Stochastics and Service Network Design

Teodor Gabriel Crainic* Arnt-Gunnar Lium† Stein W. Wallace†

*École des sciences de la gestion, U.Q.À.M., and
Centre de recherche sur les transports, Université de Montréal
C.P. 6128, succ. Centre-ville, Montréal QC, Canada H3C 3J7
theo@crt.umontreal.ca

†Molde University College
P.O. Box 2110
NO-6402 Molde, Norway
Arnt.G.Lium@himolde.no, Stein.W.Wallace@himolde.no

1 Introduction

Much of the transportation of freight is performed by so called consolidation carriers, where one vehicle or convoy serves to move freight for different customers with possibly different initial origins and final destinations (Cordeau, Toth, and Vigo 1998; Crainic (1988, 2000, 2003; Delorme, Roy, and Rousseau 1988; etc.). Freight consolidation transportation is performed by Less-Than-Truckload (LTL) motor carriers, railways, shipping lines, regular and express postal services, etc.

Consolidation transportation carriers are organized as so-called hub-and-spoke networks, where service is offered between a much larger number of origin-destination points than that of direct, origin to destination services operated by the carrier. Lower demands are then moved first to an intermediate point, a hub, to be consolidated together with loads from other customers into vehicles and convoys and moved to other hubs by high frequency and capacity services. More than one consolidation-transfer operation may occur during a trip. The carrier thus operates a series of services, collected in a transportation plan and often accompanied by a schedule. The term *service network design* refers to the construction of a transportation plan that ensures customer satisfaction through efficient and profitable operations and utilization of resources. Main questions addressed thus are: on what routes to provide service? what type of service to use? how often to offer service on each route and according to what schedule? how to route the loads through the physical and service networks?

Several efforts have been directed toward the formulation of tactical models and most yielded *fixed cost, capacitated, multicommodity network design formulations* (Crainic and Rousseau 1986; Roy and Delorme 1989; Crainic and Roy 1988; Powell and Sheffi 1983, 1989; Farvolden and Powell 1994; Kim, Barnhart, and Ware 1999; Armacost, Barnhart, and Ware 2002;

Grünert and Sebastian 2000; Buedenbender, Grünert, and Sebastian 2000; etc.). Integer decision variables are used to represent service selection decisions, while product-specific variables capture the commodity flows. Fixed-costs are associated to the inclusion of services into the plan. Costs that vary with the intensity of service and commodity traffic are associated to the movements of commodities and services. The goal of the formulation is to minimize the total system cost (or to maximize the net profit) under constraints enforcing demand, service, and operation rules and goals. A *space-time* representation of the system operations and decisions is used when schedules are contemplated. Nodes stand for terminals replicated at each period. Services induce temporal arcs replicated at all periods when the corresponding service may start. Holding activities at terminals generate temporal arcs (and costs) between the associated nodes in different periods. The sheer size of the resulting time-dependent network design problem, and the additional constraints usually required by the time dimension, makes this class of problems harder to solve than the static one.

Most of the proposed static and time-dependent service network design formulations are deterministic. This is not to say that the many sources of uncertainty (demand, congestion, weather, etc.) are ignored. On the contrary, the need to integrate stochastic elements into service network design models has been identified as an important research challenge for the field. Not many contributions have been made yet, however, and the issue is addressed through scenario analysis and adjustment of the plan during operations.

A number of challenging questions may then be asked: What are we losing by not including stochasticity into the planning model? Are we losing anything at all? Could we achieve "better" plans, in terms of robustness and less adjustment during operations, if we considered stochastic elements in the formulations (Wallace 2000, Higle and Wallace 2003)? This presentation aims to explore some of these issues and illustrate how the service design depends on the problem definition in terms of stochastics.

Specifically, we consider a simple version of a multi-period service network design formulation and assume that origin to destination demands are random variables with known marginal distributions. In a first step, we calculate deterministic demand estimations and solve the network design problem by using standard mixed-integer programming software. The next step is to formulate a two-stage stochastic network design problem, where the first stage corresponds to finding the design, while the second stage uses the design under varying demands. The stochastic formulation is solved under different assumptions regarding the correlation structure of the demands. The goal is to show how the design varies as one goes from the deterministic formulation to the different stochastic cases.

2 A Simple Service Network Design Model

We build a simple version of a multi-period service network design model for less-than-truckload motor carriers. To build the model, we make the following simplifying assumptions: 1) Homogeneous fleet of capacitated vehicles with no restrictions on how many vehicles are used; 2) Transport movements take one period, while terminal operations are instantaneous (within the period); 3) Demand cannot be delivered later than the due date, but may arrive earlier, in which case, holding costs are paid; 4) One product only; 5) There is a (fixed) cost associated to operating a vehicle (service), but no cost is associated to moving freight (trucks cost the same whether they move loaded or empty); 6) The plan is repeated periodically.

The space-time network is built by repeating the set of nodes (terminals) $\mathcal{N}$ in each of the periods $t = 0, \ldots, T-1$. Each arc (i, j) represents either a service, if $i \neq j$, or a holding activity if $i = j$. A cost c_{ij} is associated to each arc (i, j), equal to the cost of driving a truck from terminal i to j if $i \neq j$, or to the cost of holding a truck at terminal i if $i = j$. A complete network is assumed. For each commodity $k \in \mathcal{K}$, we define its demand $\delta(k)$, origin $o(k)$, destination $d(k)$, and the periods $s(k)$ and $\sigma(k)$ when it becomes available at its origin and must be delivered (at the latest) at its destination, respectively. The truck capacity is denoted M (same units as for demand). The decision variables and model are

$Y^t_{ij}(k)$: Amount of commodity k going from terminal i in period t to terminal j in period $t+1$, for $\{i = o(k), t = s(k), \forall j\} \cup \{j = d(k), t = \sigma(k), \forall i\} \cup \{s(k) < t < \sigma(k), \forall i, j\}$;

X^t_{ij}: Number of trucks from node i in period t to in node j in period $t+1$, $\forall i, j, t$,

$$\min \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{N}} \sum_{t\in T} c_{ij} X^t_{ij} \tag{1}$$

$$t = 0, \ldots, T-1 \qquad \sum_{i\in\mathcal{N}} X^t_{ij} = \sum_{i\in\mathcal{N}} X^{t+1}_{ji} \quad \forall j \in \mathcal{N} \tag{2}$$

$$t = 0, \ldots, T-1 \qquad X^t_{ij} \geq 0 \text{ and integer} \quad \forall i, j \in \mathcal{N} \tag{3}$$

$$(o(k), s(k)) \ : \qquad \sum_{i\in\mathcal{N}} Y^{s(k)}_{o(k)i}(k) = \delta(k) \qquad \forall k \in \mathcal{K} \tag{4}$$

$$(j, s(k)+1) \ : \qquad Y^{s(k)}_{o(k)j} = \sum_{i\in\mathcal{N}} Y^{s(k)+1}_{ji}(k) \qquad \forall j \in \mathcal{N}, \forall k \in \mathcal{K} \tag{5}$$

$$(1 + s(k) < t < \sigma(k) - 1) \ : \qquad \sum_{i\in\mathcal{N}} Y^{t-1}_{ij}(k) = \sum_{i\in\mathcal{N}} Y^t_{ji}(k) \qquad \forall j \in \mathcal{N}, \forall k \in \mathcal{K} \tag{6}$$

$$(j, \sigma(k) - 1) \ : \qquad \sum_{i\in\mathcal{N}} Y^{\sigma(k)-2}_{ij}(k) = Y^{\sigma(k)-1}_{j,d(k)}(k) \qquad \forall j \in \mathcal{N}, \forall k \in \mathcal{K} \tag{7}$$

$$(d(k), \sigma(k)) \ : \qquad \sum_{i\in\mathcal{N}} Y^{\sigma(k)-1}_{i,d(k)}(k) = \delta(k) \qquad \forall k \in \mathcal{K} \tag{8}$$

$$t = 0, \ldots, T-1 \ : \qquad \sum_{k\in\mathcal{K}} Y^t_{ij}(k) \leq M X^t_{ij} \qquad \forall (i, j) \in \mathcal{A} : i \neq j \tag{9}$$

$$t = 0, \ldots, T-1 \ : \qquad Y^t_{ij}(k) \geq 0 \qquad \forall (i, j) \in \mathcal{A}, \forall k \in \mathcal{K} \tag{10}$$

The objective function (1) minimizes the cost associated with the trucks moving between terminals, plus the cost of holding trucks at terminals. Constraints (2) represent conservation of flow for trucks. Constraints (4) and (8) represent conservation of flow at the origin and destination nodes of a commodity, respectively. Equations (5) represents conservation of flow conditions at nodes one period after a commodity has left its origin node. The flow may come only from node $o(k)$ in period $s(k)$, but may go to any node in period $s(k)+1$. Similarly, equations (7) enforce conservation of flow at nodes one period before flow arrives at destination. Equations (6) are the general conservation of flow constraints. They are valid from the second period after a commodity has left its origin and up to two periods before arriving at its destination. Relations (9) are the usual linking and vehicle capacity constraints. Note that commodities can be held at nodes without a truck being present (hence, $i \neq j$).

3 The Stochastic Service Network Design Formulation

Stochastics are described in terms of scenarios $s \in \mathcal{S}$. To each scenario is attached a probability $p^s \geq 0$, with $\sum p^s = 1$. A scenario is $\mathcal{K}$-dimensional, as it contains one demand for each commodity. To indicate the impact of demand variation, the Y flow variables are now indexed by s. The demand for commodity k in scenario s is given by $\delta(k, s)$. The formulation is written in a circular fashion, since we assume that the plan will be repeated. To improve readability, we present constraints according to the difference, diff, in number of time periods, between $s(k)$ and $\sigma(k)$.

$$\min \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{t \in T} c_{ij} X_{ij}^t \tag{11}$$

$$\sum_{i \in \mathcal{N}} X_{ij}^t = \sum_{i \in \mathcal{N}} X_{ji}^{(t+1 \bmod T)} \quad t = 0, \ldots, T-1, \forall j \in \mathcal{N} \tag{12}$$

$$X_{ij}^t \geq 0 \text{ and integer} \quad t = 0, \ldots, T-1, \forall i, j \in \mathcal{N} \tag{13}$$

$$\sum_{i \in \mathcal{N}} Y_{o(k)i}^{s(k)}(k, s) = \delta(k, s), \quad \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{14}$$

$$(\text{diff} \geq 3) \quad : \quad Y_{o(k)j}^{s(k)}(k, s) = \sum_{i \in \mathcal{N}} Y_{ji}^{(s(k)+1) \bmod T}(k, s) \quad \forall j \in \mathcal{N}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{15}$$

$$(\text{diff} \geq 4) \quad : \quad \sum_{i \in \mathcal{N}} Y_{ij}^{(t-1+T) \bmod T}(k, s) = \sum_{i \in \mathcal{N}} Y_{ji}^t(k, s)$$
$$t = 0, \ldots, T-1, \forall j \in \mathcal{N}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{16}$$

$$(\text{diff} \geq 3) \quad : \quad \sum_{i \in \mathcal{N}} Y_{ij}^{(\sigma(k)-2+T) \bmod T}(k, s) = Y_{j,d(k)}^{(\sigma(k)-1+T) \bmod T}(k, s)$$
$$\forall j \in \mathcal{N}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{17}$$

$$\sum_{i \in \mathcal{N}} Y_{i,d(k)}^{(\sigma(k)-1+T) \bmod T}(k, s) = \delta(k, s), \quad \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{18}$$

$$(\text{diff} = 2) \quad : \quad Y_{o(k),j}^{s(k)}(k, s) = Y_{j,d(k)}^{(s(k)+1) \bmod T}(k, s) \quad \forall j \in \mathcal{N}, \forall k \in \mathcal{K}, \forall s \in \mathcal{S} \tag{19}$$

$$\sum_{k \in \mathcal{K}} Y_{ij}^t(k, s) \leq M X_{ij}^t$$
$$\forall (i, j) \in \mathcal{A} : i \neq j, \ t = 0, \ldots, T-1, \forall s \in \mathcal{S} \tag{20}$$

$$Y_{ij}^t(k, s) \geq 0 \quad \forall (i, j) \in \mathcal{A}, \forall k \in \mathcal{K}, t = 0, \ldots, T-1, \forall s \in \mathcal{S} \tag{21}$$

4 Scenario generation and in-sample stability

In our scenario generation process we use a slightly modified version of the method proposed by Høyland, Kaut, and Wallace (2003) and detailed in Kaut and Wallace (2003; see also Høyland

and Wallace 2001 and Dupacova, Consigli, and Wallace 2001). The idea is to construct scenarios with prespecified properties. The authors show that, for their application (finance), the necessary properties are four marginal moments (for each random variable) plus correlations. To find out that these properties will do, they test what is called in- and out-of-sample stability of the optimization problem. Most scenario generation procedures are random in their own right. Hence, when run several times with the same data, they produce different scenario trees. By in-sample stability is understood that whichever of these trees are used in the optimization problem, the optimal objective function value is (approximately) the same. By out-of-sample stability is understood that the true objective function values are also the same for those decisions obtained by the different scenario trees. The latter calculations require sampling (simulation) as a tool.

In this paper we use a slightly revised version of the algorithm by Høyland, Kaut, and Wallace. The change is that instead of four marginal moments as input, we use the density functions of triangular distributions. We then create trees having the first four marginal moments in common with the trianguar distribution, in addition to making sure that the outcomes do not fall outside the supports of the random variables. The latter is not part of the original method. For integer programs, this turns out to be important. Correlations are used as before. We achieve in-sample stability in our tests. This is a minimal requirement for the testing we need to perform. Out-of-sample stability is left for later investigations.

The major goal of this paper is to *understand* how the design depends on the correlations and, in particular, what can go wrong in a deterministic analysis, where correlations by definition are irrelevant. The qualitative understanding is very important if it can be obtained, as it can be a basis for heuristics as well as general problem understanding.

5 REFERENCES

A.P. Armacost and C. Barnhart and K.A. Ware, "Composite Variable Formulations for Express Shipment Service Network Design", *Transportation Science* 36, 1-20 (2002).

K. Buedenbender and T. Grünert and H.-J. Sebastian, "A Hybrid Tabu Search/Branch and Bound Algorithm for the Direct Flight Network Design Problem", *Transportation Science* 34, 364-380 (2000).

J.-F. Cordeau and P. Toth and D. Vigo, "A Survey of Optimization Models for Train Routing and Scheduling", *Transportation Science* 32, 380-404 (1998).

T.G. Crainic, "Rail Tactical Planning: Issues, Models and Tools" in *Freight Transport Planning and Logistics*, L. Bianco and A. La Bella (eds), 463-509, Springer-Verlag, Berlin, 1988.

T.G. Crainic, "Network Design in Freight Transportation", *European Journal of Operational Research* 122, 272-288 (2000).

T.G. Crainic, "Long Haul Freight Transportation" in *Handbook of Transportation Science*, R.W. Hall (ed), 451-516, Kluwer Academic Publishers, Norwell, MA, second edition, 2003.

T.G. Crainic and J.-M. Rousseau, "Multicommodity, Multimode Freight Transportation: A

General Modeling and Algorithmic Framework for the Service Network Design Problem", *Transportation Research B: Methodology* 20B, 225-242 (1986).

T.G. Crainic and J. Roy, "O.R. Tools for Tactical Freight Transportation Planning", *European Journal of Operational Research* 33, 290-297 (1988).

L. Delorme and J. Roy and J.-M. Rousseau, "Motor-Carrier Operation Planning Models: A State of the Art", in *Freight Transport Planning and Logistics*, L. Bianco and A. La Bella (eds), 510-545, Springer-Verlag, Berlin, 1988.

J. Dupačová and G. Consigli and S.W. Wallace, "Scenarios for Multistage Stochastic Programs", *Annals of Operations Research* 100, 25-53 (2001).

J.M. Farvolden and W.B. Powell, "Subgradient Methods for the Service Network Design Problem", *Transportation Science* 28, 256-272 (2001).

T. Grünert and H.-J. Sebastian, "Planning Models for Long-haul Operations of Postal and Express Shipment Companies", *European Journal of Operational Research* 122, 289-309 (2000).

J.L. Higle and S.W. Wallace, "Sensitivity Analysis and Uncertainty in Linear Programming", *Interfaces* 33, 53-60 (2003).

K. Høyland and M. Kaut and S.W. Wallace, "A Heuristic for Moment-matching Scenario Generation", *Computational Optimization and Applications* 24, 169-185 (2003).

K. Høyland and S.W. Wallace, "Generating Scenario Trees for Multistage Decision Problems", *Management Science* 47, 295-307 (2001).

M. Kaut and S.W. Wallace, "Evaluation of Scenario-Generation Methods for Stochastic Programming", *Stochastic Programming E-Print Series, http://www.speps.info* (2003).

D. Kim and C. Barnhart and K. Ware and G. Reinhardt, "Multimodal Express Package Delivery: A Service Network Design Application", *Transportation Science* 33, 391-407 (1999).

W.B. Powell and Y. Sheffi, "The Load-Planning Problem of Motor Carriers: Problem Description and a Proposed Solution Approach", *Transportation Research A: Policy and Practice* 17, 471-480 (1983).

W.B. Powell and Y. Sheffi, "Design and Implementation of an Interactive Optimization System for the Network Design in the Motor Carrier Industry", *Operations Research* 37, 12-29, (1989).

J. Roy and L. Delorme, "NETPLAN: A Network Optimization Model for Tactical Planning in the Less-than-Truckload Motor-Carrier Industry", *INFOR* 27, 22-35 (1989).

S.W. Wallace, "Decision Making Under Uncertainty: Is Sensitivity Analysis of Any Use?", *Operations Research* 48, 20-25 (2000).