: Knowledge about Analytical Models and their Solution

Big variety of basic scenarios, e.g.

- one stage or multistage commodity flow in the system, location related decisions on one or several hierarchical stages
- multiple- or single sourcing (a customer is served by several or by exactly one of the facilities (called shortly „depots"))

decisions concerning the network topology

- capacitated or uncapacitated problems (capacity constraints at the depot locations)
- static or dynamic problems (static model based on a given demand scenario / time-varying demand and/or location decisions are time-dependent)
- transportation costs approximated by distance-dependent-costs or by vehicle-routing-based-costs
- uncertainty modeling of uncertain model parameters.

Model Classes

– Models in the Plane *(Greenfield Approaches)*
(Weber Problems, Center Problems)

– Network Models *(Graph-based Models)*
(p-Median-, p-Center-Problems)

– Mixed-Integer Models *(Discrete Models)*

→ e.g. One-stage, static models with approximated transportation costs

- Uncapacitated Facility Location Problems (UFLP)
- Maximum Covering Location Problem (MCLP)
- Capacitated Facility Location Problem (CFLP)
- Capacitated Facility Location Problem with Single Sourcing (CFLPSS)

→ ....

Combining the Domain Knowledge in the Facility Location Domain with the Solution of Analytical Models: a Case from Postal Logistics

The Case:

Slide 3: Components/Subnetworks of a Distribution Network for Letter and Parcel Mail

and

Slide 5: Overall Network – Strategic Planning Phase + Tactical Phase

> Facility Location/Allocation + Transportation Network Design for an Overall Distribution Network in Postal Logistics

Location Decisions are always combined with allocation decisions. Allocation decisions are closely related to the transportation network design. Poor modeling (rough approximation) of the transportation (allocation) costs concludes in „non-acceptance“ of the proposed strategic decisions.

Therefore, we propose a hierarchical system of models for Facility Location/Allocation + (Approximative) Transportation Network Design

Aggregated Level Strategic Model

Origins → flow of commodities → Destinations
(postal products) given quantities
Set of potential Sorting Centre Locations

Selection of Sorting Centres, Allocation of Origins/Destinations: Design of allocated geographic regions
Domain knowledge: constraints, transportation modes, fleets, cost-functions

design ↓ — redesign ↑

SC allocation regions ↓ / ↑ costs times quantities (between the strategic model and each of the three models below)

Model for Mail Collection
(strategic/tactical Model)
- Selection of consolidation points
- reallocation of mail sources to sorting centers
- route optimization in mail collection

Model for Long Haul Transportation
(strategic/tactical Model)
- Selection of hubs
- routing commodities through the long haul network
- route optimization

Model for Mail Distribution/Delivery
(strategic/tactical Model)
- ZSP/ZSB Location Optimization (TOPAS-project)
- route optimization in distribution (last mile assumed to be given)

The Analytical Model: Core of the Aggregated Level Strategic Model

Required Data

(1) (OD_i, OD_j) : origin – destination pairs $i,j \in \{1,2,...,N\}$ with assigned aggregated quantities of the postal products q_{ij}

$Q = ((q_{ij}))$ N×N-matrix

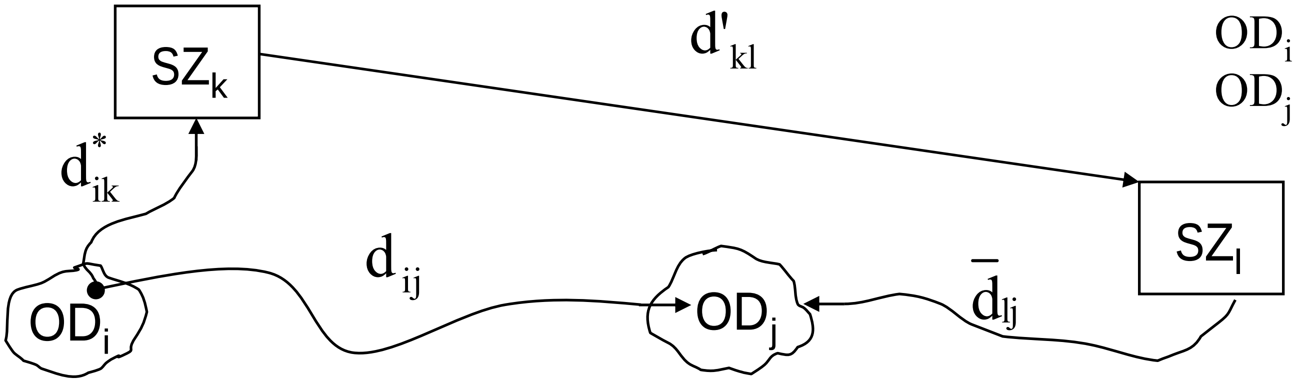
OD_i is an origin – destination object (symmetry: each origin is at the same time a destination) e.g. 5-digit zip code areas

(2) $SC_k : (C_k, L_k)$, C_k – Capacity of the potential facility (Sorting Centre candidate) SC_k

L_k – Location (in geographic coordinates) of SCk

$k = 1,2,...M$

(3) Distances



OD_i allocated to SC_k
OD_j allocated to SC_l

(distances are self explaining)

How to get these data? → see: The Knowledge Model

Decision Variables

- Selection of Sorting Centres from the Candidate Set

$$y_k = \begin{cases} 1, & \text{if } SZ_k \text{ is selected} \\ 0, & \text{otherwise} \end{cases} \quad , \; k = 1,2,\ldots,M$$

- Allocation of Origin – Destination Objects OD_i to the Candidate Set of Sorting Centres

$$z_{ik} = \begin{cases} 1, & \text{if } OD_i \text{ is allocated to } SZ_k \\ 0, & \text{otherwise} \end{cases} \quad , \quad i = 1,2,\ldots,N \ , \ k = 1,2,\ldots,M$$

$\Rightarrow$ M + N·M - binary decision variables

Analytical-type constraints

(1) Number of selected Sorting Centres

Given a required number M* of Sorting Centres, or a maximal number of Sorting Centres we get:

$$\sum_{k=1}^{M} y_k = M^* \quad (\text{or } \leq M^* \text{ respectively}) \ \text{ with } M^* << M \qquad (1)$$

(2) Single Sourcing

OD_i considered as origin and destination at the same time is allocated to exactly one sorting centre candidate SC_k

$$\sum_{k=1}^{M} z_{ik} = 1 \qquad \text{for all} \quad i = 1,2,\ldots,N \qquad (2)$$

(3) Allocation of origins / destinations to selected SC only

$$z_{ik} \leq y_k \qquad \text{for all} \quad k = 1,2,\ldots,M \quad \text{and} \quad i = 1,2,\ldots,N \qquad (3)$$

(If SC_k is not selected, $y_k = 0$, no allocation is possible $z_{ik} = 0$)

(4) Capacity constraints

SC_k has the capacity C_k at the location L_k.

(In this model, we do not allow to open a SC at different capacity levels.)

With $q_i = \sum_{j=1}^{N} q_{ij}$, the *inbound-flow* of SC_k becomes: $\sum_{i=1}^{N} q_i \cdot z_{ik}$

With $q_i^* = \sum_{j=1}^{N} q_{ji}$, the *outbound-flow* of SC_k becomes: $\sum_{i=1}^{N} q_i^* \cdot z_{ik}$

$$\sum_{i=1}^{N} q_i \cdot z_{ik} \le C_k \cdot y_k \quad \text{and} \quad \sum_{i=1}^{N} q_i^* \cdot z_{ik} \le C_k \cdot y_k \qquad (4)$$

for all k = 1,2,...,M

This analytical model (1) - (5) with binary variables y_k, z_{ik} can be identified as a generalized M*-median problem with capacity constraints (or alternatively as capacitated facility location problem with single sourcing and a given number of opened locations M*). Also, an aggregated capacity constraint

$$\sum_{k=1}^{M} C_k \cdot y_k \ge Q \quad , \qquad Q = \sum_{i,j=1}^{N} q_{ij} \qquad (5)$$

can be added.

The objective function:

We use a cost function with fix-costs f_k for „opening“ a sorting centre SC_k and allocation costs. Allocation costs for mail collection and mail distribution subnetworks are proportional z_{ik}, meaning $c_{ik} \cdot z_{ik}$. The costs for the resulting long level transportation networks depend on the used transportation mode.

For two modal (road-air) transportation the costs for the direct transportation relation $SC_k \rightarrow SC_l$ (without hub processes) are:

$$K_{kl} = \begin{cases} K_{kl}^{road} \cdot d'_{kl} \cdot \sum_{i,j=1}^{N} q_{ij} \cdot z_{ik} \cdot z_{jl} & , \text{ if } \ d'_{kl} \leq d' \\ K_{kl}^{air} \cdot d'_{kl} \cdot \sum_{i,j=1}^{N} q_{ij} \cdot z_{ik} \cdot z_{jl} & , \text{ if } \ d'_{kl} > d' \end{cases} \qquad (6)$$

(d'_{kl} distance between SC_k and SC_l, d' maximal distance for road transportation, provided there is a given time window for long haul transportation)

Therefore, we get the objective function:

$$z = \sum_{k=1}^{M} f_k \cdot y_k + \sum_{i=1}^{N} \sum_{k=1}^{M} c_{ik} \cdot z_{ik} + \sum_{k \neq l} d_{kl} \cdot \sum_{i,j=1}^{N} q_{ij} \cdot z_{ik} \cdot z_{jl} \quad \rightarrow \quad \text{Min} \qquad (7)$$

(Alternative analytical models!)

The Domain Knowledge Model

- Knowledge to derive the required data
    - origin-destination objects OD_i , $i = 1,2,\ldots,N$
    - quantities q_{ij}
    - Sorting center candidate set
- Service levels: Cut-off times and time windows
    - Rule-based constraints
- Knowledge about solution techniques, in particular: heuristics and metaheuristics

Origin-destination objects OD_i and quantities from OD_i to OD_j: q_{ij}

OD_i: - 5-digit postal zip code areas: $N \approx 10\,000$ (in Germany)
- Geographic regions assigned to the delivery depots (DD): $N \approx 3\,000$
- Artificial regions defined by pointlike mail sources (mailboxes, business customers, filials)

$\Rightarrow$ The number N determines the size of the resulting MIP considerably.

Quantity to be shipped from OD_i to OD_j: q_{ij} (most critical data)

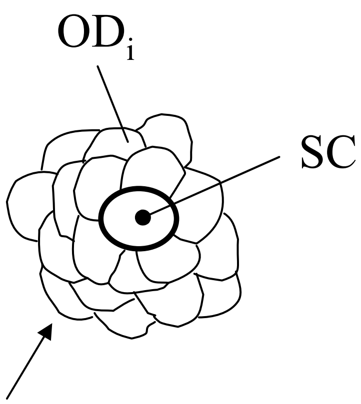
For each of the $N \cdot N$ pairs a quantity has to be determined either by:

- planning statistical experiments or
- analyzing known data
- learning about correlations between properties of the origin-destination objects and relations between each pair of them → regression functions (structure, population, infrastructure of OD_i)

q_{ij} also depends on the weekday considered, and, there is strong seasonal aspect. What if the matrix Q is changing considerably, by ≈ 20...25%?

Sorting Centre Candidate Set

The idea of the analytical model:



SC fixed sorting centre with allocated OD_i forms a geographic region, which determines both: the inbound and the outbound flow

How to determine the SC-candidates?

Clustering of OD_i objects *or* otherwise defined geographic regions: using properties / attributes such as

- character: urban or rural area (or mixed) industrial area, other
- population,
- infrastructure (road-, railnetwork)
- main business customer locations

Estimate of the *inbound* and *outbound* flow of the cluster.

→ Defining (on the basis of this clustering) a geographic region such, that a Sorting Centre should be located within this geographic region in order to serve that region.

→ Place candidates of SC's in the region, which are able to serve that region using knowledge about „good" locations

The Hybrid Knowledge-Based – Analytical Model-Based Approach

1. Define OD_i, $i = 1,...,N$ and an estimate of the quantites q_{ij} for each pair (OD_i, OD_j)

2. Cluster OD_i objects and derive geographical regions such that a Sorting Centre is needed to serve that regions.

3. Place candidates of SC‘s in each of the identified regions (all together M candidates. Use the allocation given by 2. as an initial allocation for the analytical model)

4. Perform Preprocessing by fulfilling the rule-based constraints (infeasible allocations caused by cut-off times and time windows) (not explained here in detail)

5. Solution of the analytical model

6. Replanning on the basis of the results of the analytical model by taking into account additional requests or undefined (hidden) constraints.

Conclusions and Future Work

- It needs to be shown that a hybrid Knowledge-Based and Analytical Model-Based Approach works for a real world application from Postal Logistics.

- Therefore, we develop a prototype of the hierarchical system of models for Facility Location / Allocation + Approx. Transportation Network Design (slide 21)

- The Aggregated Level Strategic Model will be solved by a metaheuristic developed by J. Wollenweber, the Model for Mail Collection was implemented by C. Hempsch and the Model for Mail Distribution / Delivery by C. Hermanns (TOPAS-Model).

- The algorithms and the overall approach will be tested by an artificial postal domain case