MAI0130
Intelligent Modeling Essential to Get Good Results
Chapter 3

Johan Hagenbjörk

Produktionsekonomi - IEI

Linköpings universitet

November 3, 2016

Operations Research theory has developed efficient algorithms for solving some single-objective optimization models that are highly structured.

Real-world decision problems tend to have many complex features making modelling a very hard task.

Wolfram [2002]

the idea of describing behavior in terms of mathematical equations works well where the behavior is fairly simple. It almost inevitably fails whenever the behavior is more complex.

LINKÖPINGS UNIVERSITET

- The model has to be an approximate representation of the real problem.
- The optimal solution must be a good approximation of the real one.

It is essential to model real-world problems intelligently.

- Approximations.
- Relaxations.
- Heuristics.
- Hierarchical modelling techniques.

LINKÖPINGS UNIVERSITET

Murty [2009] presents this topic by three case studies. This lecutre has the same outline.

1. Container Shipping Terminal.
2. Bus Rental Company.
3. Allocating Gates to Flights.
4. Concluding Remarks

- The terminal have storage yards for temporary storage of these containers.
- The storage yard is usually divided into storage blocks which have seven rows with over twenty 20-ft equivalent containers stacked lengthwise.
- Each row can contain stacked containers in up to four or five layers.

There are around 40 major shipping lines in the world and they all take their choice of container terminal very seriously.

Shipping lines rate container terminals largely based on quary crane rate, the average number of containers moved by the quary crane per hour. This might be 20-40.

With several new container terminals opening up, container terminals face growing competition and have to increase productivity to stay in business.

LINKÖPINGS UNIVERSITET

A serious problem in container terminals is that internal trucks get stuck in traffic. Quary crane rates goes down because the quary crane has to wait for

- trucks to take away unloaded import containers.
- trucks to bring export containers to load.

One of the most serious problems in terminal operations is routing the trucks optimally to minimize congestion.

The policy for allocating storage spaces to arriving containers directly influences the routing of container trucks and is therefore an important factor in controlling congestion.

- Variable and uncertain workload due to weather, road, sea conditions.
- Impossible to control the order in which container arrive or their arrival times.
- Storage must be carried out at arrival time without delay.

Container terminals work three 8 h shifts every day. It is therefore convenient to make the planning horizon equal to the length of a shift or less. Two important considerations arise.

- It should be possible to estimate the total workload in the period with reasonable accuracy.
- Total workload should be distributed as uniformly as possible over the planning period.

4 h periods come closest to meet these considerations.

LINKÖPINGS UNIVERSITET

The policy used for allocating storage spaces to arriving containers directly influence the routing of container trucks and is therefore the most important factor in controlling congestion.

Three different models were considered, where the two first were unsuccessful.

Model the decision using binary variables x_{ijkl}, which is 1 if the ith container coming from the jth quary crane is stored in the kth stack of block l.

This turned out to be an inappropriate model because

- This will lead to a huge integer problem which takes a long time to solve.
- The set of occupied storage positions changes every minute.

LINKÖPINGS UNIVERSITET

Model the problem as a two-stage problem.

Stage 1. determines the block where to store each arriving container.

Stage 2. determines the storage location in the block at the time of arrival to the block.

The first step is developing a measure for congestions. Inside the terminal, all roads have the same capacity and all vehicles are very similar. Model the road system as a directed graph $G = (\mathcal{N}, \mathcal{A})$

where

$\mathcal{N}$ is the set of nodes.

$\mathcal{A}$ is the set of arcs.

The decision variables in the flow model are f_{ij}^r = trucks of commodity r passing through arc (i, j) in the planning period for $r = 1$ to t $(i, j) \in \mathcal{A}$.

The problem is a large-scale LP with thousands of constraints but can be solved in a few minutes.

The output provides the storage position for each arriving container and the route for the truck carrying it.

The output from this model turned out to be difficult to implement as drivers resented being told which routes to take.

The model is based on estimated total workload during the entire planning period and thus only works well if the workload is evenly distributed over the planning period.

Let the fill-ratio $f_i(t)$ in block i at time t be

$$f_i(t) = \frac{\text{the number of containers in storage in block } i \text{ at time } t}{\text{total number of storage spaces for containers in block } i}$$

It was observed that the fill-ratio was highly correlated with the truck traffic in and around the block.

Maintaining the fill-ratios in all blocks equal will ensure that the traffic will be evenly distributed.

Stage 1. Determine *container quota number for block i*: That is the number of arriving containers to be stored at block i.

Stage 2. Dispatching Policy for Container Trucks: Determining which block an arriving container should be dispatched in.

Stage 3. Storage-Position Assignment Policy within a Block: How to store arriving blocks to minimize reshuffling.

The decision variable x_i is the number of containers for storage during this planning period, which will be dispatched for storage to block i.

- $a_i =$ the number of stored containers that will remain in block i at the end of this period if no additional container are sent.
- N the number of new containers expected to arrive at the terminal during the period.
- B the number of blocks in the storage yard.
- A the number of storage positions in each block.

The fill ratio in the yard at the end of the period will be

$$F = N + \sum_i \frac{a_i}{AB}$$

If all the fill ratios are equal, they will all be F.

$$\underset{x}{\text{minimize}} \quad \sum_{i=1}^{B}(u_i^+ + u_i^-) \tag{1a}$$

$$\text{subject to} \quad \sum_{i=1}^{B} x_i = N \tag{1b}$$

$$a_i + x_i - A \times F = u_i^+ + u_i^- \tag{1c}$$

$$x_i, u_i^+, u_i^- \geq 0 \tag{1d}$$

(1a) Minimize the deviation from the average fill ratio.

(1b) The sum of all containers stored in each block $i \in B$ must be equal to the number of arriving containers.

(1c) Absolute value of the deviation from the fill ratio.

LINKÖPINGS UNIVERSITET

Let $w_i(t)$ be the number of trucks waiting in block i to be served. This has to be monitored at the storage yard at all times.

Let $x_i^R(t) =$ remaining container spaces at block i at time t.

Dispatch the container at the block i satisfying
$w_i = min\left\{w_j(t) : x_j^R > 0\right\}$

That is, a block with a positive remaining quota that has the least amount of trucks waiting.

For stage three, how to store a container within the block, a heuristic online algorithm to minimize reshuffling was used.

These policies turned out to be highly effective and helped reduce congestion and reduced truck turnaround time by over 20

Two standard approached for modelling the problem led to large scale models that gave very poor results, but the third model let to a small model that was easily solved.

Moral

In order to get good results in real-world applications, it is necessary to model the problem intelligently.

Murty [2009]

Practice has shown that allocation and routing decisions made manually by human operators with long experience are usually nearly optimal, and it is very hard to beat those decisions using a computerized decision-making support system (DMSS).

This is the story about how Murty and Kim [2006] made an intelligent DMSS for a bus company in South Korea which resulted in solutions $10 - 20\%$ more economical than the manual solutions by two experienced workers who were retiring, and replaced by a young worker.

- The bus company rents buses with drivers to customer groups requesting them.
- The bus company rents their two types of buses.
    - 15-seat (5 buses).
    - 45-seat (20 buses).
- Each job specifies a route including the starting time and location, intermediate stops, and a final destination and arrival time.
- The duration of the jobs varies between 0.5 and 20 h.
- The company receives 100 requests of jobs each day. The jobs are classified into
    - large-group jobs, requiring a 45 seat bus.
    - small-group jobs, sufficient to use a 15 seat bus.
- More than one job can never be combined in the same bus.
- The company may itself rent additional buses from outside vendors.
- All job data is available the day before.

LINKÖPINGS UNIVERSITET

- Several succeeding jobs are packed into a job-sequence for a specific driver.
- Buses are kept in depots at two different locations. Driver have to pick a bus up at a depot and return it to a depot at the end of the work-sequence.
- The drivers wage are proportional to the worked hours. Drivers thus want to work a fair amount each day.
- Fatigue can lead to accidents. The company likes to keep work-sequences bellow 12 h. There seems to be no regulation however.
- Some jobs have a duration of over 24 h. However, these jobs allays have many intermediate stops where the driver can rest.

LINKÖPINGS UNIVERSITET

- Several succeeding jobs are packed into a job-sequence for a specific driver.
- Buses are kept in depots at two different locations. Driver have to pick a bus up at a depot and return it to a depot at the end of the work-sequence.
- The drivers wage are proportional to the worked hours. Drivers thus want to work a fair amount each day.
- Fatigue can lead to accidents. The company likes to keep work-sequences bellow 12 h. There seems to be no regulation however.
- Some jobs have a duration of over 24 h. However, these jobs allays have many intermediate stops where the driver can rest.

LINKÖPINGS UNIVERSITET

The problem is a multiobjective problem.

Obj 1. minimize the number of job-sequences, and hence buses used.

Obj 2. minimize the cost of empty load drives of all buses.

Obj 3. keep the percentage of long-duration work-sequences bellow 50% as far as possible. (Internal policy to avoid long-sequence (12h+) jobs two days in a row)

Set a lower bound on a work-sequence to $\delta = 6h$ to keep bus drivers happy.

Several published papers suggest using a 0-1 integer programming model with decision variables of either

- x_{ijkl} equal to 1 if the ith job is included in the jth work-sequence, for which a bus of size k from depot l is allocated.
- x_{jgkl} equal to 1 if jobs j, g are both allocated to the same bus of size k from depot l.

The third objective becomes very hard to model in such a model, and even without it, the model will have a very large number of constraints.

When n is of the order 100 the problem will take several days even for a supercomputer to solve it.

A hierarchical decomposition was used to break up the problem into several stages where objective 2 was split into

- Obj2.1 = cost of empty load drives in-between jobs on the work-sequence.
- Obj.2.2 = cost of empty load drives from and to the depot.

Split the bus allocation into two phases

Phase 1. Consider only the small-group jobs for which the 15-seat bus is suitable. For all work-sequences above the threshold value δ hours, assign 15-seat buses to the extent they are available.

Phase 2. For working-sequences bellow the threshold and all work-sequences requiring 45-seat buses, allocate 45-seat buses.

Minimizing objective 1 requires partitioning the set of n jobs into the smallest possible number of work-sequences. There is an efficient network model for this problem known as *Dilworth's minimal chain decomposition problem*.

Obj 1 and Obj 2 has not taken Obj 3 into account at all.

On each long multiple-job-sequences, the longest arc is deleted from the network and the algorithm is applied to the remaining network. This almost always turned out to satisfy the long working-sequence policy with only a slight increase in the number of work-sequences. It may be repeated until it does satisfy the condition.

LINKÖPINGS UNIVERSITET

Let

- p = number of work-sequences in the final set.
- c_t, d_t = cost of the empty load drive at the beginning and the end of the tth work-sequence.
- e_t cost of renting a bus for the tth work-sequence from outside.
- N_1, N_2 = number of buses available at depot 1 and 2

As e_t typically is much larger than c_t or d_t the number of buses rented from outside will be $(p - N_1 - N_2)^+$. Since only one bus is assigned to each work-sequence, the problem can be modelled as a $3 \times p$ transportation problem.

Relaxations, hierarchical decomposition and heuristics were used to model and analyse the complex problem.

The new worker responsible for making allocations can use this system to make the allocations in a couple of hours each evening. Two retiring workers were replaced and the solutions obtained are $10 - 20\%$ more economical than manual ones.

Methods to solve those models have been studied widely in OR literature.

- Flights use planes of different size.
- Gates are classified into different types depending on size and location.

LINKÖPINGS UNIVERSITET

- Minimize walking distance?
- Minimize walking distance for passengers with short connection time?

No real cost for the airport.

- Minimize penalties from
  - waiting time before allocated a gate.
  - reallocating a gate.

At the same time, all airports in the world claim to apply the policy of first arrived, first assigned.

$$
\begin{aligned}
r = \max \quad & \sum_{j\in J}\sum_{i\in G_j} x_{ij} && \text{(2a)}\\
\text{subject to} \quad & \sum_{i\in G_j} x_{ij} = 1, \quad j \in J_1 && \text{(2b)}\\
& \sum_{i\in G_j} x_{ij} \le 1, \quad j \in J_2 && \text{(2c)}\\
& \sum_{i\in F_i} x_{ij} \le 1, \quad i \in I && \text{(2d)}\\
& x_{ij} \ge 0, j \in J, i \in G_j && \text{(2e)}
\end{aligned}
$$

(2a) Maximize the number of eligible fligt-gate assignments.
(2b) Assign all first preference gates to flights.
(2c) Assign less important flights to second priority gates.
(2d) Flights have max one gate.

LINKÖPINGS UNIVERSITET

If this model turns out to be infeasible, it is an indication that there are not enough eligible gates available. The model then must be relaxed. This happens at high workloads.

The above model does only find the maximum number of gate assignments. It does not try to find the optimal gate allocation.

The optimal gate allocation is found by solving another mathematical model which is also a network flow model which tries to minimize the penalty function.

$$r = \min \quad \sum_{j \in J} \sum_{i \in G_j} c_{ij} x_{ij} \tag{3a}$$

subject to

$$\sum_{i \in G_j} x_{ij} = 1, \quad j \in J_1 \tag{3b}$$

$$\sum_{i \in G_j} x_{ij} \leq 1, \quad j \in J_2 \tag{3c}$$

$$\sum_{i \in F_i} x_{ij} \leq 1, \quad i \in I \tag{3d}$$

$$\sum_{j \in J} \sum_{i \in G_j} c_{ij} x_{ij} = r \tag{3e}$$

$$x_{ij} \geq 0, j \in J, i \in G_j \tag{3f}$$

(3a) Minimize sum of penalties.

(3b) Assign all first preference gates to flights.

(3c) Assign less important flights to second priority gates.

(3d) Flights have max one gate.

(3e) Assign r flights to gates.

Allocating Gates to Flights - Conclusion

- Many models in the literature seems to focus on optimizing objecctives decision makers do not consider important.
- An optimization model is divided into two stages.
- The first stage maximizes the number of flights to which gates can be allocated.
- The second stage allocates the gates to each flight based on minimizing penalties set by decision makers.
- Airport officials stated that top priority should be given to maximize the number of flights that are gated.

LINKÖPINGS UNIVERSITET

The following are important for successfully decision making:

- Looking at the problem from all possible angles and not just one way.
- Constructing an intelligent mathematical model to analyze the problem and to solve it.
- Being very tactful in selling the optimum solution obtained from the model and in implementing it.

LINKÖPINGS UNIVERSITET

Bibliography I

Katta G Murty. *Optimization for decision making*. Springer, 2009.

Katta G Murty and Woo-Je Kim. An i-dmss based on bipartite matching and heuristics for rental bus allocation. In *Intelligent Decision-making Support Systems*, pages 219–235. Springer, 2006.

Stephen Wolfram. *A new kind of science*, volume 5. Wolfram media Champaign, 2002.

LINKÖPINGS UNIVERSITET